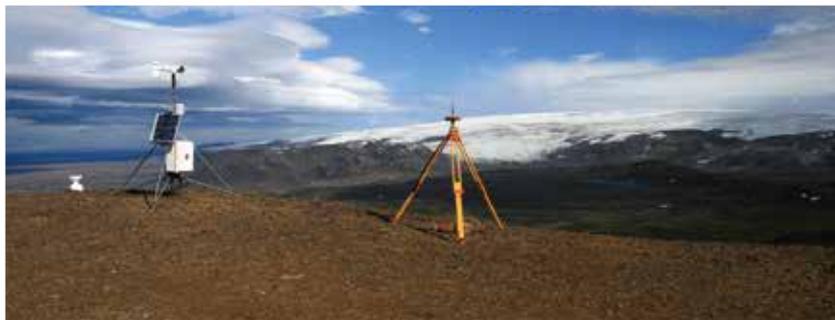


Geomorphological Techniques

online edition
ISSN: 2047-0371



BSG

British Society for Geomorphology

registered charity 1054260



Geomorphological Techniques (Online Edition) (ISSN 2047-0371)

Geomorphological Techniques is a free online resource, outlining the current methods utilised in the particular specialist area of interest. The aim is to provide an up-to-date, evolving resource for geomorphologists to use and contribute towards. All articles are peer reviewed.

Contribute your expertise

Are you an expert on a particular geomorphological technique? Have you just written a methods chapter for your PhD and would like to share it with the world? Would you like to help others to avoid the pitfalls you have uncovered, or get a head start at applying a method to further geomorphic research?

The BSG are looking for authors to volunteer articles to contribute to an updated, online version of Geomorphological Techniques. Each article should be approximately 2000 – 3000 words in length, outlining the current methods utilised in your specialist area and will be peer reviewed by at least 2 specialists in the field. Honoraria of £200 are available for BSG members that author articles.

Please contact the Publications Sub-Committee Secretary, via our website contact us page, if you are interested in contributing. The current table of contents is listed below, articles are allocated on a first come basis. Items in **bold** are completed sections which are now downloadable.

How to reference:

These are edited book chapters and so should be referenced to the chapter author and using the year listed on the pdf, the book details are as follows:

In: Clarke, L.E. & Nield, J.M. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology, London. ISSN: 2047-0371

So for example, for Section 2.1.3 would be:

Young, E.J. 2012. Section 2.1.3: dGPS. In: Clarke, L.E & Nield, J.M. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology; London, UK. ISSN: 2047-0371.

Contents

1 Composition of Earth Materials (generic, not environment-specific)

1.1 Clast properties

1.1.1 Particle size analysis

1.1.2 Particle form analysis (shape, roundness etc.)

1.1.3 Particle texture (SEM etc.)

1.2 Bulk properties of soils and sediments

1.2.1 Porosity

1.2.2 Permeability

1.2.3 Moisture content and suction

1.3 Strength of materials

1.3.1 Shear stress

1.3.2 Rock hardness

1.4 Mineralogical and chemical composition

1.4.1 Environmental magnetism

1.4.2 Major element geochemistry (e.g. XRF, AAS)

1.4.3 Minor and REE geochemistry (e.g. ICP)

1.4.4 Mineral crystallography XRD

1.4.5 Mineral inference (FTIR, NIR, VIS and UV spectroscopy)

1.5 Form and Structure of Sediment Bodies

1.5.1 Describing and logging sedimentary sequences

1.5.2 Fabric and structure of clastic sediments

1.5.3 Thin section micromorphology

1.5.4 Imaging sediment structures

1.5.5 GPR

1.5.6 Thin section micromorphology

1.6 Rock Properties

1.6.1 Thin section petrography

1.6.2 Scanning Electron Microscopy (SEM)

1.6.3 Environmental Electron Microscopy (ESEM)

2 Topographic and Spatial Analysis

2.1 Direct acquisition of elevation data

2.1.1 Surveying basics

2.1.2 Total station

2.1.3 dGPS

2.1.4 Airborne LiDAR

2.1.5 Terrestrial laser scanning

2.1.6 Bathymetric methods

2.1.7 UAVs

2.2 Photogrammetric techniques

2.2.1 Aerial photogrammetry

2.2.2 Structure for Motion (SfM)

2.3 Digital Elevation/Terrain Models

2.3.1 Creating DEMs from survey data

2.3.2 DEMs of difference

2.4 Geospatial analysis

2.4.1 GIS platforms and tools

2.4.2 Terrain analysis and landform recognition

2.4.3 Network delimitation and analysis

2.4.4 Geospatial statistics for analysis of form

2.5 Sediment fingerprinting

2.6 Geomorphological mapping

3 Processes, Forms and Materials in Specific Environments

3.1 Aeolian

3.1.1 Local form (cross-section and slope)

3.1.2 Plan geometry

3.1.3 Velocity and flow properties (e.g. anemometers, towers etc.)

3.1.4 Sediment transport (saltation probes, sediment traps, etc.)

3.2 Coastal

3.2.1 Beach morphology (physical and remote sensing techniques)

3.2.2 Cliff and shore platform geometry

3.2.3 Surf-zone hydrodynamics

3.2.4 Swash zone sediment transport

3.2.5 Swash zone sediment transport tracers

3.2.6 Field measurement & remote sensing of the swash zone

3.2.7 Coastal cell sediment balance

3.3 Fluvial

3.3.1 Local form (cross-section and slope)

3.3.2 Plan geometry

3.3.3 Bed material sampling

3.3.4 Velocity and flow properties (PIV)

3.3.5 Discharge estimation and stream gauging

3.3.6 Suspended sediment sampling

3.3.7 Bed load sediment (traps, samplers, movement detectors)

3.3.8 Tracing fluvial sediments (bed load tracing techniques)

3.3.9 Tracing particles and organisms in rivers

3.3.10 River bed scour and fill (scour chains, bed disturbance)

3.3.11 Bank Erosion (PEEPs etc.)

3.4 Glacial

3.4.1 Sampling and describing ice

3.4.2 Meltwater sampling and analysis

3.4.3 Tracer investigations

3.4.4 Borehole drilling and instrumentation

3.4.5 Glacier movement

3.4.6 Glacier energy balance

3.4.7 Basal processes

3.4.8 Glacial sedimentology

3.4.9 Glacier reconstruction (e.g. ELA, palaeoclimate reconstructions etc.)

3.5 Hill slope/Mass movement

3.5.1 Sensors

3.5.2 Sediment tracing

3.6 Karst

3.6.1 Landform classification techniques

3.6.2 Local form measurements

3.6.3 Dissolutional denudation rates

3.6.4 Tracer techniques

3.7 Lacustrine

3.7.1 Contemporary environment

3.7.2 Recent sediment accumulation (sediment water interface)

3.7.3 Sediment provenance (catchment – sink linkages)

3.8 Periglacial

3.8.1 Environmental factors

3.8.2 Sediment transport

3.9 Dryland

3.9.1 Erosion monitoring

3.9.2 Dryland slope processes

3.9.3 Monitoring ephemeral streams

3.10 Marine

3.10.1 Shallow water morphology

3.10.2 Deep water morphology

3.10.3 Marine remote sensing techniques

3.11 Hyporheic zone

3.11.1 In-Situ sampling

3.11.2 Coring

3.11.3 Tracer techniques

4 Long-term Environmental Change (dating techniques, etc.)

4.1 Environmental reconstruction techniques (geomorphology relevant aspects)

4.1.1 Coring methods

4.1.2 Diatoms

4.1.3 Chironomids

4.1.4 Pollen

4.1.5 Marine organisms (forams)

4.1.6 Isotope analysis

4.1.7 Core scanning

4.2 Chronological techniques

4.2.1 General context (absolute, relative, radiometric, errors)

4.2.2 Radiocarbon

4.2.3 Short isotopes (e.g. Pb210, Cs137)

4.2.4 U series

4.2.5 Amino acid racemisation

4.2.6 Luminescence

4.2.7 Lichenometry

4.2.8 Dendro

4.2.9 Archaeo and Palaeomagnetism

4.2.10 Cosmogenics

4.2.11 Tephra

5 Modelling Geomorphic Systems

5.1 Generic and conceptual

5.2 Numerical modelling

5.3 Physical modelling

5.4 Statistical modelling

5.5 Evaluating and testing models

5.6 Environment Specific Models

5.6.1 Weathering

5.6.2 Aeolian

5.6.3 Coastal

5.6.4 Fluvial

5.6.5 Glacial

5.6.6 Hill slope/Mass movement

5.6.7 Karst

5.6.8 Lacustrine

5.6.9 Periglacial

5.6.10 Dryland

5.6.11 Marine

5.6.12 Landscape evolution

The Use of Equivalent Quartz Size and Settling Tube Apparatus to Fractionate Soil Aggregates by Settling Velocity

Yaxian Hu¹, Wolfgang Fister¹, Hans-Rudolf Rüegg¹, Peter A. Kinnell², Nikolaus J. Kuhn¹

¹ Physical Geography and Environmental Change Research Group, Department of Environmental Sciences, University of Basel, Switzerland (yaxian.hu@unibas.ch)

² Institute of Applied Ecology, University of Canberra, Australia



ABSTRACT: In a given layer of surface runoff, particle transport distance declines with increasing settling velocity. Settling velocity itself is determined by the size, density and shape of the particles. For sediment composed of aggregates, settling velocity does not only vary due to texture, but also due to aggregation, aggregate size and stability. Therefore, aggregation can strongly affect the transport distance of the sediment and the substance specific redistribution of the eroded material, such as organic matter. Understanding the effect of aggregation, for example, on redistribution of eroded organic matter is therefore essential for understanding local, regional and global carbon cycles. To capture and establish the relationship between aggregation, settling velocity and aggregate specific organic matter content, a settling tube apparatus, based on a previous design, was constructed and applied to fractionate soils by water stable aggregate size classes. To illustrate the effect of aggregation on settling velocity, the results were compared with mineral grain sizes after ultrasound dispersion. Five settling velocity classes were distinguished based on the Equivalent Quartz Size (EQS) of particles $\geq 250 \mu\text{m}$, 125 to $250 \mu\text{m}$, 63 to $125 \mu\text{m}$, 32 to $63 \mu\text{m}$, and $\leq 32 \mu\text{m}$. Fractionation of a silty loam by settling tube illustrates that aggregation strongly affects settling velocities and should be considered in erosion models, as opposed to the texture of mineral grains. An analysis of sediment organic matter in the five settling velocity classes also showed that settling velocity is a suitable parameter to physically connect the redistribution of eroded soil organic matter to overland flow transport processes.

KEYWORDS: settling tube apparatus, settling velocity, transport distance, aggregate fractionation

Introduction

Soil particles displaced by erosion experience selective deposition along their flow paths across watersheds (Walling, 1983). Understanding the effect of this selective deposition on the redistribution of particle-bound substances (e.g. soil organic carbon, phosphorous or other contaminants) within watersheds requires a discrimination of particles and their properties by their respective transport distances. The transport distances of displaced soil particles are related to their settling velocities (Dietrich, 1982; Kinnell, 2001; Kinnell, 2005). For eroded soil particles composed of aggregates, settling velocities generally do

not correspond to the average or median mineral grain size, because aggregates differ in size, density and shape from mineral grains (Johnson *et al.*, 1996; Tromp-van Meerveld *et al.*, 2008). The average or median grain size can be the same for a range of soils, but the aggregate size distribution can differ, e.g. when clay enhances the formation of aggregates. The distribution of settling velocities therefore can provide more accurate information on the quality and behavior of eroded and aggregated sediment than just texture (Loch, 2001). The distribution of settling velocities based on grain size classes has already been included into some erosion / deposition models (Morgan *et al.*, 1998; van Oost *et al.*,

2004; Fiener *et al.*, 2008; reviewed by Aksoy and Kavvas, 2005). However, inconsistencies such as over-prediction of clay in sediment fractions or under-prediction of sand and silt in sediment samples (Beuselinck *et al.*, 1999; van Oost *et al.*, 2004) are often present in their results, because soil particles are mostly eroded in the form of aggregates rather than as mineral grains (reviewed by Walling, 1988; Slattery and Burt, 1997; Beuselinck *et al.*, 2000). Aggregation potentially increases settling velocities and reduces transport distances. As a consequence, aggregation can lead to aggregate specific, rather than mineral grain specific, distribution of particle-bound substances across a landscape by selective deposition (Kuhn, 2007; Kuhn and Armstrong, 2012). The settling velocities of aggregates are therefore crucial to determine the effect of erosion on redistribution of substances (such as eroded soil organic carbon, phosphorous, nitrogen or metals) across landscapes, as well as their delivery into aquatic systems. By further identifying the lability of the eroded soil organic carbon, and quantifying the relative proportion mobilised into or out of different ecosystems, it also can substantially improve our understanding of the role of erosion / deposition on global carbon cycling.

The settling velocity of mineral particles is determined by their size, density and shape (Dietrich, 1982). For aggregated soils, their irregular shape, porosity, permeability, interaction with organic matter of low density, aggregation, and the relative fragility of wet aggregates (Dietrich, 1982; Le Bissonnais *et al.*, 1989; Johnson *et al.*, 1996; Tromp-van Meerveld *et al.*, 2008) can all affect their settling velocities. Therefore, a conceptual approach based on Equivalent Quartz Sizes (EQS), modified from the equivalent sand size used by Loch (2001), is developed to address the effect of aggregation on the potential redistribution of eroded soil organic matter across a hill-slope. EQS represents the diameter of a spherical quartz particle that would fall with the same velocity as the aggregated particle for which fall velocity is measured (Loch, 2001). Therefore, EQS represents an integrated index to indicate the settling behaviour rather than to represent the specific size of the soil particles. In this manner, our current understanding of the effects of mineral grain size on sediment behavior can be applied to aggregated sediment particles based on the concept of

EQS. Although the accurate size of aggregated particles needs to be validated by field data, the accuracy of soil erosion models can also be largely improved by applying the distribution of settling velocities based EQS compared to grain size distribution.

Use of settling tubes to fractionate sediment particles

The settling tube (column) is a traditional technique used to measure the settling characteristics of aquatic solids (Droppo *et al.*, 1997; Wong and Piedrahita, 2000; Rex and Petticrew, 2006), but the settling tubes used in river and marine environment are often short and with small openings. Consequently, they are unable to allow coarse particles to pass through, so they cannot be directly applied to fractionate sediment that is often in the form of aggregates. Settling tubes, such as the 20 cm long example used in Johnson *et al.* (1996), cannot be used to fractionate the aggregated sediments either, because such a short settling distance is not capable of accurately distinguishing the velocities of fast settling particles.

In order to fractionate aggregated soils, Hairsine and McTainsh (1986) designed a top-entry settling tube apparatus (The "Griffith Tube"), which was adapted from the "Siltometer" developed by Puri (1934). In this design, soil samples were introduced into a 200 cm long vertical tube from the top by an injection barrel. After falling through the static water column by gravity, soil fractions were collected over predetermined time intervals into sampling dishes situated in a turntable under the widely-open bottom of the tube. This design was then improved by Kinnell and McLachlan (1988) using a more reliable injection barrel, and further by Loch (2001), who employed an electric motor to raise the tube and rotate the turntable. Unlike other physical fractionation methods, for instance, wet and dry sieving (Cambardella and Elliott, 1993; Christensen, 2001), where aggregates would inevitably experience abrasion, settling through a water column preserves fragile aggregates. However, such a technique has not been widely implemented, because the lack of details in describing the existing settling tube apparatus makes it very difficult to reconstruct one without detailed knowledge of their design specifications.

Such information could only be obtained by personally contacting the authors, which is often not possible. In particular, measurements linked to modeling the redistribution of organic carbon and their implications on the carbon cycle are missing (Kuhn, 2013). It is also noteworthy that soil particles depositing through a column of static water neglects the potential effects of, for instance, flow turbulence during transport processes. Other information (e.g. topography, and flow velocity) is required, therefore, in order to further calibrate the realistic transport distances of eroded soil particles.

Settling tube apparatus developed by Basel University

The settling tube apparatus built at Physical Geography and Environmental Change Research Group from Basel University consists of four components (Figure 1a): the settling tube, through which the soil sample settles (Figure 1b); the injection device, by which the soil sample is introduced into the tube (Figure 2); the turntable, within which the fractionated subsamples are collected (Figure 3); and the control panel, which allows an operator to control the rotational speed and resting/moving intervals of the turntable (Figure 4).

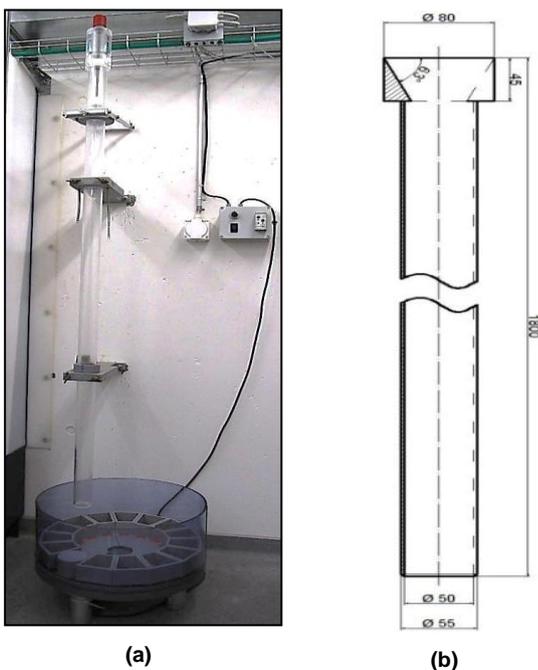


Figure 1. (a) Overview of the complete setup of the Basel University settling tube apparatus; (b) The settling tube. Measurement units in mm.

The settling tube

The settling tube is made of transparent PVC, has a length of 180 cm and an internal diameter of 5 cm (Figure 1b). The tube can hold approx. 4 liters of water, through which the soil sample is free to settle. In most cases, the soil particles are smaller than 2 mm in diameter, so the diameter ratio of the tube (50 mm) to a particle (< 2 mm) is greater than 25 to 1. According to Loch (2001), such a ratio largely eliminates concerns associated with edge effects and the variability introduced by wall effects is expected to be < 10%.

The injection device

An injection device is used to insert the soil sample into the top of the settling tube. It consists of a central chamber and a ca. 30 cm long metal rod connected to a Teflon cone at the lower end, a Teflon piston in the middle and a rubber plug attached to a handle at the upper end (Figure 2, based on the design of Kinnell and McLachlan, 1988).

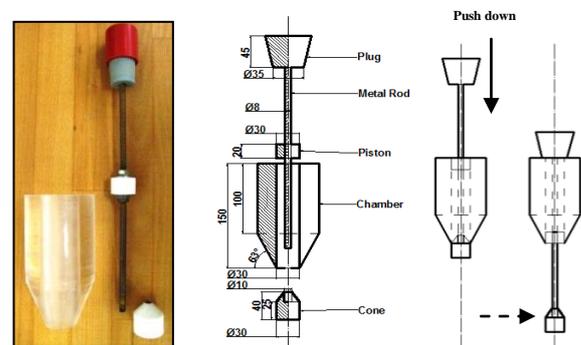


Figure 2. The injection device used in the Basel University settling tube apparatus. Measurement units in mm.

The soil sample is placed into the central chamber before injection. The capacity of the injection chamber (ca. 80 cm³) limits the mass of dry soil to 25 g. This leads to a soil concentration of approximately 6 g L⁻¹ in the water column. Following Loch (2001), the concerns associated with particle interactions during settling are therefore minimal. The metal rod passing through the chamber opens the Teflon cone at bottom of the chamber, releasing the soil sample while the piston and plug keep the chamber sealed to prevent water flowing out at the bottom of the

pipe. Kinnell and McLachlan (1988) used a pin inserted into the rod to prevent the piston from moving downwards before release. In our injection device the deformation of the Teflon is used to seal the chamber and prevent the movement of the piston (Figure 2). This design seals the chamber more effectively; however, it requires a much greater force to open it. A slow opening of the chamber can lead to inaccurate settling times. It is also noteworthy that the cone frequently represents an obstacle in the pathway of falling particles and small amounts (< 1 g) have been observed on the surface of the cone. Further improvements on the design of the injection device are required in order to solve this problem.

The turntable

The turntable is placed under the settling tube and is used to collect the soil fractions that settle out of the tube. It consists of a circular tank (Figure 3a, PVC transparent), and a set of sampling dishes (Figure 3b, PVC grey). The circular tank is 50 cm in diameter, 20 cm deep, and has a volume of 40 L. The net volume of each sampling dish is ca. 290 cm³. When settling, the water level in the tank must be higher than the bottom opening of the tube to prevent the water from flowing out of the tube. The turntable tank rests on a layer of plastic ball bearings placed in a tray beneath the tank. This tray rests on three pillars (Figure 3c). An electric motor, affixed to the pillars, enables a timed and stepwise rotation of the turntable and thus places each respective sampling dish precisely underneath the settling tube, e.g. at time intervals corresponding to the settling times of the EQS. Where motor installation is not available, manual operation to replace the sampling dishes is also feasible.

The control panel

A plug-in time delay relay (© Comat, RS 122-H) is used to control the rotational speed and resting/moving intervals of the turntable (Figure 4). The control panel primarily consists of three parts: the main switch, the speed-control knob, and the interval-control buttons.

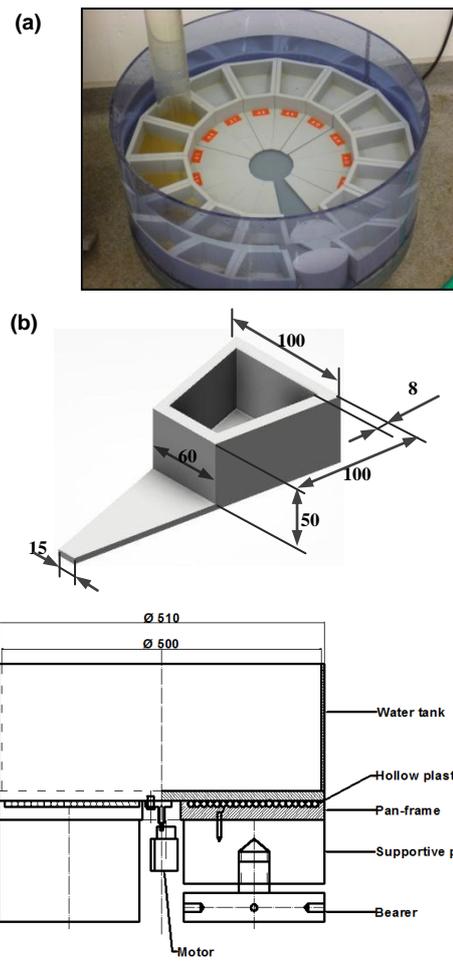


Figure 3. The turntable (a), the sampling dish (b), and the supporting frame. Measurement units in mm. (c) of the Basel University settling tube apparatus.

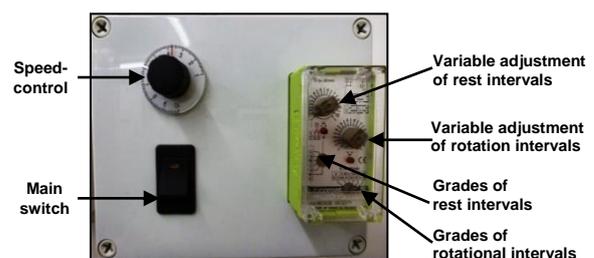


Figure 4. The control panel of the Basel University settling tube apparatus

Potential transport distance of eroded organic carbon based on texture and aggregation

Soil selection and preparation

A silty loam from Möhlin, in northwest Switzerland (47° 33' N, 7° 50' E) was used to compare the differences between the potential transport distance of eroded organic carbon predicted based on soil texture and

that predicted by aggregate fractionation. The soil was sampled from Bäumlhof Farm with a wheat-grass-maize rotation in August 2011. Sampling directly from the field, rather than from depositional sites after a certain extent of preferential transportation, provides an opportunity to evaluate the likely transport distance of all classes of eroded particles. Its total organic carbon concentration is 10.9 mg g⁻¹, and the aggregate stability (based on Nimmo and Perkins, 2002) is 66.85 %. This degree of aggregation and organic carbon content were considered appropriate to investigate the effects of aggregation onto the potential redistribution of eroded soil organic carbon by deposition across the landscape.

Calculation of Equivalent Quartz Size

Stokes' Law covers the range of the mineral grain sizes dominating the silty loam used for this study. Their terminal settling velocities can be calculated by:

$$V = \frac{h}{t} = \frac{d^2 g (D_s - D_f)}{18\eta} \quad (1)$$

Where: V = settling velocity (m·s⁻¹), h = settling distance = 180 cm with this settling tube apparatus, t = settling time (s), d = diameter of settling particle (mm), g = gravitational force = 9.81 N·kg⁻¹, η = viscosity of water at 20° C = 1 × 10⁻³ Ns·m⁻², D_s = average density of the solid particles, for most soils = 2.65 × 10³ kg·m⁻³, D_f = density of water = 1.0 × 10³ kg·m⁻³.

The use of Stokes' Law to calculate EQS is, in the strictest sense, limited to particles < 0.07 mm (Rubey, 1933). For soils dominated by larger mineral grains, different relationships should be used (e.g. Ferguson and Church, 2004; Wu and Wang, 2006). Five size fractions were selected according to their likely transport distance once eroded (Starr *et al.*, 2000): ≥ 250 μm, 125 to 250 μm, 63 to 125 μm, 32 to 63 μm, ≤ 32 μm (Table 1).

Soil fractionation by settling tube and wet sieving

The soil samples were dried at 40 °C until constant weight was achieved and then gently dry-sieved with an 8 mm sieve to avoid over-sized clods. Prior to settling / wet-sieving, 25 g of dry soil were immersed into

50 ml tap water for 15 min. This fast-wetting emphasizes slaking and slight clay dispersion and simulates the destruction of aggregates during an erosion event (Le Bissonnais, 1996). For all tests, tap water was used. The electric conductivity of tap water was 2220 μs·cm⁻¹, which was five times higher than the rainwater in Basel (462 μs·cm⁻¹). In general, increased electric conductivity of tap water generally enhances dispersion during rainfall simulation tests (Borselli *et al.*, 2001). A comparative aggregate stability test (Wet Sieving Apparatus, Eijkelkamp, Netherlands) using tap water and rainwater from Basel had shown that tap water only had a minor effect on aggregates greater than 250 μm after 20 min of continuous up-and-down movement (67.24 % in rainwater, 73.59 % in tap water). Therefore, the use of tap water was considered to be acceptable.

*Table 1. Settling times and velocities of the Möhlin silty loam, based on the Equivalent Quartz Size (EQS) classes, and the likely transport distance of soil particles once eroded and based on the conceptual function developed by Starr *et al.*, 2000. The settling distance is 180 cm.*

EQS (μm)	Settling velocity (m·s ⁻¹)	Settling time (s)	Likely transport distance
> 250	> 4.5 × 10 ⁻²	< 40	Deposited across landscapes
125 - 250	1.5 × 10 ⁻² - 4.5 × 10 ⁻²	40 - 120	
63 - 125	3.0 × 10 ⁻³ - 1.5 × 10 ⁻²	120 - 600	
32 - 63	1.0 × 10 ⁻³ - 3.0 × 10 ⁻³	600 - 1800	Possibly transferred into rivers
< 32	< 1.0 × 10 ⁻³	> 1800	Likely transferred into rivers

A 25 g of fast-wetted soil sample was fractionated using the Basel University settling tube apparatus into five settling velocity classes (Table 1). A typical settling pattern of soil particles in the water column is shown in Figure 5. The finest particles correspond to those that remain in suspension after 1800 s of settling (i.e. EQS < 32 μm). Fractionated samples were dried at 40°C and dry weights as well as total organic carbon concentration (by Leco RC 612 at 550°C) were measured.



Figure 5. A typical settling pattern of soil particles through the water column: coarse particles settle fastest, while the fine particles stay suspension at upper part.

A second 25 g of fast-wetted soil sample was subjected to ultrasound using a Sonifier 250 from Branson, USA. The energy dissipated in the water/soil suspension was ca. $60 \text{ J}\cdot\text{ml}^{-1}$ (i.e. Energy = output power 70 W \times time 85 s / suspension volume 100 ml) (North, 1976). According to Kaiser *et al.* (2012), although the aggregates were probably not thoroughly dispersed at this level of energy, the coarse mineral and organic particles ($> 250 \mu\text{m}$) were prone to be damaged if higher energy than $60 \text{ J}\cdot\text{ml}^{-1}$ was further applied. The dispersion energy level of $60 \text{ J}\cdot\text{ml}^{-1}$ was thus considered to be satisfactory, in the context of distinguishing the size distribution of aggregates from grains. The dispersed fractions were then wet-sieved into the five size classes corresponding to the five EQS classes used for the fractionation by settling tube. The weights and total organic carbon concentrations of each class were then measured in the same way as for the settling tube fractionated samples.

Effect of aggregation on settling velocity

The results of the two fractionation approaches are shown in Figure 6. The effect of aggregation on settling velocity is pronounced: 68.61 % of the aggregated soil behaved like particles of EQS greater than $63 \mu\text{m}$ (Figure 6a). The mineral particle size

distribution, on the other hand, shows that 89.65 % of soil grains were smaller than $63 \mu\text{m}$ (Figure 6a). This difference between proportion of EQS and corresponding mineral grain size classes illustrates that aggregation has a great effect on the particle settling velocity of the silty loam tested in this study.

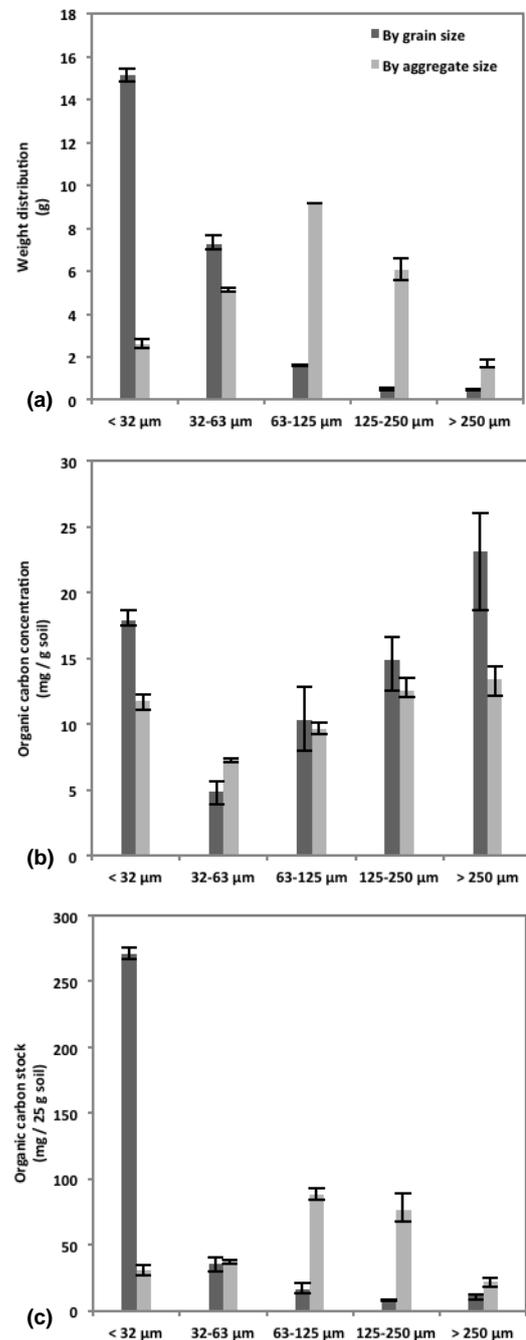


Figure 6. (a) Weight distribution, (b) organic carbon concentration distribution, and (c) organic carbon stock distribution of the Möhlin silty loam across aggregate size classes fractionated by the Basel University settling tube apparatus and across grain size classes dispersed by ultrasound. Error bars indicate the range of minimum and maximum values. $n=3$.

The relevance of this finding is further illustrated by the distribution of total organic carbon in aggregates and mineral grains. The distribution of total organic carbon concentration follows a similar pattern for both grain size and aggregate size classes (Figure 6b). However, multiplying the organic carbon concentration of each size class with its weight (Figure 6c) shows that 73 % of the organic carbon stock is contained in particles > 63 µm, while 79 % of the organic carbon stock was associated with grains < 32 µm. This implies that aggregation strongly affects the potential transport distance of the eroded organic carbon. Basically, the amount of deposition across the landscape would be tripled as the soil texture suggests. By contrast, the exportation of eroded organic carbon to watercourses would be reduced to a third. In consequence, the effect of aggregation on transport distances would fundamentally change our perspective on the environmental impact of eroded organic carbon as well as other nutrients and contaminants.

Conclusion and Implication

The settling tube fractionation provides settling velocity as a 'tool' to physically connect the redistribution of eroded soil organic carbon in an agricultural landscape with a soil transportability parameter used in current erosion models. While the initial test reported in this paper is very limited in its applicability to real erosion events, the results prove that a more accurate settling velocity of aggregated soil particles can be measured based on Equivalent Quartz Size than solely based on grain size distribution. The results also indicate that selective transport of aggregated sediment potentially has a great impact on the redistribution of eroded organic carbon in terrestrial ecosystems and its delivery to aquatic ecosystems. Such differentiation of sediment movement could, in turn, be highly significant for the effect of soil erosion on the global carbon cycle (Kuhn *et al.*, 2009). As a consequence of the results of the settling velocity fractionation procedure presented in this study, we conclude that the settling tube apparatus can be further applied to determine the realistic settling velocities of eroded soil generated in the field. The application of a settling tube to fractionate sediment particles according to settling

velocities makes a significant contribution to our understanding of local, regional and global geochemical fluxes within terrestrial ecosystems, and also of their interaction with atmospheric and aquatic systems.

Acknowledgement

We gratefully acknowledge the financial support granted by Chinese Scholarship Council and the University of Basel. The contributions of Marianne Caroni, Ruth Strunk and Mathias Würsch in carrying out the laboratory measurements are also thankfully acknowledged. The draft of the manuscript was substantially improved by comments from Philip Greenwood.

References

- Aksoy H, Kavvas ML. 2005. A review of hillslope and watershed scale erosion and sediment transport models. *Catena* **64**: 247-271.
- Beuselinck L, Govers G, Steegen A, Quine TA. 1999. Sediment transport by overland flow over an area of net deposition. *Hydrological Processes* **13**: 2769-2782.
- Beuselinck L, Steegen A, Govers G, Nachtergaele J, Takken I, Poesen J. 2000. Characteristics of sediment deposits formed by intense rainfall events in small catchments in the Belgian Loam Belt. *Geomorphology* **32**: 69-82.
- Borselli L, Torri D, Poesen J, Salvador Sanchis P. 2001. Effects of water quality on infiltration, runoff and interrill erosion processes during simulated rainfall. *Earth Surface Processes and Landforms* **26**: 329-342.
- Cambardella CA, Elliott ET. 1993. Methods for physical separation and characterization of soil organic matter fractions. *Geoderma* **56**: 449-457.
- Christensen BT. 2001. Physical fractionation of soil and structural and functional complexity in organic matter turnover. *European Journal of Soil Science* **52**: 345-353.
- Dietrich WE. 1982. Settling velocity of natural particles. *Water resources research* **18**: 1615-1626.

- Droppo IG, Leppard GG, Flannigan DT, Liss SN. 1997. The freshwater floc: A functional relationship of water and organic and inorganic floc constituents affecting suspended sediment properties. *Water Air Soil Pollut.* **99**: 43–53, doi:10.1007/BF02406843.
- Ferguson RI, Church M. 2004. A simple universal equation for grain settling velocity. *Journal of sedimentary Research* **74 (6)**: 933-937.
- Fiener P, Govers G, Van Oost K. 2008. Evaluation of a dynamic multi-class sediment transport model in a catchment under soil-conservation agriculture. *Earth Surface Processes and Landforms* **33**: 1639-1660.
- Hairsine PB, McTainsh, G. 1986. The Giffith Tube: A simple settling tube for the measurement of settling velocity of soil aggregates. School of Australian Environmental Studies, Griffith University, AES Working Paper 3/86.
- Johnson CJ, Li XY, Loan BE. 1996. Settling velocities of fractal aggregates. *Environ. Sci. Technol.* **30**: 1911-1918.
- Kaiser M, Berhe AA, Sommer M, Kleber M. 2012. Application of ultrasound to disperse soil aggregates of high mechanical stability. *J. Plant Nutr. Soil Sci.* **175**: 521-526.
- Kinnell PIA. 2001. Particle Travel distances and bed and sediment compositions associated with rain-impacted flows. *Earth Surface Processes and Landforms* **26**: 749-758.
- Kinnell PIA. 2005. Raindrop-impact-induced erosion processes and prediction: a review. *Hydrology Process* **19**: 2815-2844.
- Kinnell, PIA, McLachlan C. 1988. An Injection Barrel for the Top Entry Sedimentation Tube. Technical memorandum 43/1988, CSIRO Division Soils.
- Kuhn NJ, Armstrong EK. 2012. Erosion of organic matter from sandy soils: solving the mass balance. *Catena* **98**: 87-95.
- Kuhn NJ, Hoffman T, Schwanghart W, Dotterweich M. 2009. Agricultural soil erosion and global carbon cycle: controversy over? *Earth Surface Processes and Landforms* **34**: 1033-1038.
- Kuhn NJ. 2007. Erodibility of soil and organic matter: independence of organic matter resistance to interrill erosion. *Earth Surface Processes and Landforms* **32**: 794-802.
- Kuhn NJ. 2013. Assessing lateral organic Carbon movement in small agricultural catchments. In: Graf, C. (Red.) *Mattertal - ein Tal in Bewegung. Publikation zur Jahrestagung der Schweizerischen Geomorphologischen Gesellschaft 29. Juni-1. Juli 2011, St. Niklaus, Birmensdorf, Eidg. Forschungsanstalt WSL*: 151-164.
- Le Bissonnais Y, Bruand A, Jamagne M. 1989. Laboratory experimental study of soil crusting: relation between aggregate breakdown mechanisms and crust structure. *Catena* **16**: 377-392.
- Le Bissonnais Y. 1996. Aggregate stability and assessment of soil crustability and erodibility: I. Theory and methodology. *European Journal of Soil Science* **47**: 425-437.
- Loch RJ. 2001. Settling velocity – a new approach to assessing soil and sediment properties. *Computers and Electronics in Agriculture* **31**: 305-316.
- Morgan RPC, Quinton JN, Smith RE, Govers G, Poesen JWA, Auerswald K, Chisci G, Torri D, Styczen ME. 1998. The European soil erosion model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments. *Earth Surface Processes and Landforms* **23**: 527-544.
- Nimmo JR, Perkins KS. 2002. Aggregate stability and size distribution. in Dane J.H. and Topp G.C., eds., *Methods of soil analysis, Part 4-Physical methods*. Madison, Wisconsin, *Soil Science Society of America*: 317-328.
- North PF. 1976. Towards an absolute measurement of soil structural stability using ultrasound. *J. Soil Sci.* **27**: 451-459.
- Puri AN. 1934. A siltometer for studying size distribution of silts and sands. *Punjab Irrigation Institute Research Publication* **2(7)**: 10pp.
- Rex JF, Petticrew EL. 2006. Pacific salmon and sediment flocculation: nutrient cycling and intergravel habitat quality. *Sediment Dynamics and the Hydromorphology of Fluvial Systems* (Proceedings of a symposium held in Dundee, UK, July 2006) **306**: 1-8.

- Rubey W. 1933. Settling velocities of gravel, sand and silt particles. *Am. J. Sci.* **225**: 325-338.
- Slattery MC, Burt TP. 1997. Particle size characteristics of suspended sediment in hillslope runoff and stream flow. *Earth Surface Processes and Landforms* **22 (8)**: 705-719.
- Starr GC, Lal R, Malone R, Hothem D, Owens L, Kimble J. 2000. Modeling soil carbon transported by water erosion processes. *Land Degradation and Development* **11**: 83-91.
- Tromp-van Meerveld HJ, Parlange JY, Barry DA, Tromp MF, Sander GC, Walter MT, and Parlange MB. 2008. Influence of sediment settling velocity on mechanistic soil erosion modeling. *Water Resources Research* **44**: W06401, doi:10.1029/2007WR006361.
- Van Oost K, Beuselinck L, Hairsine PB, Govers G. 2004. Spatial evaluation of a multi-class sediment transport and deposition model. *Earth Surface Processes and Landforms* **29**: 1027-2044.
- Walling DE. 1983. The sediment delivery problem. *Journal of Hydrology* **65**: 209-237.
- Walling DE. 1988. Erosion and sediment yield research – some recent perspectives. *Journal of Hydrology* **100**: 113-141.
- Wong KB, Piedrahita RH. 2000. Settling velocity characterization of Aquacultural solids. *Aquacultural Engineering* **21**: 233-246.
- Wu WM, Wang SY. 2006. Formulas for Sediment Porosity and Settling Velocity. *Journal of Hydraulic Engineering* **8**: 858-862.

1.2.3. Determination of Moisture and Total Organic Content within Basin Environments: Loss-on-Ignition

Jamie C. Wood¹

¹ Centre for Environmental Change and Quaternary Research, University of Gloucestershire, UK.
(JamieWood@connect.glos.ac.uk)



ABSTRACT: Basin environments capture material from the surrounding catchment and preserve this within their sedimentology, enabling them to act as environmental archives. These archives provide geomorphologists with a wide range of proxy data from which to reconstruct palaeoenvironmental change. Of the available proxies, total organic content (TOC) remains one of the simplest to assess through loss-on-ignition (LOI). This technique adopts two phases; samples are initially dried and are then ignited at elevated temperatures to burn off all organic carbon, thus providing a measure of moisture and organic content respectively. The ease of assessment and links between organic content and climatic change make TOC an attractive proxy for researchers focusing on basin environments. Although the technique is simple, LOI has caveats which can lead to either under- or overestimation of TOC. This chapter provides an overview of the LOI process and highlights the potential issues which users could face when adopting the technique.

KEYWORDS: Loss-on-Ignition; organic; moisture; cores; basin environments

Introduction

Terrestrial basin environments (e.g. lakes and peat bogs) can act as rich archives when exploring palaeoenvironmental change. The ability of basins to preserve material captured from surrounding areas as stratigraphic profiles enables geomorphologists to use a wide range of proxy data to reconstruct palaeoenvironmental change. Despite likely variations in sedimentation rates, many basins will constantly accumulate material through background sedimentation (Gilli *et al.*, 2013); enabling them to produce continuous, high-resolution sequences. If sampled carefully, cores from these environments can produce records that are decadal or even sub-decadal in resolution (Battarbee, 2000).

The composition of basin deposits is particularly dependent upon the climate, as it can heavily influence both sediment supply and organic productivity (Leeder *et al.*, 1998; Ariztegui *et al.*, 2001). A large proportion of the organic material that is deposited into basins is derived from plant detritus (Meyers

and Lallier-Vergès, 1999). Variations in the organic productivity and detrital availability of an area are typically recorded within basin profiles as fluctuations in Total Organic Carbon (TOC). The linkages between environmental conditions and the organic content of sediment (Meyers, 1997) make TOC an attractive proxy when reconstructing both abrupt and long-term climatic events from core samples.

Loss-on-Ignition (LOI) is a relatively quick and inexpensive method for determining the total organic carbon (TOC) throughout core lengths (Heiri *et al.*, 2001). Samples are initially dried and then held at high temperatures to ignite all present organics, the loss in mass at each of these stages providing an evaluation of moisture content (%) and TOC (%) respectively. Since LOI results are expressed as percentages it can be difficult to distinguish whether fluctuations are a result of absolute changes in organic or mineral matter (Birks and Birks, 2006), indicating that the technique is not suitable as an independent proxy. However, when used

in conjunction with other proxies, such as pollen (e.g. Beer *et al.*, 2007), LOI can provide valuable support when reconstructing palaeoenvironmental change.

Although this article addresses LOI for use upon terrestrial basin samples, the method is not restricted to this environment. The technique has also been used to assess forest surface soils (De Vos *et al.*, 2005), to identify sedimentological changes associated with tsunami phases (Hawkes *et al.*, 2007), and has been adopted in many other environmental settings. Regardless of the focus, studies adopting LOI will use the same, or a very similar, procedure to the one highlighted within this chapter (see Figure 2).

Despite being a relatively simple method, there are caveats which have led to the accuracy and precision of LOI being questioned within the literature (e.g. Santisteban *et al.*, 2004). But providing that the method is not used independently and that each stage is standardised, LOI remains a useful tool for geomorphological studies. This chapter provides an overview of the LOI technique and highlights some of the key decision-making aspects of the methodology.

Methodological Considerations

Sample size

There are various recommendations for how much sample to use for the LOI process, with some studies adopting given masses and others volumes. Although neither sample mass nor volume has been highlighted as superior measure for sampling, it is essential that sample size is always standardised. This is particularly important when dealing with organic rich samples and short ignition times (i.e. 2-4 hour) as increases in sample size have been shown to reduce the TOC estimated by LOI (Schulte *et al.*, 1991). If considering this factor, users may choose to use samples between 1 and 4 grams.

Ignition temperature

When selecting an ignition temperature it is essential to consider not just the combustion temperature of organics but the effects of these elevated temperatures upon the mineral fractions. Many peatland and lacustrine studies adopt a 2-4 hour 550°C

ignition treatment for LOI (LOI₅₅₀) and commonly cite the methodology used by Heiri *et al.* (2001). However, the most suitable ignition time and temperature for samples can vary between core lengths, depending on their composition and sedimentary source (Wang *et al.* 2011).

Boyle (2004) highlights 4 simplified stages of thermal decay within organic matter (summarised in Table 1). On initial observation it would seem that an LOI₅₅₀ treatment would be a logical protocol for assessing TOC, as the LOI₅₅₀ combusts the most stable humified organics. However, it is clearly observed within the literature that clay minerals lose most of their structural water (dewater) above 450°C (De Vos *et al.*, 2005; Salehi *et al.*, 2011), meaning LOI₅₅₀ would produce an overestimate of TOC should a sample be clay rich. Additional overestimation of TOC can be created through the loss of volatile salts, metal oxides and when using LOI temperatures of >500°C, through the loss of CO₂ from carbonates (Heiri *et al.*, 2001; Hirota and Szyper, 1975). As a result of this it has been suggested that the LOI₅₅₀ should only be used for samples which are predominantly organic (Bhatti and Bauer, 2002) with only a small clay mineral fraction. Within the literature there appears to be no recommended organic to mineral ratio for ignition temperature selection. Users should therefore visually assess samples and consider their environmental setting, to determine the most likely composition of samples (i.e. organic-rich or mineral-rich).

Table 1. Stages of thermal organic decay (adapted from Boyle, 2004)

Stage	Temp (°C)	Process
1	40 - 135	Dehydration of plant matter
2	250 - 350	Breakdown of fatty acids, peptides and carbohydrates within organics
3	370 - 420	Decomposition of less stable humified organics
4	530 - 540	Decomposition of more stable humified organics

Many have adopted LOI procedures with lower ignition temperatures to avoid TOC overestimation created by clay minerals and carbonates. An example of this comes from Davies (1974) who tested the use of a 24 hour LOI treatment at 430°C and found excellent comparability between LOI₄₃₀ results and the more traditional Walkley-Black method (Walkley and Black, 1934) for analysing organic content. The study also found that carbonates that were added to samples within the laboratory, were not impacted by the LOI₄₃₀ procedure (Davies, 1974).

The method used by Davies (1974) adopts a long duration ignition which may not be suitable for processing a large quantity of samples, however, there are many variants of the LOI procedure with shorter ignition times that can be adopted (e.g. Table 2). Schulte and Hopkins (1996) and Schulte *et al.* (1991) provide an overview of various LOI ignition temperatures and durations adopted throughout the literature. They also highlight an R^2 value which highlights the comparability of the adopted LOI techniques with other more traditional methods for assessing TOC. Table 2 provides a selection of the ignition treatments highlighted within these studies.

Table 2. Example LOI treatments and comparability (R^2) with more traditional methods of obtaining TOC (adapted from Schulte and Hopkins, 1996; Schulte *et al.*, 1991).

Temp (°C)	Ignition Duration (h)	R^2
360	2	0.97
400	8	0.97
430	24	0.99
500	4	0.87
600	6	0.86

N.B.: Not all data was used from original source. For full data tables see: Schulte and Hopkins (1996 - p23.) and Schulte *et al.* (1991 - p.161).

When selecting an LOI treatment from the highlighted literature, it remains essential to consider the likely composition of samples,

ensuring that those which are not organic rich are ignited below 500°C. The selected temperature should always be reported within any write-up (Heiri *et al.* 2001).

Most LOI methods are likely to either overestimate or underestimate TOC, owing to the overlapping combustion and dewatering temperatures highlighted above.

Crucible positioning

Although difficult to accommodate when processing a large number of samples, the positioning of crucibles within a furnace can impact the LOI results. Heiri *et al.* (2001) found differences of up to 3.7% between crucible furnace positions. Differences between crucibles positioned centrally and on the outer margins of a furnace were seen to be more significant with longer ignition periods (Figure 1). Despite the errors associated with crucible positioning appearing relatively small, they still contribute to the intrinsic uncertainties associated with the LOI technique and users should be aware of this potential error. When dealing with a small number of samples, users may choose to consult the results provided by Heiri *et al.* (2001) to minimise these errors by selecting the most reproducible positions within a furnace.

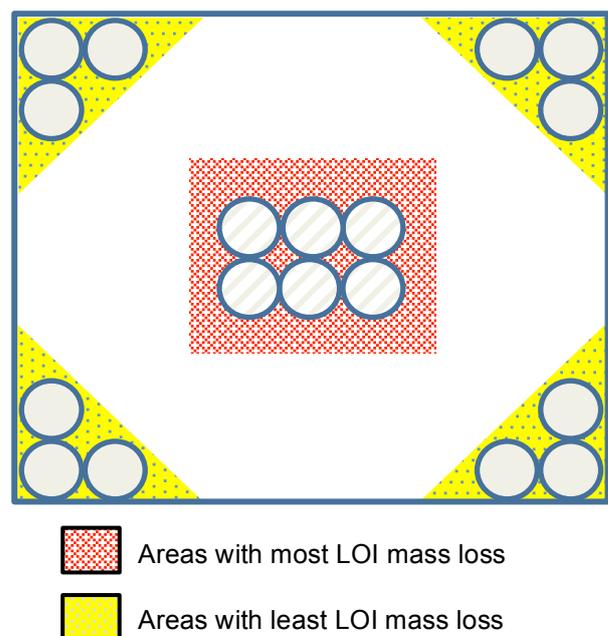


Figure 1. Plan view of crucible furnace positions and differences in LOI₅₅₀ mass loss (adapted from Heiri *et al.*, 2001).

Loss-On-Ignition Process

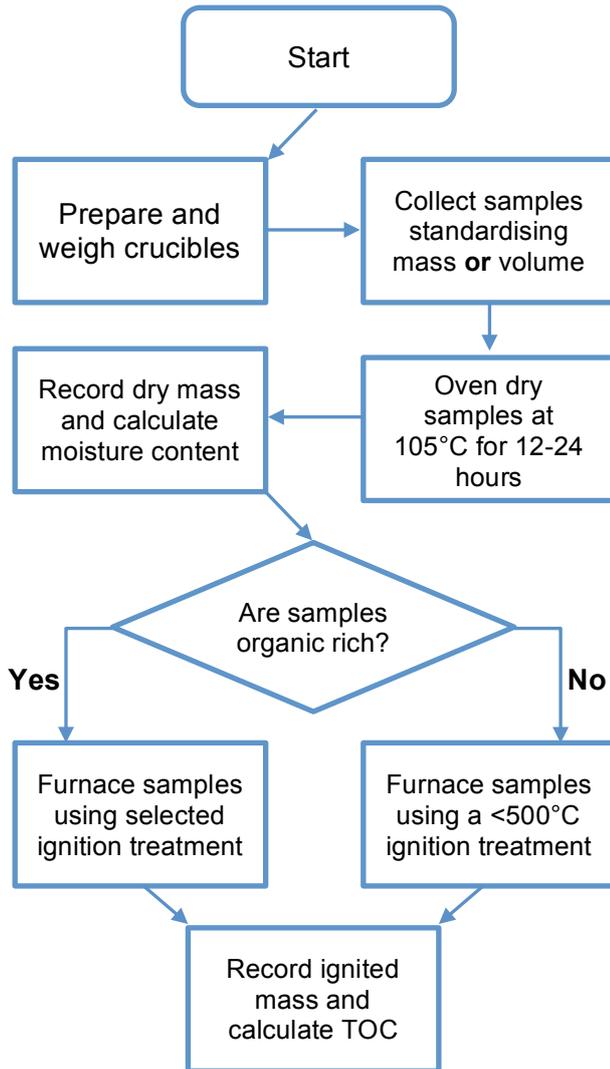


Figure 2. LOI procedure flow diagram.

Preparation

Porcelain crucibles are required for the LOI process as they can withstand the elevated temperatures used for ignition. Despite evidence suggesting that crucible volume does not impact LOI (Schulte *et al.*, 1991), it may be preferable to use the same size/shape crucibles to eradicate any potential inter-sample variations.

Prior to sampling for LOI, crucibles require some preparatory steps; a crucible for each sample, plus a few spares, should be prepared. New or soiled crucibles should be exposed to the selected ignition treatment prior to sampling, this ensures that crucibles can withstand the elevated temperatures and that all thermally unstable contaminants have been burnt off (Gale and Hoare, 1991). Following this initial stage crucibles can be

labelled appropriately (e.g. depth of sample within core) using a permanent marker and weighed to at least three decimal places.



Figure 3. Obtaining crucible mass using a 4.d.p balance. Sampling spatula (left) and 1.25 cm³ measuring spoon (right) included.

Core sampling

When dealing with samples of limited sample mass it may be preferable to sample volumetrically. This ensures that low density strata are not depleted when attempting to standardise sample size. For studies adopting a volumetric sampling strategy, it is common to use 1 cm³ subsamples for LOI (e.g. Beer *et al.*, 2007; Dapples *et al.*, 2002; Shuman, 2003). However, if 1 cm³ volumetric samplers are not available, there is the potential to use a spatula and 1.25 cm³ ($\frac{1}{4}$ teaspoon) sampling spoons (Figure 3). These are affordable, easy to clean, produce replicable samples and are available from most homeware stores.

As core lengths will begin to dry out immediately upon opening, it is essential that LOI subsamples are collected swiftly, particularly if moisture content is required (e.g. for dose rate assessment in luminescence dating). It is recommended that crucibles are laid out on glass trays in numerical order prior to sampling, this helps users keep track when sampling and ensures that sample collection is as swift as possible. Before sampling material for LOI analysis the surface of cores should be removed, to

ensure that sample impacted by corer noise does not influence results. In this instance corer noise describes the contamination and distortion of stratigraphic units created during the coring process; this can be caused by frictional interactions between samples and the collection chamber. A spatula can be used to scrape away a thin layer of the unwanted material and, as with all core samples, users should scrape horizontally across cores to preserve their lithostratigraphic integrity.

Once the core surface is cleaned sampling can be conducted at the required resolution. Material should be removed from core tubing using a spatula and then placed into a measuring spoon. Note that compaction should be avoided, so material should not be pressed with force into sampling spoons. Collected LOI samples should be transferred from the measuring spoon into the prepared crucibles and weighed immediately. This provides the wet mass of samples.

Assessment of moisture content

Once the wet mass has been obtained, crucibles should be placed in an oven at 105°C for between 12 and 24 hours to dry samples (Gale and Hoare, 1991; Ben-Dor and Banin, 1989; Heiri *et al.* 2001). A 105°C oven treatment is adopted by a majority of studies as it will dehydrate most components within a sample. Also this drying temperature does not coincide with the ignition temperature of organics, meaning an underestimation of TOC will not occur during this process. After the drying period crucibles should be removed from the oven and immediately placed into desiccators to cool, this prevents samples from rehydrating. Once cool, samples can be weighed. An attempt should be made to restrict the amount of time the crucibles are exposed to the laboratory air to prevent rehydration. To calculate moisture content the following equation can be used:

$$\text{Moisture Content (\%)} = 100 \times \frac{WS - LOI_{105}}{WS}$$

Where: LOI_{105} and WS are dry sample mass (g) and wet sample mass (g) respectively.

Assessment of TOC

Key pieces of equipment are required for the ignition process (Figure 4). These include:

- Muffle furnace
- Tongs
- Heat proof gloves
- Safety Visor
- Heat proof trays



Figure 4. Furnace and safety equipment.

Before loading, users should preheat the furnace to the selected LOI temperature. When this temperature has been reached crucibles can be loaded using the tongs and safety equipment highlighted. Crucibles should be loaded into the furnace as a series of rows, to ensure that crucibles are easily identifiable. It is critical for users to map the position of samples within a furnace as the permanent marker may burn off.

Once the selected LOI treatment time has been achieved crucibles can be removed and placed onto heatproof tray. Samples should not be allowed to cool completely as they may begin to rehydrate, so should be relabelled and transferred into a desiccator (using tongs) prior to reweighing.

When the samples have been reweighed the TOC can be calculated using the following equation:

$$TOC(\%) = 100 \times \frac{LOI_{105} - LOI_{Ign.}}{LOI_{105}}$$

Where: LOI_{105} and $LOI_{Ign.}$ are dry sample mass (g) and ignited sample mass (g) respectively.

Further Uses

The environmental setting may influence the ignition temperature but it does not influence the LOI process and many studies using LOI will adopt a similar procedure to the one outlined here.

As an additional step, many studies employ an elevated 950°C ignition treatment following TOC acquisition, to assess the carbonate content of samples (Heiri *et al.* 2001). This provides an understanding of the minerogenic component of core samples. Carbonate content provides an additional aspect for multi-proxy studies and can be used to infer changes in a range of variables; salinity, catchment stability and climatically driven changes in mineral availability. (Cassina *et al.*, 2013; Birks and Birks, 2006).

Conclusion

This chapter provides an overview of the LOI technique and the associated methodological procedures. It is apparent that there is not an optimum ignition temperature for LOI and that users should take the composition of core lengths into consideration before selecting a thermal treatment. It is noted that LOI is not restricted to basins or core samples and that it can be adopted in many environmental settings. The overlapping dewatering and combustion temperatures of various materials mean that LOI will not provide a high precision estimate of organic or moisture content; however, if used in conjunction with other methods the results obtained from LOI can form a good basis for geomorphological interpretation.

Acknowledgements

I would like to thank Dr Sue Dawson and Dr Karen Scott for their time and constructive comments which helped to improve this article. I would also like to thank the current editor of Geomorphological Techniques, Dr Lucy Clarke, for her assistance and the time that she has given to process this article.

This article was completed during a funded PhD studentship, final thanks are extended to the University of Gloucestershire and my supervisory team for their support.

References

- Ariztegui D, Chondrogianni C, Lami A, Guilizzoni P, Lafargue E. 2001. Lacustrine organic matter and the Holocene paleoenvironmental record of Lake Albano (central Italy). *Journal of Paleolimnology* **26**: 283-292. DOI: 10.1023/A:1017585808433
- Battarbee R. 2000. Palaeolimnological approaches to climate change, with special regard to the biological record. *Quaternary Science Reviews* **19**: 107-124. DOI: 10.1016/S0277-3791(99)00057-8
- Beer R, Heiri O, Tinner W. 2007. Vegetation history, fire history and lake development recorded for 6300 years by pollen, charcoal, loss on ignition and chironomids at a small lake in southern Kyrgyzstan (Alay Range, Central Asia). *The Holocene* **17**: 977-985. DOI: 10.1177/0959683607082413
- Ben-Dor E, Banin A. 1989. Determination of organic matter content in arid-zone soils using a simple "loss-on-ignition" method. *Communications in Soil Science and Plant Analysis* **20**: 1675-1695. DOI: 10.1080/00103628909368175
- Bhatti J, Bauer I. 2002. Comparing loss-on-ignition with dry combustion as a method for determining carbon content in upland and lowland forest ecosystems. *Communications in Soil Science and Plant Analysis* **33**: 3419-3430. DOI: 10.1081/CSS-120014535
- Birks H, Birks J. 2006. Multi-proxy studies in palaeolimnology. *Vegetation History and Archaeobotany* **15**: 235-251. DOI: 10.1007/s00334-006-0066-6
- Boyle J. 2004. A comparison of two methods for estimating the organic matter content of sediments. *Journal of Paleolimnology* **31**: 125-127. DOI: 10.1023/B:JOPL.0000013354.67645.df
- Cassina F, Dalton C, Dillane M, de Eyto E, Poole R, Sparber K. 2013. A multi-proxy palaeolimnological study reconstruct the evolution of a coastal brackish lake (Lough Furnace, Ireland) during the late Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* **383-384**: 1-15. DOI: 10.1016/j.palaeo.2013.04.016
- Dapples F, Lotter A, van Leeuwen F, van der Knaap W, Dimitriadis S, Oswald D. 2002. Paleolimnological evidence for increased landslide activity due to forest clearing and

- land-use since 3600 cal BP in the western Swiss Alps. *Journal of Paleolimnology* **27**: 239-248. DOI: 10.1023/A:1014215501407
- Davies B. 1974. Loss-on-ignition as an estimate of soil organic matter. *Soil Science of America Journal* **38**: 150-151. DOI: 10.2136/sssaj1974.03615995003800010046x
- De Vos B, Vandecasteele B, Deckers J, Muys B. 2005. Capability of loss-on-ignition as a predictor of total organic carbon in non-calcareous forest soils. *Communications in Soil Science and Plant Analysis* **36**: 2899-2921. DOI: 10.1080/00103620500306080
- Gale S, Hoare P. 1991. Quaternary sediments: Petrographic methods for the study of unlithified rocks. Wiley: New York.
- Gilli A, Anselmetti F, Glur L, Wirth S. 2013. Lake Sediments as Archives of Recurrence Rates and Intensities of Past Flood Events. In *Dating Torrential Processes on Fans and Cones*, Schneuwly-Bollschweiler M, Stoffel M, Rudolf-Miklau F (eds). Springer: Netherlands; 225-242. DOI: 10.1007/978-94-007-4336-6_15
- Hawkes A, Bird M, Cowie S, Grundy-Warr C, Horton B, Hwai A, Law L, Macgregor C, Nott J, Eong Ong J, Rigg J, Robinson R, Tan-Mullins M, Tiong Sa T, Yasin Z, Aik L. 2007. Sediments deposited by the 2004 Indian Ocean tsunami along the Malaysia-Thailand Peninsula. *Marine Geology* **242**: 169-190. DOI: 10.1016/j.margeo.2007.02.017
- Heiri O, Lotter A, Lemcke G. 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: reproducibility and comparability of results. *Journal of Paleolimnology* **25**: 101-110. DOI: 10.1023/A:1008119611481
- Hirota J, Szyper J. 1975. Separation of total particulate carbon into inorganic and organic components. *Limnology and Oceanography* **20**: 896-900. DOI: 10.4319/lo.1975.20.5.0896
- Leeder M, Harris T, Kirkby M. 1998. Sediment supply and climate change: implications for basin stratigraphy. *Basin Research* **10**: 7-18. DOI: 10.1046/j.1365-2117.1998.00054.x
- Meyers P. 1997. Organic geochemical proxies of paleoceanographic, paleolimnologic, and paleoclimatic processes. *Organic Geochemistry* **27**: 213-250. DOI: 10.1016/S0146-6380(97)00049-1
- Meyers P, Lallier-Vergès E. 1999. Lacustrine sedimentary organic matter records of Late Quaternary paleoclimates. *Journal of Paleolimnology* **21**: 345-372. DOI: 10.1023/A:1008073732192
- Salehi M, Hashemi Beni O, Beigi Harchegani H, Esfandiarpour Borujeni I, Motaghian H. 2011. Refining soil organic matter determination by loss-on-ignition. *Pedosphere* **21**: 473-482. DOI: 10.1016/S1002-0160(11)60149-5
- Santisteban J, Mediavilla R, López-Pamo E, Dabrio M, Zapata M, García M, Castano S, Martínez-Alfaro, P. 2004. Loss on ignition: a qualitative or quantitative method for organic matter and carbonate mineral content in sediments?. *Journal of Paleolimnology* **32**: 287-299. DOI: 10.1023/B:JOPL.0000042999.30131.5b
- Schulte E, Hopkins B. 1996. Estimation of soil organic matter by weight loss-on-ignition. In *Soil organic matter: Analysis and interpretation*, Magdoff F, Tabatabai M, Hanlon E (eds). Soil Science Society of America: Madison; 21-31.
- Schulte E, Kaufmann C, Peter J. 1991. The influence of sample size and heating time on soil weight loss-on-ignition. *Communications in Soil Science and Plant Analysis* **22**: 159-168. DOI: 10.1080/00103629109368402
- Shuman B. 2003. Controls on loss-on-ignition variation in cores from two shallow lakes in the northeastern United States. *Journal of Paleolimnology* **30**: 371-385. DOI: 10.1023/B:JOPL.0000007226.68831.e3
- Walkley A, Black I. 1934. An estimation of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science* **37**: 29-38. DOI: 10.1097/00010694-193401000-00003
- Wang Q, Yuncong L, Wang Y. 2011. Optimizing the weight loss-on-ignition methodology to quantify organic and carbonate carbon of sediments from diverse sources. *Environmental Monitoring & Assessment* **174**: 241-257. DOI: 10.1007/s10661-010-1454-z

1.3.1. Measuring the shear strength of cohesive sediment in the field

Robert C. Grabowski^{1,2}

¹ School of Geography, Queen Mary, University of London, UK (r.c.grabowski@gmail.com)

² Cranfield Water Science Institute, Cranfield University, UK



ABSTRACT: Shear stress is a fundamental driver of geomorphic change and is at the centre of a range of geomorphological processes including the deformation of glacier tills, mass movements on hillslopes, riverbank erosion and the stability of intertidal sediment. Consequently, measurements of a sediment's resistance to shear stress, i.e. shear strength, is essential in our field. Numerous techniques have been developed to quantify both surface and internal shear strength for soils and sediments, but most are limited to the laboratory. Measurements of shear strength, particularly for cohesive sediment, are best conducted in the field in the first instance, because the extraction, transport and storage of cores prior to analysis in the laboratory cause physical, chemical, and biological changes to the sediment characteristics that alter its shear strength. The handheld shear vane and the cohesive strength meter (CSM) are portable devices that can be used to measure internal and surface shear strength, respectively, in the field. The shear vane quantifies the undrained geotechnical shear strength of the sediment, i.e. resistance to deformation / fracture, and the CSM measures resistance to surface erosion by water. These tools can be combined with geomorphological mapping, stratigraphy, and sedimentological and biological analyses to support a range of investigations into geomorphological forms and processes.

KEYWORDS: shear vane, erosion threshold, erodibility, jet test, CSM

Introduction

Sediment is continually subjected to physical stress in the environment, whether from the crushing weight of a glacier, the scouring flows of water, or the relentless pull of gravity. Forces push, pull and manipulate sediment. When these forces are imposed parallel to the sediment surface, the stress generated is termed a shear stress (force per m²; Pascal, Pa) (Figure 1).

Shear stress is of paramount importance to geomorphologists because it is a driver of geomorphic change. The term 'shear stress' is used in various ways in earth science, and a clarification is needed at this early point in the chapter (for more information, see Fookes et al. 2007, Ch. 3). Shear stress can describe forces applied parallel to the sediment surface (e.g. surface drag by

flowing water) that result in a surface response (e.g. entrainment of surficial sediment) (Figure 1b). It can also be used to describe forces applied at the sediment surface (e.g. a moving glacier), but which are transmitted into the bed to exert internal shear stress that results in sediment deformation or fracture (Figure 1c). Finally, shear stress can be used to describe a force that acts upon the whole sediment bed (e.g. gravity) and which induces a shear stress that can result in sediment deformation or mobilisation (e.g. hillslope mass movement) (Figure 1d). In this chapter, I take an inclusive view of shear stress. Geomorphic processes in nature often involve a combination of stresses. A flowing glacier not only exerts a shear stress on the sediment bed, but the glacier's weight also imposes an extraordinary compressive stress normal to

the sediment surface. Likewise, the erosion of a sediment bed in a river is linked not only to shear stress generated by fluid drag along the surface but also to small-scale unsteady flow events (e.g. bursts and sweeps) that impart forces normal to the sediment surface. Therefore, the chapter focuses on methods that quantify the shear strength of sediment as it relates generally to either internal shear stress or surface shear stress.

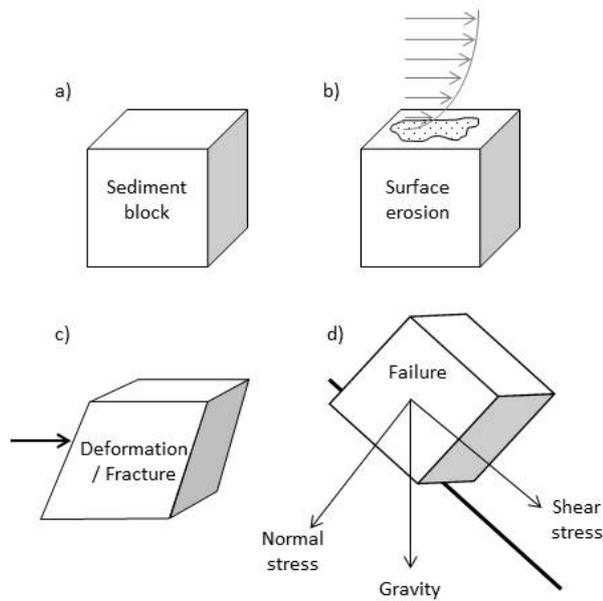


Figure 1: Shear stress imposed on a block of sediment resulting in (b) surface erosion, (c) sediment deformation / fracture, or (d) slope failure

Shear strength is a measure of a sediment's resistance to shear stress, or in other words it is a measure of the sediment's stability in space and time. When an applied shear stress exceeds the shear strength of a sediment, failure occurs and the sediment deforms, fractures and / or mobilises (Figure 1). Shear strength is an attribute of the sediment itself and is determined by the physical, chemical and biological characteristics that influence particle interactions within the sediment (cohesion, adhesion and friction).

Many excellent methods have been developed to quantify shear strength for soils and sediments. The most accurate of these are conducted within a laboratory using specialist equipment. However, cohesive sediment undergoes significant changes in sediment properties when it is cored, transported, stored and finally analysed in the

laboratory, and these changes can significantly alter its shear strength (Knappett and Craig, 2012). Consequently, field approaches are recommended in the first instance for the quantification of *situ* shear strength of cohesive sediment.

This chapter provides an introduction to the *in situ* quantification of shear strength for cohesive sediment specifically with the geomorphologist in mind. It is divided into two sections. The first pertains to geotechnical shear strength, in which internal shear stress applied to sediment causes it to deform or fracture. The second section relates to stresses induced at the sediment surface specifically by flowing water and which results in erosion of the sediment surface. In each section, a single tool is presented in detail, one which is ideally suited for fieldwork by a geomorphologist. Complimentary techniques that can be applied in the laboratory are listed in each section.

Geotechnical shear strength

Many laboratory-based tests have been developed over the years to quantify the shear strength of soil and sediment, including the triaxial and direct shear tests. Detailed information on the theory and application of these approaches can be found in a soil engineering textbook (e.g. Terzaghi et al., 1996; Knappett and Craig, 2012) and the British Standard 1377.

Shear Vane

For *in situ* applications, the shear vane is the most reliable and readily-available device for measuring the undrained shear strength of cohesive sediment. It has been used extensively for the analysis of shear strength in soils (Serota and Jangle, 1972; Knappett and Craig, 2012), and has been applied to glacial till (e.g. Khan and Kostaschuk, 2011), riverbanks (e.g. Chen et al., 2012), and fine bed sediments in riverine (e.g. Grabowski, 2010) and marine environments (e.g. Hauton and Paterson, 2003).

Handheld shear vanes are compact, portable devices that can be easily carried into the field and used *in situ* (Figure 2). They typically come equipped with vanes of varying sizes to measure shear strength

across a wide range of sediment bulk densities. For example, the Pilcon Hand Vane Tester (EDECO, England) has two vanes, a 19 mm diam. vane for use in consolidated sediment and a 33 mm diam. vane for softer sediment.

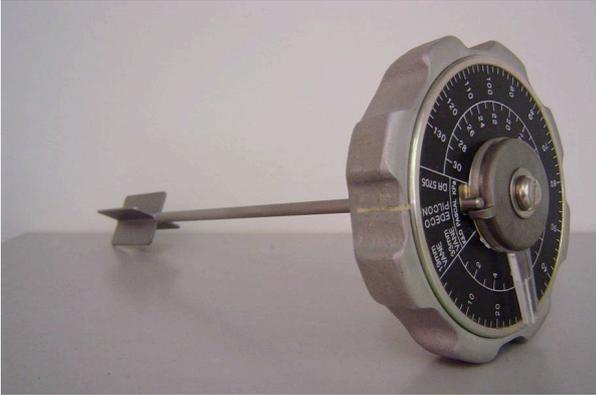


Figure 2: A shear vane tester equipped with a 33 mm diameter vane. The vane is inserted into the sediment, and the dial is rotated until the sediment fails, and the vane spins freely in the sediment.

To operate the device, the vane is inserted into the sediment perpendicular to the sediment surface, avoiding any lateral movement. Standard protocol for soil testing sets a minimum depth of insertion (70-80 mm for the Pilcon model); however the device can be inserted to a range of depths to investigate vertical changes in shear strength for cohesive sediment. The vane head is then rotated at a constant rate (e.g. 1 rev min⁻¹) until the sediment fails (i.e. the vane rotates freely in the sediment). The maximum shear stress (τ_v ; kN/m²) resisted by the sediment is a function of the torque (M) applied at failure and the size of the vane:

$$\tau_v = \frac{M}{K}$$

$$K = \pi D^2 \left(\frac{H}{2} + \frac{D}{6} \right)$$

where K is a constant that depends on the height (H) and depth (D) of the vane. Commercial shear vane testers will typically come calibrated for their standard size vanes so that τ_v can be recorded directly in the field. Several readings are recommended, and the mean used to characterize a sample. If τ_v differs substantially between readings for a sample, this may indicate that presence of a non-cohesive sediment lens or an obstruction

(a large clast or root), and additional readings should be taken. For more information on conducting a shear vane test, see British Standard 1377 – Part 9.

Sediment samples should be collected for the analysis of sediment properties. At a minimum, sediment grain size distribution and water content should be measured.

It is important to remember that the shear vane is only recommended for use in cohesive soils and sediments. This does not mean that the sample must be composed entirely of clay and silt, though, as fine sediments with relatively low clay contents can be classified as cohesive (van Ledden et al., 2004; Winterwerp and Van Kesteren, 2004; for a recent review see, Grabowski et al., 2011). As stated above, samples with highly heterogeneous sediment grain sizes can be problematic, due to the physical interaction of larger clasts. The presence of other types of material in the sediment, such as fine plant roots, is not as problematic as their tensile strength contributes to the shear strength of the sample (Gyssels et al., 2005). Grabowski (2010) found that the shear vane was able to detect differences in the shear strength of fine riverine sediment caused, in part, by differences in the amount of plant roots in the sediment, which would impact its susceptibility to fracturing and bulk erosion (Figure 3).

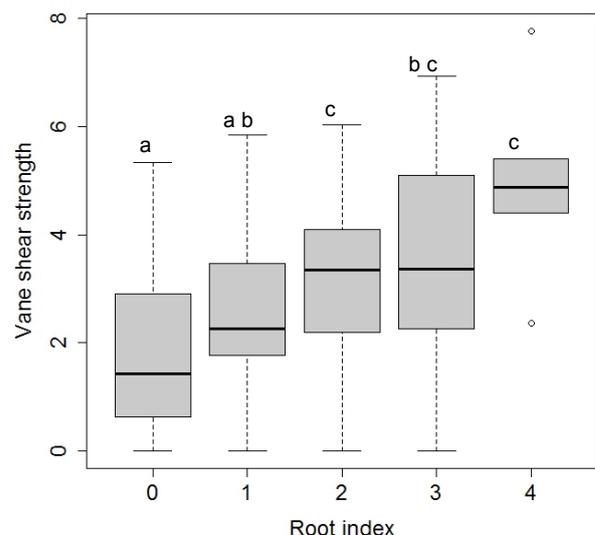


Figure 3: Vane shear strength for fine riverine sediment by root index category (0 = none, 1 = few, 2 = intermediate, 3 = abundant, and 4 = roots dominant) Bere Stream, Dorset, UK. (Kruskal Wallis, Mann-Whitney, $p < 0.05$) (Grabowski, 2010)

A final note about a related device. The pocket shear meter (e.g. Torvane) is similar in principle to the handheld shear vane, but the volume of sediment that is sheared is thinner and is limited to the surface of the sample. It is a useful tool to aid the visual classification of soil but is not suitable for the quantification of geotechnical shear strength of sediment (Head, 1994).

Resistance to erosion

The shear strength of sediment as related to erosion by flowing water is expressed as the shear stress required to initiate surface erosion, i.e. critical shear stress, or erosion rates once erosion has begun.

Flumes are considered the gold standard in this area of research. They generate steady water flows parallel to the sediment surface to impart consistent shear stresses. A wide range of flumes have been developed to quantify erosion resistance: unidirectional and oscillatory; straight and annular; recirculating and flow-through; and laboratory-based and field-deployable systems (e.g. Aberle et al., 2004; Amos et al., 2004; Bale et al., 2007; Droppo et al., 2007; Gerbersdorf et al., 2005; Roberts et al., 2003; Widdows et al., 1998).

Concerns over the impacts of extraction, transport and storage of sediment on sediment properties for laboratory investigations, combined with field evidence of significant spatial and temporal variations in erosion resistance has driven the development of field-deployable systems with small footprints, such as the Gust erosion chamber (Thomsen and Gust, 2000), EROMES (Schünemann and Köhl, 1993) and the cohesive strength meter (CSM) (Paterson, 1989). The CSM is the only system that is commercially available, and is the focus of this section. Whilst the CSM does not generate a shear stress *per se*, it quantifies sediment's resistance to erosion by water, and as was discussed in the introduction this process often involves both shear and normal stresses in nature. Furthermore, the forces generated by the CSM have been shown to correlate with surface shear stress, at least at the lower range of the instrument (Grabowski et al., 2010).

Cohesive Strength Meter (CSM)

The CSM is a portable device that measures the *in situ* erosion resistance of surficial sediments. It is a modern, portable version of the laboratory jet test used by soil scientists for over 50 years to measure the shear strength of clayey soil samples (c.f. Paterson, 1989). Since its conception 20 years ago, the CSM has been increasingly used, particularly in the estuarine environment (e.g. Yallop et al., 1994; Tolhurst et al., 1999; Friend et al., 2005) but more recently in rivers (Grabowski et al., 2012).

The CSM is composed of a main instrument unit containing the computer and a reservoir of water pressurised by an external air cylinder, which is connected to a sensor head that is inserted into the sediment (Figure 4). The device is designed to be deployed *in situ* and the sensor head can be submerged under water. There is a maximum depth of deployment of the sensor head, which is dependent of the rigidity of the tubing, thus the use of core tubes in Figure 4. If the sensor head is being deployed sub-aerially, as is commonly the case for intertidal marine studies, the chamber must be filled with water gently using a syringe. To run a test, the CSM fires a jet of water downward onto the sediment surface within the sensor head chamber. Jet pressure is increased incrementally, and erosion is monitored based on sediment resuspension using an infrared transmissometer.



Figure 4: The cohesive strength meter (CSM) being used in the Bere Stream (Dorset, UK).

The initiation of erosion is typically identified by a 10% drop in light transmission within the test chamber (Figure 5). The CSM records the internal pressure (P_i) for each step in PSI,

but the pressure of the impinging jet on the sediment surface (i.e. stagnation pressure, P_{stag}) is dependent on the specifics of the CSM model, the sampling routine, the length and diameter of tubing, and the height of the jet orifice above the sediment surface (Vardy et al., 2007). The conversion of P_1 to P_{stag} allows erosion thresholds to be compared between different CSM units and models. To be able to incorporate these estimates of erosion thresholds into sediment transport models, P_{stag} must be converted to horizontal shear stress (τ). An empirical calibration was developed for the CSM by Grabowski et al. (2010) by comparing erosion thresholds generated in the CSM and annular flume using homogeneous sand/mud mixtures. For P_{stag} range of 30-110 Pa, critical shear stress was found to relate to P_{stag} as:

$$\tau_c = 0.0013P_{\text{stag}} + 0.047$$

$$(R^2_{\text{adj}} = 0.87, P < 0.01)$$

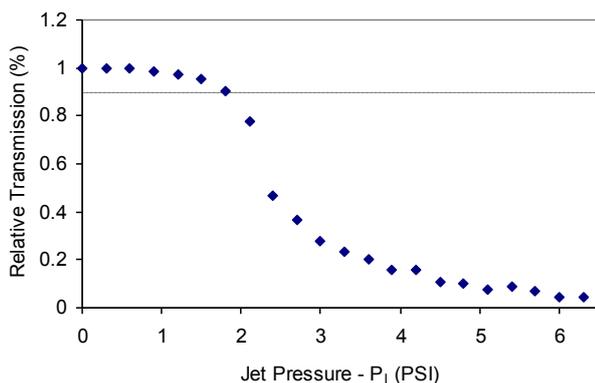


Figure 5: An example of a transmission curve from a cohesive strength meter run. A 10% reduction in transmission was used as the operation definition of erosion.

The cohesive strength meter, as the name suggests, should only be used to estimate erosion thresholds for cohesive sediment. This is primarily due to the method by which the CSM identifies erosion. A minimum clay/silt content is needed to ensure that erosion is synonymous with resuspension. For example, if the CSM was used to measure the erosion threshold of pure sand, the sand would need to be resuspended to a height of 1 cm off the bed to be registered by the transmissometer. In nature, sand grains would be entrained and transported as bedload at shear stress lower than necessary to suspend the sand grains to that height. The minimum clay/silt content will vary depending on the type and amount of clay.

For pure kaolin the threshold minimum was found to be around 5% clay content (Grabowski et al., 2010), which is comparable to minimum content need to form a cohesive structural matrix in the sediment (4 – 10%) (Mitchener and Torfs, 1996; van Ledden et al., 2004). Grabowski et al. (2012) found that a minimum silt/clay content of 10% was appropriate for fine riverine sediment.

Finally, it is worth re-emphasising that the CSM measures only the erosion of the surficial sediment. It does not directly quantify erosion rates or depths of scours, though these could be obtained by an analysis of the rate of change in the slope of the transmission vs jet pressure plots (e.g. Figure 5; Tolhurst et al. 1999) and an empirical relationship between transmission and suspended sediment concentration, or by measuring scour depths after exposure to high jet pressures. A related device, the *in situ* jet test uses a depth-integrated approach to calculate critical shear stress based on the equilibrium scour depth at a set jet pressure. It has been used to quantify the stability of riverbanks (Hanson and Simon, 2001; Simon et al., 2006).

Conclusion

Numerous methods exist to quantify the shear strength of soils and sediments. The tools highlighted in this chapter, the shear vane and CSM, are two that are well suited to field geomorphological research. When used correctly, they generate shear strength readings that can support research in a range of geomorphological forms and processes.

References

- Aberle J, Nikora V, Walters R. 2004. Effects of bed material properties on cohesive sediment erosion. *Marine Geology* 207: 83-93.
- Amos CL, Bergamasco A, Umgieser G, Cappucci S, Cloutier D, DeNat L, Flindt M, Bonardi M, Cristante S. 2004. The stability of tidal flats in Venice Lagoon - The results of in-situ measurements using two benthic, annular flumes. *Journal of Marine Systems* 51: 211-241.

- Bale AJ, Stephens JA, Harris CB. 2007. Critical erosion profiles in macro-tidal estuary sediments: Implications for the stability of intertidal mud and the slope of mud banks. *Continental Shelf Research* 27: 2303-2312.
- Chen Y, Thompson CEL, Collins MB. 2012. Saltmarsh creek bank stability: Biostabilisation and consolidation with depth. *Continental Shelf Research* 35: 64-74. DOI: 10.1016/j.csr.2011.12.009.
- Droppo IG, Ross N, Skafel M, Liss SN. 2007. Biostabilization of cohesive sediment beds in a freshwater wave-dominated environment. *Limnology and Oceanography* 52: 577-589.
- Fookes PG, Lee EM, Griffiths JSD. 2007. *Engineering geomorphology: theory and practice*. Whittles, CRC: Dunbeath, Scotland, UK.
- Friend PL, Lucas CH, Rossington SK. 2005. Day-night variation of cohesive sediment stability. *Estuarine Coastal and Shelf Science* 64: 407-418.
- Gerbersdorf SU, Jancke T, Westrich B. 2005. Physico-chemical and biological sediment properties determining erosion resistance of contaminated riverine sediments - Temporal and vertical pattern at the Lauffen reservoir/River Neckar, Germany. *Limnologica* 35: 132-144.
- Grabowski RC. 2010. *The erodibility of fine sediment deposits in lowland chalk streams*. PhD Thesis submitted In Geography: Queen Mary, University of London, pp266.
- Grabowski RC, Droppo IG, Wharton G. 2010. Estimation of critical shear stress from cohesive strength meter-derived erosion thresholds. *Limnology and Oceanography-Methods* 8: 678-685.
- Grabowski RC, Droppo IG, Wharton G. 2011. Erodibility of cohesive sediment: the importance of sediment properties. *Earth Science Reviews* 105: 101-120.
- Grabowski RC, Wharton G, Davies GR, Droppo IG. 2012. Spatial and temporal variations in the erosion threshold of fine riverbed sediments. *Journal of Soils and Sediments* 12: 1174-1188. DOI: 10.1007/s11368-012-0534-9.
- Gyssels G, Poesen J, Bochet E, Li Y. 2005. Impact of plant roots on the resistance of soils to erosion by water: a review. *Progress in Physical Geography* 29: 189-217.
- Hanson GJ, Simon A. 2001. Erodibility of cohesive streambeds in the loess area of the midwestern USA. *Hydrological Processes* 15: 23-38.
- Hauton C, Paterson DM. 2003. A novel shear vane used to determine the evolution of hydraulic dredge tracks in sub-tidal marine sediments. *Estuarine Coastal and Shelf Science* 57: 1151-1158.
- Head KH. 1994. *Manual of soil laboratory testing*. Halsted: New York.
- Khan I, Kostaschuk R. 2011. Erodibility of cohesive glacial till bed sediments in urban stream channel systems. *Canadian Journal of Civil Engineering* 38: 1363-1372. DOI: 10.1139/I11-099.
- Knappett J, Craig RF. 2012. *Craig's soil mechanics*. Spon Press: Abingdon, Oxon; New York.
- Mitchener H, Torfs H. 1996. Erosion of mud/sand mixtures. *Coastal Engineering* 29: 1-25.
- Paterson DM. 1989. Short-term changes in the erodibility of intertidal cohesive sediments related to the migratory behavior of epipelagic diatoms. *Limnology and Oceanography* 34: 223-234.
- Roberts JD, Jepsen RA, James SC. 2003. Measurements of sediment erosion and transport with the adjustable shear stress erosion and transport flume. *Journal of Hydraulic Engineering-ASCE* 129: 862-871.
- Schünemann M, Kühl H. 1993. Experimental investigations of the erosional behaviour of naturally formed mud from the Elbe estuary and the adjacent Wadden Sea. In: Mehta A (Ed.) *Coastal and Estuarine Studies*, p314-330.
- Serota S, Jangle A. 1972. *Direct-Reading Pocket Shear Vane*. *Civil Engineering* 42: 73-74.
- Simon A, Pollen N, Langendoen E. 2006. Influence of two woody riparian species on critical conditions for streambank stability: Upper Truckee River, California. *Journal of the American Water Resources Association* 42: 99-113.
- Terzaghi K, Peck RB, Mesri G. 1996. *Soil Mechanics in Engineering Practice (3rd Edition)*. John Wiley & Sons: Chichester, UK.

Thomsen L, Gust G. 2000. Sediment erosion thresholds and characteristics of resuspended aggregates on the western European continental margin. *Deep-Sea Research Part I-Oceanographic Research Papers* 47: 1881-1897.

Tolhurst TJ, Black KS, Shayler SA, Mather S, Black I, Baker K, Paterson DM. 1999. Measuring the in situ erosion shear stress of intertidal sediments with the Cohesive Strength Meter (CSM). *Estuarine Coastal and Shelf Science* 49: 281-294.

van Ledden M, van Kesteren WGM, Winterwerp JC. 2004. A conceptual framework for the erosion behaviour of sand-mud mixtures. *Continental Shelf Research* 24: 1-11. DOI: 10.1016/j.csr.2003.09.002.

Vardy S, Saunders JE, Tolhurst TJ, Davies PA, Paterson DM. 2007. Calibration of the high-pressure cohesive strength meter (CSM). *Continental Shelf Research* 27: 1190-1199.

Widdows J, Brinsley MD, Bowley N, Barrett C. 1998. A benthic annular flume for in situ measurement of suspension feeding/biodeposition rates and erosion potential of intertidal cohesive sediments. *Estuarine Coastal and Shelf Science* 46: 27-38.

Winterwerp JC, Van Kesteren WGM. 2004. *Introduction to the Physics of Cohesive Sediment in the Marine Environment*. Elsevier B. V.: Amsterdam.

Yallop ML, de Winder B, Paterson DM, Stal LJ. 1994. Comparative structure, primary production and biogenic stabilization of cohesive and non-cohesive marine sediments inhabited by microphytobenthos. *Estuarine Coastal and Shelf Science* 39: 565-582.

Measuring rock hardness in the field

Lisa Mol^{1,2}

¹ School of Earth and Ocean Sciences, Cardiff University (MOLL@cardiff.ac.uk)

² Rock Breakdown Laboratory, School of Geography and the Environment, Oxford University



ABSTRACT: Rock surface hardness is often used as an indicator of the degree to which a rock surface has weathered. As the surface deteriorates the loss of cohesion results in crumbling of the surface, increased pore water circulation and dislodging of sections such as flakes. It is widely assumed that this results in a lowering of rock surface hardness. However, hardness can also increase if weathering leads to cementation of the surface due to the deposition of solutes such as quartz, clays and small quantities of carbonates. A number of different instruments are available to map out hardness distributions and changes over time. This chapter outlines the use of a simple field test (Moh's hardness test), three rebound devices (Picolo, Equotip and Schmidt Hammer) and resistance drilling as possible methods for assessing rock surface hardness.

KEYWORDS: weathering, field methodology, surface hardness

Introduction

Landscape development and rock weathering go hand in hand; whether it is through weakening and removal of material creating new landforms, the rate of sediment production, accumulation and removal dictating the flow of rivers, or the supply of sediment for sand dunes. All of these are heavily influenced by the ability of material to withstand weathering processes. Understanding the rate of weathering, and associated source and rate of production of sediment, is therefore paramount when investigating wider landscape development.

Geomorphological investigations often are complicated by issues of scale, process rates, and a disconnect between laboratory and field observations. Rock weathering studies are no exception to this, as field measurements and laboratory measurements often don't corroborate (see for example Matsuoka, 2001; White and Brantley, 2003) and questions arise when extrapolating the small scale observations of granular disintegration to larger scale rock face development, or even landscape development (Viles, 2001). Moreover, non-linearity, threshold development and

changing environmental conditions add further to the complexities experienced by researchers. Figure 1 illustrates the multitude of scales on which weathering operates and the presence of positive and negative feedback systems.

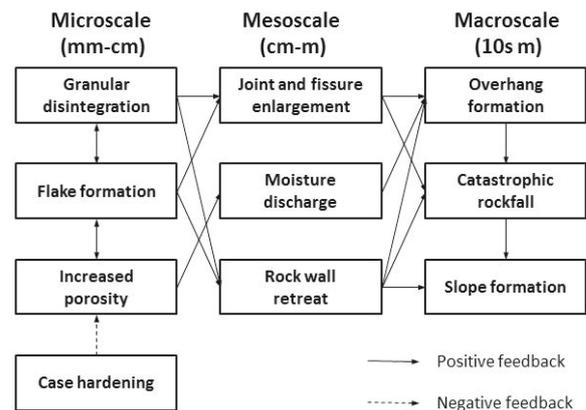


Figure 1: A simple representation of some of the complex feedback mechanisms within weathering. (Adapted from: Mol and Viles, 2012a, p.224).

However, these challenges do not mean that we should not attempt to understand rock weathering rates, quite the contrary as the

vast body of published research indicates that the methodology is rapidly becoming more sophisticated, and diversifying, enabling researchers to approach weathering studies from a multitude of angles such as the impacts of temperature (Smith *et al.*, 2011) and precipitation ingress (McCabe *et al.*, 2013). In addition, methodological precautions such as taking repeat measurements of surface hardness (Coombes *et al.*, 2013), and carrying out chemical analysis to monitor changes in the rock (Buj *et al.*, 2011) aid in reducing the uncertainties caused by the complexities outlined previously. We can therefore use rock surface hardness to pinpoint areas of accelerated weathering. As surface weathering progresses the bonds between minerals weaken, allowing for more moisture absorption and general loss of surface hardness. Hardness can therefore be used as a tool for mapping weathering progression as well as dating exposure of rock surfaces (see Goudie, 2006 for an overview). This chapter discusses some of the methods available to measure rock hardness in situ and their advantages and drawbacks.

Rock weathering and surface hardness measurements; theory

Rock weathering is influenced by many different factors such as: (1) the amount of thermal stress a rock face is subjected to; (2) the presence or absence of water; (3) the development of microbial activity on or under the surface; and, (4) the cycling of chemicals and salts which cause disintegration of the mineral matrix. All of these factors have a net result in common; weakening of the rock surface and subsequent susceptibility to erosion. As surface deterioration sets in, cementation between grains gradually weakens, individual grains disintegrate, and surface volume loss becomes evident. This deterioration can take a number of forms, such as flaking, blistering and accumulation of biological activity (such as algal colonies) and moisture beneath the surface (Mol, 2014), all of which reduce the hardness of the rock surface as subsurface weathering influences porosity and changes the compressive strength and elasticity of a rock (Moses *et al.*, 2014). We can therefore use rock hardness to estimate the progression of weathering processes and specific areas of weakness.

However, there are three complicating factors that should be monitored. Firstly, increasing weakness can lead to an increase in surface roughness, which in turn affects rebound values (McCarroll, 1991), thereby creating a slightly distorted image of rock weathering variations. Secondly, weathering processes could strengthen a rock surface through case hardening. This is created through cementation of the surface by the precipitation of elements such as iron and magnesium within the subsurface. This then leads to a thin, discoloured hardened skin (see for example Viles and Goudie, 2004). Thirdly, variability within the strength of bedding planes in sedimentary rock can lead to differential weathering that can only partially be contributed to outcrop exposure.



Figure 2: Active weathering along the coastal road, Longyearbyen (Svalbard). Clearly visible are the harder sandstones which have not weathered at the same rate as the weaker shale. Also visible is the post-deposition folding of the facies.

When assessing weathering rates you should therefore take into account both external factors such as temperature fluctuations and precipitation levels as well as predisposition to weathering of the individual bedrock bedding planes in a heterogeneous outcrop. Especially in areas where sea level has fluctuated significantly over millions of years the stratigraphy across a rock face can change dramatically vertically. This then leads to complex weathering features, such as found along the coastal road in Longyearbyen, Svalbard (see Figure 2). In this section not only do the depositional phases result in very different weathering rates, the subsequent tectonic activity has created a fold that has altered the physical

structure of the shale, leading to crumbling of the section rather than flaking which dominates the rest of the section. Rock surface hardness should therefore be taken as a relative measurement of weathering and placed within the geological and environmental context of the site.

A simple field method: Moh's test

Moh's test is a very simple procedure in the field that can provide an initial assessment of rock hardness. All that is needed to hand is a rock sample, whether a loose hand sample or an exposed rock surface, and one of a small array of potential objects; a coin, a piece of glass or even your nail. Moh's test uses a scale of known hardness (see Table 1), which you can then compare to the impact of an object when scratched over the surface of

the rock (see Table 2). Glass also is commonly used, where a piece of rock that scratches the surface of the glass is considered hard ($H=5.5$ or above), whereas a piece of rock that does not impact the glass surface is considered soft ($H=5.5$ or below). This field test can be used to give a first indication of the hardness of the test site and, if multiple samples are used, variability of hardness within a limited area. Though this test is often used to distinguish between minerals, it can also be used to test various sections of a homogenous material to determine which sections are most badly affected by weathering processes. It is however, only semi-quantitative and does not provide an exact figure for hardness.

Table 1: Mineral hardness according to Moh's test

1. Talc (H=1)	2. Gypsum (H=2)	3. Calcite (H=3)	4. Fluorite (H=4)	5. Apatite (H=5)
6. Orthoclase (H=6)	7. Quartz (H=7)	8. Topaz (H=8)	9. Corundum (H=9)	10. Diamond (H=10)

Table 2: Commonly available materials and their hardness index

1. Finger nail (H= 2.5)	2. A copper coin (H=3.0)	3. Steel blade (H=5.5)
-------------------------	--------------------------	------------------------

Rebound devices

One group of equipment available are 'rebound devices', a group of devices that use a spring-loaded mechanism to measure rebound of a metal object against a rock face.

The principle is very simple and can be compared to throwing a ball against a wall (though ball velocity in flight towards the wall would be constant); if a wall is very hard the ball will bounce back fast and quite far. Now imagine that same wall covered in a layer of pillows; the ball will bounce back much more slowly and not as far. Rebound devices work on exactly the same principle – the 'cushioning' effect of a reduction in surface hardness dampens the return of the impact device, thereby slowing the arrival time and thus lowering the value measured (Sumner and Nel, 2002). There are a number of different devices available. In particular, the Schmidt Hammer and the Equotip are

commonly used as they are quick, simple, relatively cheap and portable (Goudie, 2013), and non-destructive. They should all be calibrated before the fieldwork takes place using a calibration block, which can be obtained from the supplier.

Schmidt Hammer

The Schmidt Hammer is a rebound device first used in the 1950s but gained momentum in the 1960s (see for example Rusakov and Mavrodi, 1968), and since then has been extensively used in geomorphological research for example in dating rock surface exposure time (Kellerer-Priklbauer *et al.*, 2007) and estimating the effect of environmental controls such as aspect on rock weathering (Hansen *et al.*, 2013). The device contains a spring-loaded piston which is pressed against the rock surface. The dial on the side of the device indicates the rebound value (R-value) on a scale from 0 to

100. The higher the rebound value, the harder the rock surface. To use the device in the field you simply need to press the metal rod against the surface until it is fully compressed within the main chamber (see Figure 3). By pressing the button (on the analogue version there is only 1!) you release the spring mechanism and the rebound is measured on the dial on the side. Care needs to be taken that the device is placed perpendicular to the rock surface, as placing it at an angle will result in the metal impact rod glancing off the surface during rebound instead of measuring the rock surface hardness.



Figure 3: Schmidt Hammer in action.

This device is particularly good for harder rock surfaces such as granite. When using a Schmidt Hammer on softer sedimentary rock care needs to be taken when selecting the model: the N-type works with high impact

energy and could actually be damaging to the point of leaving a mark where it has actively indented the surface. An example of this can be seen in Figure 4 where the white marks on the rock face indicate N-type Schmidt Hammer readings. If there is a risk of this type of damage then the user might want to consider using an L-type Schmidt Hammer instead, which has a low impact energy and is suitable for brittle objects or structures less than 100mm thick. This is important also if the Schmidt Hammer is used in the laboratory to monitor small test blocks. To further reduce the impact a research can attach a mushroom plunger which decreases point-specific impact and instead spreads the impact over a larger area. The outlined advantages and limitations of the different models therefore need to be taken into account when selecting the rock hardness method for the measurement of a specific site or sample.



Figure 4: Impact damage from Schmidt Hammer (photo: Lisa Mol).

Equotip

The Equotip (Figure 5) is essentially the 'little brother' of the Schmidt Hammer. Rather than a loaded piston the Equotip uses a small rebound 'bullet' made of carbon tungsten which is fired by first compressing and then releasing the coil within the impact device (seen to the right in Figure 5). When this bullet rebounds into the device it recompresses this coil, thereby generating a measurements of both impact and rebound velocity. The software then uses the following formula to calculate Leeb value:

$$L = V_r/V_i \times 1000$$

where L is Leeb value, V_r is rebound velocity and V_i is impact velocity (Aoki and Matsukura, 2008). Higher values indicate higher rock surface hardness.

It can be programmed for the appropriate rebound test (i.e. concrete or steel) and impact angle, or used on a fully automatic setting. It can, however, store hundreds of measurements which eliminates the need to painstakingly copy over every measurement by hand in the field, which is not an option on the older Schmidt Hammer models. Care has to be taken when comparing Equotip to Schmidt Hammer as the scales used for measurement are different (R-value vs Leeb-value).

One of the drawbacks of the Equotip is that the piston can collect dust when used repeatedly on heavily weathered surfaces which affects the readings. In addition, because of the lower impact the Equotip is more sensitive to small scale irregularities such as edges and cracks, which needs to be taken into account when surveying either small test blocks or larger surfaces that are heavily cracked.



Figure 5: Equotip with impact device (right). Photo: Lisa Mol.

However, this technique has been successfully applied in a number of studies (see for example Aoki and Matsukura, 2008; Hansen et al, 2013; Matsukura *et al.*, 2007). Figure 6 illustrates this type of research. Here you can see a plot of Equotip data gathered on concave sandstone surfaces in the Golden Gate Highlands National Park, where

variability and development of rock surface hardness was combined with internal moisture distributions to quantify the role of case hardening in tafoni development over time.

Piccolo

The Piccolo is the smallest of the impact devices and, as it is part of the Equotip family, also measures hardness on the Leeb scale allowing for direct comparison of measurements. Due to its size it is even more sensitive to changes in the surface and could equally create noise in the dataset. However, because of its sensitivity it is very suited to small scale investigations into hardness loss such as small building blocks or small samples in environmental simulations. It is also 'pocket sized' which makes it very useful for fieldwork where mobility is a primary concern due to long hikes or difficult terrain.

Drawbacks and limitations of rebound devices

One of the biggest issues with rebound devices is the 'edge effect', where the measurement registered is influenced by the proximity of a rock edge or major crack. This effect is extensively discussed in Viles *et al.* (2011) where it was shown that the smaller the device the less influence the edge effect has. However, the smaller devices (such as the Piccolo) are more sensitive to surface irregularities, leading to potentially a large variability in the data set. When selecting a particular rebound device one has to keep in mind the size of the rock structure (i.e. outcrop vs small laboratory sample) and the surface structure as changes in surface smoothness can lead to variability in the measurements (McCarroll, 1991).

In addition, the user can also influence the readings; small changes such as the pressure exerted on the piston during the rebound measurements and angle of impact can cause fluctuations in the readings, especially if multiple researchers work on the same data set. These small inconsistencies can be traced through statistical analysis of the data set and repeat measurements by multiple users on the same site (see Viles *et al.*, 2011 Figure 4 for an illustration of these small inconsistencies).

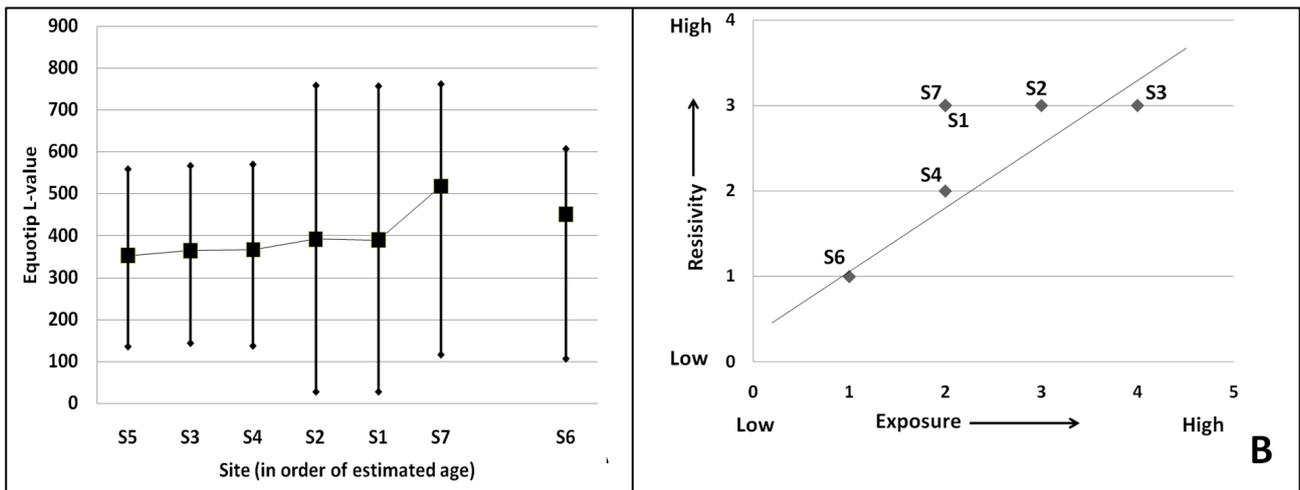


Figure 6: Rock surface hardness and resistivity distributions in relation to tafoni development over time. From: Mol, L and Viles, H. (2012b).

Drilling resistance

Resistance drilling is a method frequently used to assess for example the condition of a building stone or wall. A drill is placed against a stone surface and controlled pressure is applied (Figure 7). The time it takes for the drill to travel a pre-determined distance into the stone surface can be plotted against force used to create a resistance plot, outlining variations in rock hardness through the subsurface. However, even though resistance drilling is classified as 'non-destructive', it leaves a rather noticeable hole in the surface of the stone. This makes it unsuitable for sensitive sites such as heritage and listed buildings if the test specimen is to be returned to the original structure. All tests have to be carried out in the laboratory as the drill is computer controlled, as shown in Figure 6. To minimise the risk of wear and tear of drill heads influencing measurements only diamond-tip drill can be used.

Microdrilling equipment is also available for in situ measurements, which measure up to 50mm depth (see for example Török *et al.*, 2007 and Cnudde *et al.*, 2009 for more information) and can be used to map the presence of weathering crusts. This technique works on the same principle as the larger scale laboratory based equipment, where a small drill head is placed against the rock surface and progression through the stone vs pressure exerted is measured via specialised software connected to the drill.



Figure 7: Resistance drilling in action. Photo: Laboratorio de Petrofísica del Instituto de Geociencias IGEO (CSIC-UCM).

Rock surface hardness and weathering rates; a word of caution

While it is possible to make a straightforward correlation between weathering and surface hardness, as many studies (such as the ones referenced in this article) have successfully shown, a word of caution is needed. There are a number of factors that influence rock surface hardness readings that need to be taken into consideration. For example, water content, flake formation and hardening and the presence of microbial colonies can all dampen the impact in which case the readings are as much a representation of the environmental conditions (such as a recent rainfall) as the actual deterioration of the surface. Researchers should therefore always take note of any irregularities

observed on the rock surface that could influence hardness readings and combine rock surface hardness measurements with, for example, saturation measurements.

Concluding remarks

This chapter gives a brief overview of rock surface hardness methods and theory. The tools highlighted in this chapter are very suitable for monitoring and mapping rock surface weathering, though the drawbacks have been outlined. It has to be stressed that the selection of the method is of utmost importance in light of acceptable noise in the data set, size of the sample area and fragility of the surface. If applied properly, these methods can give a good insight into weathering processes and their impact on surface deterioration.

References

- Aoki H, Matsukura Y. 2008. Estimating the unconfined compressive strength of intact rocks from Equotip hardness. *Bulletin of Engineering Geology and the Environment*, **67** (1): 23 – 29.
- Buj O, Gisbert J, McKinley JM, Smith B. 2011. Spatial characterization of salt accumulation in early stage limestone weathering using probe permeametry. *Earth Surface Processes and Landforms*, **36** (3): 383 – 394.
- Cnudde V, Silversmit G, Matthieu B, Dewanckele J, De Samber B, Schoonjans T, Van Loo D, De Witte Y, Elburg M, Laszlo V, Van Hoorebeke L, Jacobs P. 2009. Multi-disciplinary characterisation of a sandstone crust. *Science of the Total Environment*, **407**: 5417 – 5427.
- Coombes MA, Feal-Pérez A, Naylor L, Wilhelm K. 2013. A non-destructive tool for detecting changes in the hardness of engineering materials: Application of the Equotip durometer in the coastal zone. *Engineering Geology*, **167**: 14 – 19.
- Goudie AS. 2006. The Schmidt Hammer in geomorphological research. *Progress in Physical Geography*, **30**: 703 – 718.
- Goudie AS. 2013. The Schmidt Hammer and Related Devices in Geomorphological Research. In: Shroder J, Switzer AD, Kennedy DM. (Eds.) *Treatise on Geomorphology*. Academic Press, San Diego, CA, Vol 14, pp 338 – 345.
- Hansen CD, Meiklejohn KI, Nel W, Loubser MJ, Van Der Merwe BJ. 2013. Aspect-controlled weathering observed on a blockfield in Dronning Maud Land, Antarctica. *Geografiska Annaler: Series A, Physical Geography*, **95** (4): 305 – 313.
- Kellerer-Pirklbauer A, Wangensteen B, Farbrót H, Etzelmüller B. 2007. Relative surface age-dating of rock glacier systems near Hólar in Hjaltadalur, northern Iceland. *Journal of Quaternary Science*, **23**: 137 – 151.
- Matsuoka N. 2001. Microgelivation versus macrogelivation: towards bridging the gap between laboratory and field frost weathering. *Permafrost and Periglacial Processes*, **12** (3): 299 – 313.
- Matsukura Y, Hattanji T, Oguchi CT, Hirose T. 2007. Ten year measurements of weathering rates of rock tablets on a forested hillslope in a humid temperate region, Japan. *Zeitschrift für Geomorphologie, Supplementary Issues*, **51** (1): 27 – 40.
- McCabe S, Brimblecombe P, Smith BJ, McAllister D, Srinivasan S, Basheer PA. 2013. The use and meaning of ‘time of wetness’ in understanding building stone decay. *Quarterly Journal of Engineering Geology and Hydrogeology*, **46**: 469 – 476.
- McCarroll D. 1991. The Schmidt Hammer, weathering and rock surface roughness. *Earth Surface Processes and Landforms*, **16**: 477 – 480.
- Mol L. 2014. Investigations into the relationship between changes in internal moisture regimes and rock surface deterioration in cavernous sandstone features. *Earth Surface Processes and Landforms*, **39** (7), 914 – 927.
- Mol L, Viles HA. 2012a. Conserving History in Changing Arid Environments. In: Mol L, Sternberg T (Eds.) *Changing Deserts; integrating people and their environment*. Cambridge, White Horse Press, pp 215 – 236.
- Mol L, Viles H. 2012b. The role of rock surface hardness and internal moisture in tafoni development in sandstone. *Earth Surface Processes and Landforms*, **37** (3): 301 – 314.

Moses C, Robinson D, Barlow J. 2014. Methods for measuring rock surface weathering and erosion: A critical review. *Earth-Science Reviews*, **135**: 141 – 161.

Rusakov NG, Madrovi PI. 1968. Determination of the strength of broken rock by the impact method. *Soviet Mining*, **4 (6)**: 571 – 575.

Smith BJ, Srinivasan S, Gomez-Heras M, Basheer PA, Viles H. 2011. Near-surface temperature cycling of stone and its implications for scales of surface deterioration. *Geomorphology*, **130 (1-2)**: 76-82.

Sumner P, Nel W. 2002. The effect of rock moisture on the Schmidt Hammer rebound: tests on rock samples from Marion Island and South Africa. *Earth Surface Processes and Landforms*, **27 (10)**: 1137 – 1142.

Török A, Siegesmund S, Müller C, Hüpers A, Hoppert M, Weiss T. 2007. Differences in texture, physical properties and microbiology of weathering crust and host rock: a case study of the porous limestone of Budapest (Hungary). In: Pýrkryl R, Smith BJ (Eds.) *Building Stone Decay: From Diagnosis to Conservation*. Geological Society, London, Special Publications 271: 261 – 276.

Viles HA. 2001. Scale issues in weathering studies. *Geomorphology*, **41 (1)**: 63 – 72.

Viles HA, Goudie A. 2004. Biofilms and case hardening on sandstones from Al-Quwayra, Jordan. *Earth Surface Processes and Landforms*, **29**: 1473 – 1485.

Viles H, Goudie A, Grab S, Lalley J. 2011. The use of the Schmidt Hammer and Equotip for rock hardness assessment in geomorphology and heritage science: a comparative analysis. *Earth Surface Processes and Landforms*, **36 (3)**: 320 – 333.

White AF, Brantley S. 2003. The effect of time on weathering of silicate minerals: why do weathering rates differ in the laboratory and field? *Chemical Geology*, **202 (3-4)**: 479 – 450.

1.4.1. Environmental Magnetism: Sediment Source Tracing

Sophie C. Sherriff^{1,2}

¹ School of the Environment, University of Dundee, Dundee, DD1 4HN, Scotland
(s.c.sherriff@dundee.ac.uk)

² Johnstown Castle Environment Research Centre, Teagasc, Johnstown Castle, Wexford, Ireland



ABSTRACT: Environmental magnetism is a technique used to analyse the natural characteristics of materials. The mineral magnetic array exhibited by a sample e.g. a single rock-based or composite soil, sediment or core material, are differentiated according to the iron oxide assemblage. Such accumulations of iron oxides principally reflect the parent material characteristics, but also any subsequent environmental processes e.g. erosion, deposition, and biogenic processes. Consequently, environmental magnetism has been utilised in sediment source tracing investigations. Sediment accumulations in impacted locations, i.e. ‘*targets*’ such as rivers and lakes, are assumed to possess the characteristics of the soils within the contributing catchment area. Mineral magnetic measurements, including magnetic concentration, mineralogical composition and magnetic grain size, can therefore be used as sediment tracers to indicate subtle differences between potential source samples. The contribution of sources to the target is determined using multivariate statistical analysis or un-mixing models. The results of source tracing studies are essential to inform cost-effective management strategies targeted at reducing soil erosion and improving water quality.

KEYWORDS: environmental magnetism, remanence, susceptibility, sediment sources

Introduction

Environmental magnetism utilises the mineral magnetic behaviour of a material to interpret the environmental processes acting upon it. The magnetic character is dependent upon the assemblage of minerals within a material. Although all materials will respond to magnetic fields, iron oxides are particularly sensitive. The near-universal occurrence of iron oxides in the environment e.g. haematite, magnetite and goethite, provides many opportunities to apply the technique. The assemblage of magnetic minerals is primarily dependent upon the parent material, specifically rock-formation, and any subsequent modification due to geomorphic processes e.g. transportation, deposition and post-depositional processes (Liu *et al.*, 2012). Due to the range of electron structures within different oxides, responses from each type will differ when exposed to externally applied magnetic fields, and this enables

discrimination between contrasting sources and processes.

The principles of environmental magnetism, initially introduced as a research area by Thompson and Oldfield (1986), were founded on rock magnetic techniques. Since then, the approach has been applied to a variety of applications (Table 1), including: estuarine and fluvial sediment source tracing (Caitcheon, 1993; Jenkins *et al.*, 2002); reconstruction of ice extent in glacial environments (Walden and Ballantyne, 2002); climatic reconstruction of loess sequences (Heller and Evans, 1995); reconstruction of paleo-environmental geomorphology (Sun *et al.*, 2011); classification of soils (Booth *et al.*, 2005); and, pollution studies to identify spatial patterns of vehicle pollution (Maher *et al.*, 2008).

Table 1: Summary of environmental magnetism studies for multiple applications

Application	
Glacial	Nolan <i>et al.</i> (1999), Walden and Ballantyne (2002), Gurney and White (2005), Ojala <i>et al.</i> (2011)
Palaeoclimatic	Sagnotti <i>et al.</i> (1998), Dearing <i>et al.</i> (2001), Gathorne-Hardy <i>et al.</i> (2009), Bradák <i>et al.</i> (2011)
Paleomagnetism	Zijderveld <i>et al.</i> (1991), Dupont-Nivet <i>et al.</i> (2002), Lisé-Pronovost <i>et al.</i> (2009), Sun <i>et al.</i> (2011)
Pollutants	Maher <i>et al.</i> (2008), Blundell <i>et al.</i> (2009), Zhang <i>et al.</i> (2008a), Wang (2013), Crosby <i>et al.</i> (2014)
Sediment sources	Caitcheon (1993), Foster and Walling (1994), Caitcheon (1998), Jenkins <i>et al.</i> (2002), Hounslow and Morton (2004), Blake <i>et al.</i> (2006)
Soil	Maher (1998), Booth <i>et al.</i> (2005), Neff <i>et al.</i> (2005), Hannam and Dearing (2008), Lourenço <i>et al.</i> (2014)

containing iron or iron-based substances, e.g. iron-oxides and iron-sulphides. When exposed to a small magnetic field, the concentration of magnetic material termed the magnetic susceptibility (χ) is measured. Characteristics of χ for different magnetic behaviours are summarised in Table 2, however, composite samples such as soils and sediments will vary according to the constituent minerals (Hatfield and Maher, 2008).

On application of stronger magnetic field, ferromagnetic minerals additionally respond such that they hold a permanent magnetic moment or natural remanent magnetisation. A sample cannot return to the original pre-field magnetisation measurement (Dekkers, 1997). As field strength increases, remanent magnetisation also increases until it reaches 'saturation', i.e. the maximum magnetisation possible which is known as the Saturation Isothermal Remanent Magnetisation (SIRM). Once the material reaches the saturation magnetisation, an 'in-field' hysteresis loop (Figure 1) is established. To remove the remanence, de-magnetisation must be conducted (Walden, 1999a). The magnetic field required to return a sample to full demagnetisation is defined as the coercivity.

Mineral magnetism provides great utility for techniques such as source tracing studies. Identification of sources of fluvial sediment is crucial to inform targeted and cost-effective soil and water quality mitigation programmes. Characteristics of sediments collected from a catchment outlet, e.g. lake or river, are compared to the soils within the contributing catchment area and the source provenance un-mixed. By determining the principal sources, these areas may be targeted for management. This chapter summarises the principles of magnetic behaviour and application of the technique to fluvial source tracing.

Mineral magnetism

Three main types of magnetic behaviour exist; diamagnetism, paramagnetism and ferromagnetism (Smith, 1999). Diamagnetic minerals include those lacking in iron such as calcite, quartz and halite (Booth *et al.*, 2005), organic substances and water (Maher *et al.*, 2009). Paramagnetic minerals, for example olivine, biotite and pyrite may contain iron in their structures (Dearing, 1999). Ferromagnetic behaviour occurs in minerals

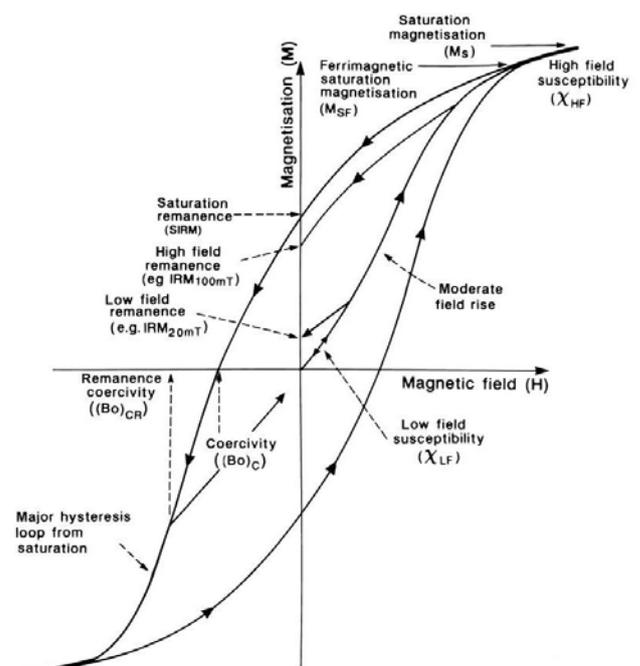


Figure 1: Theoretical remanence acquisition and hysteresis of a ferrimagnetic mineral denoting key susceptibility and iso-thermal remanence terminology. (From Maher *et al.*, 2009).

Table 2: Response of magnetic mineral sub-classes to susceptibility and remanence measurements. (Summarised from Dearing, 1999; and Smith, 1999).

	Susceptibility ($10^{-6} \text{ m}^3 \text{ kg}^{-1}$)	Remanence
Diamagnetism	Weak negative magnetisation (<0)	No remanence
Paramagnetism	Weak positive magnetisation ($0.01-1$)	No remanence
Canted anti-ferromagnetism (magnetically hard)	Moderate positive magnetisation ($0.27-1.19$)	Low fields: little or no remanent magnetisation High fields: 300 mT required to obtain response
Ferrimagnetism (magnetically soft)	Strong positive magnetisation ($50-1116$)	Remanent magnetisation exhibited at lower applied fields e.g. <100 mT
Ferromagnetism	Very strong positive magnetisation ($68850-276000$)	Abundance limited in natural samples

The character of remanence acquisition in a ferromagnetic material alters according to the mineralogy and the domain arrangement within it. In general, magnetically 'softer' minerals known as ferrimagnets (e.g. magnetite) exhibit a large proportion of the SIRM in lower applied fields. Magnetically 'harder' materials known as canted anti-ferromagnets (e.g. haematite) require much higher applied fields to achieve a substantial remanent response (Table 2). As such, multiple fields of measurement are investigated to determine the mineralogical composition.

Magnetic domains or grain size additionally affects remanence measurements. Five grain size classes are defined (from the finest): super-paramagnetic, viscous magnetism, stable single domain, pseudo-single domain, and multi-domain (Smith, 1999). For a single mineral larger, multi-domain grains will respond to remanence measurements differently to smaller, single domain grains which are more resistant to external current (Peters and Dekkers, 2003; Maher and Thompson, 2009). An hysteretic remanent magnetisation (ARM) is used to detect ultra-fine magnetic mineral grains (Banerjee *et al.*, 1981) and has been utilised in multiple investigations.

Application of technique

Tracing sediment sources

Sediment source tracing is a geomorphological technique to determine the provenance of sediments within a catchment. The method has been applied on a series of scales, from catchments less than 15 km^2 (Gruszowski *et al.*, 2003) to those greater than $100,000 \text{ km}^2$ (Maher *et al.*, 2009). Sediment sources are determined according to their particulate characteristics, therefore the range of parameters or 'tracers' measured by mineral magnetics provides a suitable technique for source tracing.

A number of assumptions underlie sediment tracing techniques and must be considered: (1) a tracer must distinguish at least two sources; (2) a tracer is transported with sediment and not altered by selective erosion; (3) source properties are stable over time; and, (4) no post-depositional tracer modification (Foster and Lees, 2000).

Desktop study

Source tracing programmes require subdivision of a contributing catchment area into definable and contrasting source areas. The specific research hypothesis will dominate source area definition, e.g. to target sub-catchment basins, erosion processes or using a pilot study to determine appropriate source areas. In order to effectively sub-divide, it is essential to understand the sources and generation of magnetic minerals in the environment.

Naturally, the variety of abundance and characteristics of magnetic materials occurs due to the combined effect of parent materials and secondary processes. The formation of crystals in igneous rocks is dependent on the speed of magma cooling, subsequently determining the exact components of metallic compounds. In metamorphic and sedimentary rocks however, the materials available for formation, in addition to the nature and intensity of formation, will determine the magnetic character (Liu *et al.*, 2012). Consequently, contrasting catchment geology will create a primary magnetic 'imprint'.

Field assessment

In some cases, parent signatures are sufficient to distinguish between soil types or lithologies (Booth *et al.*, 2005; Hatfield and Maher, 2008). However, a range of secondary processes may affect the consistency of the magnetic signature (e.g. ex-situ material delivery), bio-mineralisation, in-situ authigenesis, and anthropogenic additions (Thompson *et al.*, 1980; Hilton *et al.*, 1986; Dekkers, 1997).

Bio-mineralisation, for example, occurs in topsoils which enhance the fine ferrimagnetic concentration in water-logged soils or those frequently exposed to wetting and drying cycles (Maher and Taylor, 1988; Fassbinder *et al.*, 1990). Burning of vegetation for vegetation control or by wildfire increases the fine magnetic component in topsoils due to extreme soil temperature (Blake *et al.*, 2006). Such secondary signatures modify the primary, parent material signature, providing either an opportunity to define a source according to such characteristics, or adding complication to source interpretations. Additionally, small quantities of strong magnetic material may overpower the signal of weaker minerals, in some cases this will make the technique unsuitable for the particular study (Lees, 1999). It is therefore essential to complete a visual field assessment and a preliminary magnetic susceptibility study of the monitoring area, to determine which areas may be dominated by secondary magnetic processes (Lees, 1999).

Sample collection

Collection of source samples must be carefully conducted to prevent contamination

of the natural mineral magnetic signal. Non-metallic equipment must be used where possible, and where unavoidable, potentially contaminated particles in contact with a metallic surface must be removed. Relevant site specific information such as, in the case of soils, the sampled horizon, which may explain primary or secondary mineral forming processes, should be recorded in order to aid later interpretation. Where surface measurements are made with a field-sensor, high vegetation cover and soil moisture may cause fluctuations in results due to the addition of diamagnetic materials. Source areas must be adequately sampled in order to fully determine the characteristics of each area. Where intra- and inter- unit variability overlap, the definitive capabilities of the final result are compromised (Lees, 1999).

Target sample collection for un-mixing is dependent on the depositional environment. Suspended sediments, river-bed, lake, estuarine and sea-shelf deposits have all been collected in source provenance studies (Collins *et al.*, 1998; Jenkins *et al.*, 2002; Collins and Walling, 2007; Maher *et al.*, 2009).

Signature modification may also occur in the post-depositional environment due to the bacterial formation of magnetite. Jelinowska *et al.* (1997) discovered that bacteria in anoxic lake sediments reduced iron-oxides such that the magnetic domains were irreversibly altered. Maher *et al.* (2009) used acid dissolution to remove secondary magnetic material to expose particles within host grains of target samples, i.e. the primary magnetic signature for source tracing. Measurements performed on treated and un-treated samples were found to be distinctly different, supporting the concept that the impact of secondary magnetic processes is significant.

Post-collection, samples must be made airtight and refrigerated to prevent biological activity altering the mineral magnetic composition. Depending on ambient temperature during field sampling, it may be sensible to store samples in a cool box to prevent temperature related modifications of the magnetic signature.

Sample preparation

Samples can be processed as a composite sample or core. Where samples need to be

dried: air-drying, freeze-drying, or oven drying below 40°C is advised. Temperatures exceeding this threshold may modify the natural magnetic assemblage (Walden, 1999b). Manual grinding is recommended as mechanical grinders pose greater risk of cross contamination and mineral breakdown. Sink samples, particularly those in suspension require freeze-drying, centrifuging or sedimentation methods to extract particulates for analysis.

Particle size separation should also be considered for target samples. Hatfield and Maher (2008) sub-divided suspended sediment samples into four particle size classes (<2 µm, 2-8 µm, 8-31 µm, 31-63 µm) to determine inconsistencies in magnetic imprint. The 8-31 µm and 31-63 µm classes were found to be most effective for source tracing, whereas <2 µm and 2-8 µm reflected post-depositional bacterial magnetite formation.

The majority of magnetic analysis instruments, (such as the MS2 laboratory sensor, pulse magnetiser, and magnetometer) accept samples in 10 cc sample pots. It should be ensured that the sample completely fills the pot to prevent movement of particles during analysis, which may produce spurious results.

Where sample quantity is limited, the sample can be immobilised in cling film and packed using cotton wool. Some instruments cannot accommodate sample lids and tape can be used to seal the pot. For core based samples, specific equipment attachments can be used for magnetic susceptibility measurements or non-metallic blades used to split cores into sub-sections.

Where mass specific measurements are intended, each sample must be weighed and recorded. Once the sample pot is filled and clearly labelled, it is helpful to draw an axis direction on each pot. This will aid Advanced Remanence Measurements (ARM) and Iso-thermal Remanence (IRM) measurements which are made on, or reversed to, a single axis.

Measurement

Mineral magnetic measurements must be approached systematically with increasing field strength (Figure 2). Units are generally reported as volume specific magnetisation, however conversion to mass specific measurements are beneficial for interpretation of results and comparison with the wider literature.

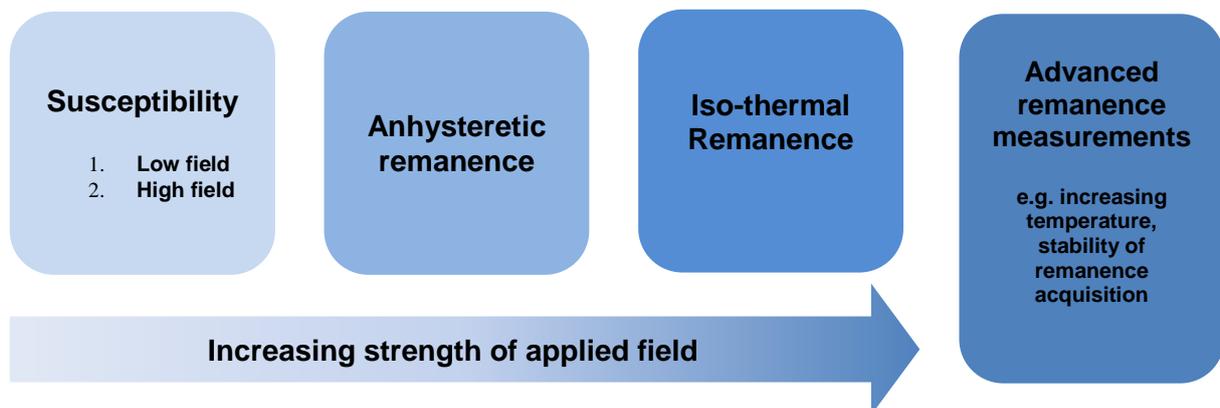


Figure 2: Schematic diagram showing order of magnetic measurements and relative field strength

Magnetic susceptibility

Magnetic susceptibility meters (e.g. MS2/MS3 Bartington Instruments Ltd shown in Figure 3) are used to collect lab- and field- based susceptibility data. Magnetisation is reported as volume specific (dimensionless κ), mass specific susceptibility, χ , in SI units of

$10^{-6} \text{ m}^3 \text{ kg}^{-1}$ are converted, shown by Equation 1:

$$\text{Mass specific susceptibility} = \frac{\kappa \times 10}{\text{Sample mass}} \quad (\text{Eq. 1})$$

Where the sample mass is in grams. The laboratory based sensor features both low- ($\chi_{\text{LF}}=0.465 \text{ kHz}$) and high- ($\chi_{\text{HF}}=4.65 \text{ kHz}$)

susceptibility frequencies and two sensitivity categories. The frequency dependent susceptibility ($\% \chi_{FD}$) is a ratio between χ_{HF} and χ_{LF} measurements that reflects the super-paramagnetic grain size component (Blundell *et al.*, 2009) and is calculated by Equation 2:

$$\left(\frac{\chi_{LF} - \chi_{HF}}{\chi_{LF}} \right) \times 100 \quad (\text{Eq.2})$$

Magnetic susceptibility equipment must be situated in a magnetically stable environment. Readings are sensitive, therefore, care must be taken to remove metallic or electronic components in the vicinity of the sensor. To an extent, background interference can be accounted for by 'zeroing' the sensor before a sample is measured and taking a subsequent air-measurement to monitor and account for drift (Dearing, 1999).

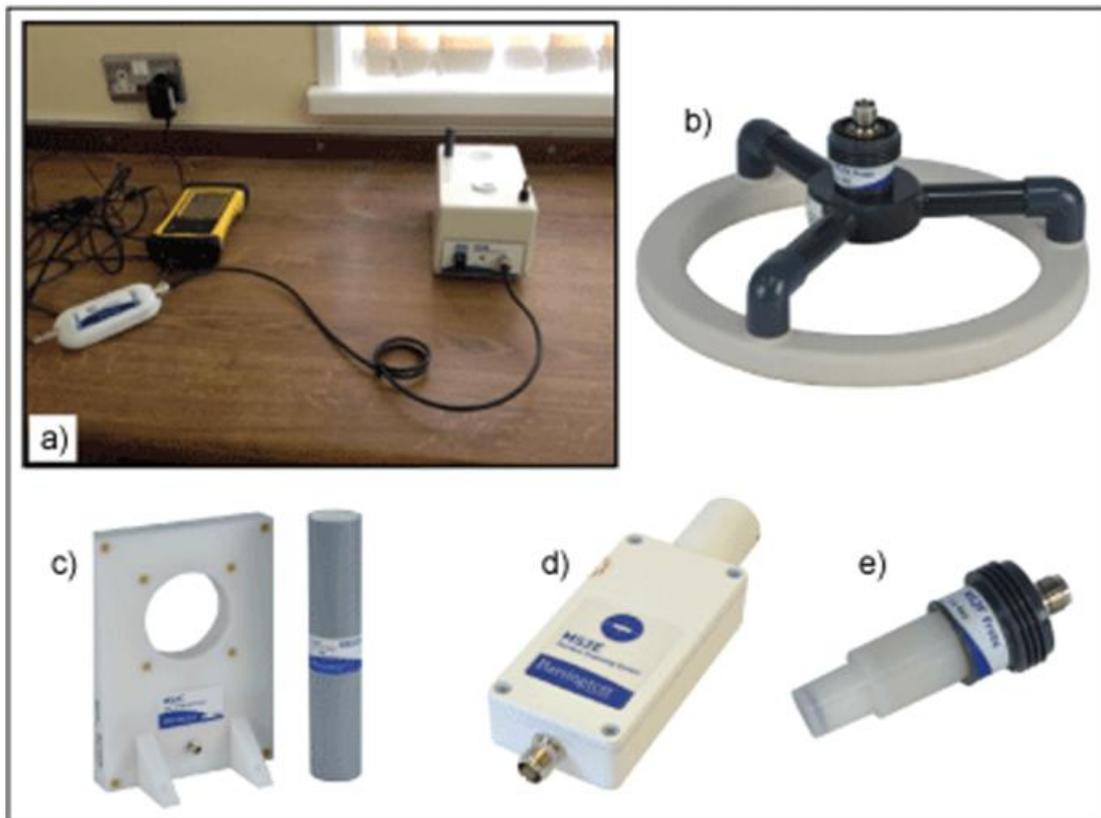


Figure 3: Bartington MS3 magnetic susceptibility meter with attachments: a) laboratory sensor, b) surface sensor, c) core scanning sensor, d) core logging sensor, e) surface point probe (The images for b-e are taken from www.bartington.com).

Remanence acquisition

Isothermal remanence measurements are made using two instruments: (1) a magnetiser or electromagnet to apply a magnetic field; and, (2) a magnetometer to measure the intensity of magnetisation (Walden, 1999a). The SIRM is commonly defined as 1 Tesla. It is sufficient to distinguish mineralogical differences between soft and hard materials, however, if full saturation of hard minerals is required, a larger field may be required. Values are most commonly reported in 10^{-3} Am^{-1} . Conversion to mass specific measurements uses Equation 3 below:

$$\text{Mass specific magnetisation} = \frac{\text{Magnetisation} \times 1.29}{\text{Sample weight}} \quad (\text{Eq. 3})$$

Where sample weight is in grams and the mass specific magnetisation is reported in units of $10^{-5} \text{ Am}^2 \text{ kg}^{-1}$.

Measurement of multiple fields is recommended between zero and SIRM, to characterise the remanence acquisition curve. The exact number is dependent upon study requirements, access to facilities and time (Walden, 1999a). These intermediate fields can be performed as forward or backward fields. Forward fields are repeated at increasing strength from low ($\sim 20 \text{ mT}$) through medium ($\sim 100 \text{ mT}$) to high ($\sim 300 \text{ mT}$) denoted as $\text{IRM}_{20\text{mT}}$, $\text{IRM}_{100\text{mT}}$ and $\text{IRM}_{300\text{mT}}$ respectively.

Table 3: Summary of common interpretations of mineral magnetic parameters (Summarised from: Blake et al., 2006; Blundell et al., 2009; Booth et al., 2005; Dekkers, 1997; Liu et al., 2012; Maher, 1998; Maher and Taylor, 1988; Walden, 1999a; Walden and Ballantyne, 2002).

Parameter	Description	Examples
Concentration		
Magnetic susceptibility Mass specific χ	Concentration of magnetic material	a,b,c,d,e,f
SIRM Mass specific SIRM	Concentration of remanence holding material (ferromagnetic grains)	a,b,c,d,e,f
Domain state		
Frequency dependent susceptibility $\% \chi_{FD}$	Proportion of super-paramagnetic magnetite	a,b,d,f
Susceptibility of anhysteretic remanence χ_{ARM}	Ultra-fine magnetite particles near to the super- paramagnetic/single domain boundary	a,b,c,f
$\chi_{ARM}/SIRM$	Relative grain size of magnetite, high values suggest superparamagnetic grains whereas low values indicate coarser multi domain minerals	b,e,f
χ_{ARM}/χ_{LF}	Stable single domain ferromagnetic grains- super paramagnetic and fine viscous grains can affect parameter	b,f
Mineralogical composition		
Forward IRM fields Mass specific ($IRM_{x_{mT}}$)	Abundance of minerals capable of remanence acquisition at the denoted field Low fields (<100 mT) magnetically 'soft' material High fields (>300 mT) magnetically 'hard' material	b,c,f
IRM ratios (%) Forward field/SIRM	Percentage of SIRM acquired at the specified field. Large values at low fields indicate 'soft' minerals. Increases at higher fields indicated 'hard' minerals.	e
Hard IRM (HIRM) $SIRM-IRM_{300mT}$	Mass specific indication of hard magnetic minerals	d
Backfield ratios	Demagnetisation of remanence, magnetically 'soft' minerals demagnetise at low backfield fields, 'hard' minerals demagnetise at higher applied backfields	a,d
S-ratio $IRM-100_{mT}/SIRM$	Ratio of saturated to non-saturated minerals at - 100mT. Range of values -1 to +1; values close to -1 are dominated by 'soft' minerals, higher values reflect an increasing 'hard' component.	b,d,f
$SIRM/\chi_{LF}$	Proportion of ferromagnetic material in a sample, lower ratio suggests a greater proportion of diamagnetic and paramagnetic minerals	b

* Examples of application in source tracing studies: a. Russell et al. (2001); b. Jenkins et al. (2002); c. Hatfield and Maher (2008); d. Zhang et al. (2008b); e. Maher et al. (2009); f. Blake et al. (2006)

Conversely, backfield remanent magnetism can be investigated by initially imparting a forward saturating field to obtain the SIRM, and subsequent applications of fields in the opposite direction, i.e. 180° of the primary axis (Walden, 1999a).

Anhyseretic remanence uses a demagnetiser, with ARM attachment to impart the measurement and a magnetometer to determine the magnetisation intensity (Walden, 1999a). Equation 3 is similarly used to determine the mass specific ARM ($10^{-5} \text{ Am}^2 \text{ kg}^{-1}$) and when normalised by the biasing field strength (A/m), is termed the anhyseretic susceptibility (χ_{ARM}) reported in m^3/kg .

The suite of parameters described can be measured quickly with little sample preparation. The parameters measured, and ratios between them, are used to interpret the characteristics of each sample (Table 3).

Statistical analysis

Analysis methods vary from bi-plot and multivariate techniques such as cluster, factor and principal components analysis to determine source area grouping (Booth *et al.*, 2005; Hatfield and Maher, 2008). Contributions from source areas are frequently determined using un-mixing models (Collins *et al.*, 1998; Franks and Rowan, 2000). Lees (1999) summarises statistical techniques in source tracing exercises.

Case study

Mineral magnetic source tracing was carried out in a small (10 km^2) intensively agricultural fluvial catchment in Ireland. Potential source areas were selected due to their agricultural or non-agricultural significance, consequently seven sources were sampled; field topsoils, field subsoils, on-farm tracks, road verges, channel banks, sub-surface drains outlets, and open field drains. Channel bank, field drain, track, drain outlet and road verge samples were taken using a plastic trowel to collect a surface scrape. Field topsoils used an auger to consistently sample the top 5 cm of soil and a second screw corer to sample at 40 cm for subsoil samples. Target samples were collected at the catchment outlet using time integrated suspended sediment samplers (TISS) (Phillips *et al.*, 2000), which were emptied and re-installed at 6-12 week intervals.

Source samples were returned to the laboratory and dried at $<40^\circ\text{C}$ in plastic trays. Once dried, samples were sieved to $125 \mu\text{m}$ to replicate the particle size distribution of TISS samples. TISS samples were refrigerated for a minimum of three days to allow sedimentation of particles. Water was siphoned without disturbing the deposited sediment. The remaining sediment and water mixture were dried at $<40^\circ\text{C}$ and manually ground using a mortar and pestle once dry.

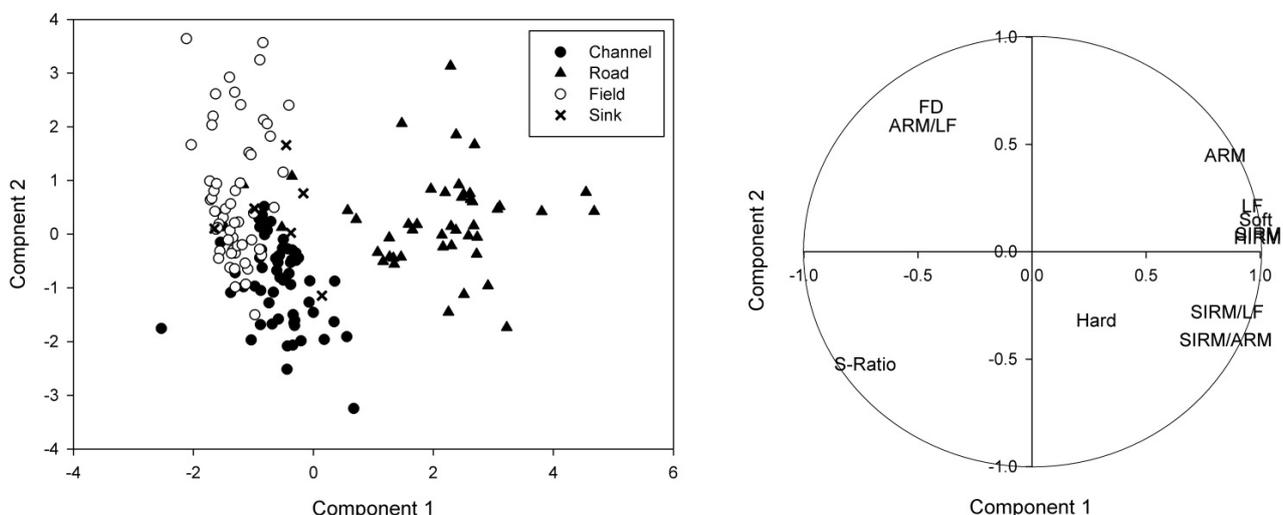


Figure 4: Principal components analysis of soil and sediment samples for source area ascription of the case study: a) score plot, and b) loading plot.

All samples were weighed and packed into plastic 10 cc pots either full of the sample, or where quantities were low, immobilised in cling film and cotton wool. Samples were analysed for χ_{LF} , χ_{HF} , χ_{ARM} , SIRM, bIRM_{40mT}, bIRM_{100mT}, and bIRM_{300mT}. The parameters χ_{LF} , % χ_{FD} , χ_{ARM} , SIRM, SIRM/ χ_{LF} , SIRM/ χ_{ARM} , χ_{ARM} / χ_{LF} , bIRM_{40mT}, bIRM_{300mT}, S-ratio and HIRM were calculated and used for statistical analysis.

Multivariate analysis showed that the seven source areas could not be satisfactorily distinguish and therefore three parent groups were identified; fields (topsoil, subsoil and sub-surface drains), roads (road verges and tracks) and channel (channel banks and open drains). Multiple discriminant analysis showed 96.3% discrimination capability with all eleven tracers (SPSS v18). Principal components analysis performed on JMP 9.0 showed that the three parent source groups are well defined, and the loading diagram showing the arrangement of parameters to produce the score plot (Figure 4a). Target sink values were enclosed within the spatial limits of the source areas, indicating that post-depositional or transportation effects were not apparent, or significant enough to affect sample composition. We can therefore conclude that suspended sediments are predominantly comprised of material from field and channel bank sources.

Acknowledgements

This chapter was completed whilst funded by the Walsh Fellowship Programme, Teagasc. The author acknowledges the Walsh Fellow Overseas Travel Award for funding, and Dr John Walden for hosting a placement at the Environmental Magnetism Laboratory at the University of St Andrews. The author would also like to thank two anonymous reviewers whose comments helped to improve the article.

References

- Banerjee SK, King J, Marvin J. 1981. A rapid method for magnetic granulometry with applications to environmental studies. *Geophysical Research Letters* **8**: 333–336.
- Blake WH, Wallbrink PJ, Doerr SH, Shakesby RA, Humphreys GS. 2006. Magnetic enhancement in wildfire-affected soil and its *Earth Surface Processes and Landforms* **31**: 249-264.
- Blundell A, Dearing JA, Boyle JF, Hannam JA. 2009. Controlling factors for the spatial variability of soil magnetic susceptibility across England and Wales. *Earth Science Reviews* **95**: 158-188.
- Booth CA, Fullen MA, Walden J, Smith JP, Hallett MD, Harris J, Holland K. 2005. Magnetic Properties of Agricultural Topsoils of the Isle of Man: Their Characterization and Classification by Factor Analysis. *Communications in Soil Science and Plant Analysis* **36**: 1241-1262.
- Bradák B, Thamó-Bozsó E, Kovács J, Márton E, Csillag G, Horváth E. 2011. Characteristics of Pleistocene climate cycles identified in Cérna Valley loess–paleosol section (Vértesacsá, Hungary). *Quaternary International* **234**: 86-97.
- Caitcheon GG. 1993. Sediment source tracing using environmental magnetism: A new approach with examples from Australia. *Hydrological Processes* **7**: 349-358.
- Caitcheon GG. 1998. The significance of various sediment magnetic mineral fractions for tracing sediment sources in Killimicat Creek. *Catena* **32**: 131-142.
- Collins AL, Walling DE. 2007. Sources of fine sediment recovered from the channel bed of lowland groundwater-fed catchments in the UK. *Geomorphology* **88**: 120-138.
- Collins AL, Walling DE, Leeks GJL. 1998. Use of composite fingerprints to determine the provenance of the contemporary suspended sediment load transported by rivers. *Earth Surface Processes and Landforms* **23**: 31-52.
- Crosby CJ, Booth CA, Appasamy D, Fullen MA, Farr K. 2014. Mineral magnetic measurements as a pollution proxy for canal sediments (Birmingham Canal Navigation Main Line) *Environmental Technology (United Kingdom)* **35**: 432-445.
- Dearing J. 1999. Magnetic susceptibility. In: Walden J, Oldfield F, Smith JP (eds). *Environmental magnetism: a practical guide. Technical Guide, No 6. Quaternary Research Association, London.*
- Dearing J, Livingstone IP, Bateman MD, White K. 2001. Palaeoclimate records from OIS 8.0–5.4 recorded in loess–paleosol sequences on the Matmata Plateau, southern

- Tunisia, based on mineral magnetism and new luminescence dating. *Quaternary International* **76–77**: 43-56.
- Dekkers, MJ. 1997. Environmental magnetism: an introduction. *Geologie en Mijnbouw* **76**: 163-182.
- Dupont-Nivet G, Guo Z, Butler RF, Jia . 2002. Discordant paleomagnetic direction in Miocene rocks from the central Tarim Basin: evidence for local deformation and inclination shallowing. *Earth and Planetary Science Letters* **199**: 473-482.
- Fassbinder JWE, Stanjek H, Vali H. 1990. Occurrence of magnetic bacteria in soil. *Nature* **343**: 161-163.
- Foster IDL, Walling DE. 1994. Using reservoir deposits to reconstruct changing sediment yields and sources in the catchment of the Old Mill Reservoir, South Devon, UK, over the past 50 years. *Hydrological Sciences* **39**: 347-368.
- Foster IDL, Lees JA. 2000. Tracers in geomorphology: theory and applications in tracing fine sediments. In: Foster IDL (ed) *Tracers in Geomorphology* 3-20.
- Franks SW, Rowan JS. 2000. Multi-parameter fingerprinting of sediment sources: Uncertainty estimation and tracer selection. In: Bentley LR, Sykes JF, Gray WG, Brebbia CA, Pinder GF. (eds) *Computational Methods in Water Resources XIII*. pp 1067-1074.
- Gathorn-Hardy FJ, Erlendsson E, Langdon PG, Edwards KJ. 2009. Lake sediment evidence for late Holocene climate change and landscape erosion in western Iceland. *Journal of Paleolimnology* **42**: 413-426.
- Gruszowski KE, Foster IDL, Lees JA, Charlesworth SM. 2003. Sediment sources and transport pathways in a rural catchment, Herefordshire, UK. *Hydrological Processes* **17**: 2665-2681.
- Gurney SD, White K. 2005. Sediment magnetic properties of glacial till deposited since the Little Ice Age maximum for selected glaciers at Svartisen and Okstindan, northern Norway. *Boreas* **34**: 75-83.
- Hatfield RG, Maher BA. 2008. Suspended sediment characterization and tracing using a magnetic fingerprinting technique: Bassenthwaite Lake, Cumbria, UK. *The Holocene* **18**: 105-115.
- Hannam JA, Dearing JA. 2008. Mapping soil magnetic properties in Bosnia and Herzegovina for landmine clearance operations. *Earth and Planetary Science Letters* **274**:285-294.
- Heller F, Evans ME. 1995. Loess magnetism. *Reviews of Geophysics* **33**: 211-240.
- Hilton J, Lishman JP, Chapman JS. 1986. Magnetic and chemical characterization of a diagenetic magnetic mineral formed in the sediments of productive lakes. *Chemical Geology* **56**: 325–335.
- Hounslow MW, Morton AC. 2004. Evaluation of sediment provenance using magnetic mineral inclusions in clastic silicates: comparison with heavy mineral analysis. *Sedimentary Geology* **171**: 13-36.
- Jelinowska A, Tucholka P, Wieckowski K. 1997. Magnetic properties of sediments in a Polish lake: evidence of a relation between the rock-magnetic record and environmental changes in Late Pleistocene and Holocene sediments. *Geophysical Journal International* **129**: 727-736.
- Jenkins PA, Duck RW, Rowan JS, Walden J. 2002. Fingerprinting of bed sediment in the Tay Estuary, Scotland: an environmental magnetism approach. *Hydrology and Earth System Sciences* **6**: 1007-1016.
- Lees J. 1999. Evaluating magnetic parameters for use in source identification, classification and modelling of natural and environmental materials. In: Walden J, Oldfield F, Smith JP (eds). *Environmental magnetism: a practical guide*. Technical Guide, No 6. Quaternary Research Association, London.
- Lisé-Pronovost A, St-Onge G, Brachfeld S, Barletta F, Darby D. 2009. Paleomagnetic constraints on the Holocene stratigraphy of the Arctic Alaskan margin. *Global and Planetary Change* **68**: 85-99.
- Liu Q, Roberts AP, Larrasoana JC, Banerjee SK, Guyodo Y, Tauxe L, Oldfield F. 2012. Environmental magnetism: Principles and applications. *Reviews of Geophysics* **50**: 1-50.
- Lourenço AM, Sequeira E, Sant'Ovaia H, Gomes CR. 2014. Magnetic, geochemical and pedological characterisation of soil profiles from different environments and geological backgrounds near Coimbra, Portugal. *Geoderma* **213**: 408-418.

- Maher BA. 1998. Magnetic properties of modern soils and Quaternary loessic paleosols: paleoclimatic implications. *Palaeogeography, Palaeoclimatology, Palaeoecology* **137**: 25-54.
- Maher BA, Taylor RM, 1988. Formation of ultrafine-grained magnetite in soils. *Nature* **336**: 368-370.
- Maher BA, Thompson R. 2009. Quaternary Climates, Environments and Magnetism. Cambridge University Press: Cambridge.
- Maher BA, Moore C, Matzka J. 2008. Spatial variation in vehicle-derived metal pollution identified by magnetic and elemental analysis of roadside tree leaves. *Atmospheric Environment* **42**: 364-373.
- Maher BA, Watkins SJ, Brunskill G, Alexander J, Fieldings CR. 2009. Sediment provenance in a tropical fluvial and marine context by magnetic 'fingerprinting' of transportable sand fractions. *Sedimentology* **56**: 841-861.
- Neff JC, Reynolds RL, Belnap J, Lamothe P. 2005. Multi-Decadal Impacts of Grazing on Soil Physical and Biogeochemical Properties in Southeast Utah. *Ecological Applications* **15**: 87-95.
- Nolan SR, Bloemendal J, Boyle JF, Jones RT, Oldfield F, Whitney M. 1999. Mineral magnetic and geochemical records of late Glacial climatic change from two northwest European carbonate lakes. *Paleolimnology* **22**: 97-107.
- Ojala AEK, Kubischta F, Kaakinen A, Salonen V-P. 2011. Characterization of diamictos on the basis of their mineral magnetic properties in Murchisonfjorden, Nordaustlandet, Svalbard. *Sedimentary Geology* **233**: 150-158.
- Peters C, Dekkers MJ. 2003. Selected room temperature magnetic parameters as a function of mineralogy, concentration and grain size. *Physics and Chemistry of the Earth* **28**: 659-667.
- Phillips JM, Russell MA, Walling DE. 2000. Time-integrated sampling of fluvial suspended sediment: a simple methodology for small catchments. *Hydrological Processes* **14**:2589-2602.
- Russell MA, Walling DE, Hodgkinson RA. 2001. Suspended sediment source in two lowland agricultural catchments in the UK. *Journal of Hydrology* **252**: 1-24.
- Sagnotti L, Florindo F, Wilson GS, Roberts AP, Verosub KL. 1998. Environmental magnetism of lower Miocene strata from the CRP-1 core, McMurdo Sound, Antarctica. *Terra Antartica* **5**: 661-667.
- Smith J. 1999. An introduction to the magnetic properties of natural materials. In: Walden J, Oldfield F, Smith JP (eds). *Environmental magnetism: a practical guide*. Technical Guide, No 6. Quaternary Research Association, London.
- Sun D, Bloemendal J, Yi Z, Zhu Y, Wang X, Zhang Y, Li Z, Wang F, Han F, Zhang Y. 2011. Palaeomagnetic and palaeoenvironmental study of two parallel sections of late Cenozoic strata in the central Taklimakan Desert: Implications for the desertification of the Tarim Basin. *Palaeogeography, Palaeoclimatology, Palaeoecology* **300**: 1-10.
- Thompson R, Oldfield F. 1986. Environmental magnetism. Allen & Unwin, London.
- Thompson R, Bloemendal RW, Dearing JA, Oldfield F, Rummery TA, Stober JC, Turner GM. 1980. Environmental applications of magnetic minerals. *Science* **207**: 481-485.
- Walden J. 1999a. Remanence measurements. In: Walden J, Oldfield F, Smith JP (eds). *Environmental magnetism: a practical guide*. Technical Guide, No 6. Quaternary Research Association, London.
- Walden J. 1999b. Sample collection and preparation. In: Walden J, Oldfield F, Smith JP (eds). *Environmental magnetism: a practical guide*. Technical Guide, No 6. Quaternary Research Association, London.
- Walden J, Ballantyne CK. 2002. Use of environmental magnetic measurements to validate the vertical extent of ice masses at the Last Glacial Maximum. *Journal of Quaternary Science* **17**: 193-200.
- Wang XS. 2013. Heavy metal pollution in urban topsoils: Mineralogical analyses and magnetic characterization. *Environmental Earth Sciences* **70**: 3155-3161.
- Zhang C, Huang B, Piper JDA, Luo R. 2008a. Biomonitoring of atmospheric particulate matter using magnetic properties of *Salix matsudana* tree ring cores. *Science of the Total Environment* **393**: 177-190.
- Zhang W, Xing Y, Yu L, Feng H, Lu M. 2008b. Distinguishing sediments from the Yangtze and Yellow Rivers, China: a mineral magnetic approach. *The Holocene* **18**: 1139-1145.

Zijderveld JDA, Hilgen FJ, Langereis CG, Verhallen PJJM, Zachariasse WJ. 1991. Integrated magnetostratigraphy and biostratigraphy of the upper Pliocene-lower Pleistocene from the Monte Singa and

Crotone areas in Calabria, Italy. *Earth and Planetary Science Letters* **107**: 697-714.

The emerging use of Magnetic Resonance Imaging (MRI) for 3D analysis of sediment structures and internal flow processes

Heather Haynes¹ & William M. Holmes²

¹ Water Academy, Heriot-Watt University, Edinburgh, UK (h.haynes@hw.ac.uk)

² Glasgow Experimental MRI Centre, Wellcome Surgical Institute, University of Glasgow, Glasgow, UK (william.holmes@glasgow.ac.uk)



Magnetic Resonance Imaging (MRI) can be used for 3D analysis of small-scale porous media structure and internal flow-related processes. It offers notable advantages over traditional sediment sampling (e.g. cores or surface-based scanning) as it is capable of high spatio-temporal resolution of the full 3D volume, including the sub-surface. Similarly, compared to X-Ray tomography, the extensive catalogue of MR pulse sequences typically provides: faster capture for imaging dynamic fluid processes; greater flexibility in resolving chemical species or tracers; and a safer radiation-free methodology. To demonstrate the relevance of this technique in geomorphological research, three exemplar applications are described: porous media structure of gravel bed rivers; measurements of fluid processes within aquifer pores and fractures; and, concentration mapping of contaminants through sand/gravel frameworks. Whilst, this emerging technique offers significant potential for visualizing many other 'black-box' processes important to the wider discipline, attention is afforded to discussion of the present constraints of the technique in field-based analysis.

KEYWORDS: Magnetic Resonance Imaging; sediment structure; porosity; permeability; 3D analysis

Introduction

Traditional geomorphological techniques for analysing small-scale sediment structure are typically constrained to 1D or 2D approaches, such as coring, photography etc. Even where more advanced techniques are available (e.g. laser displacement scanning), these tend to be restricted to imaging the surface of the sediment bed in a manner preclusive of true 3D analysis of volumetric space and the sub-surface particle characteristics and packing arrangements. Using Magnetic Resonance Imaging (MRI) overcomes these limitations, providing researchers with a technique with which to provide novel 3D spatio-temporal data on the internal structure of opaque porous media and the related fluid exchange and chemical reactions occurring within.

To date, MRI has been widely applied in the study of both porous media and mass transport phenomena in research disciplines

such as biomedicine, separation science, food science, well logging, physical science, rheology, chemical engineering and petroleum engineering. This breadth of applications is well demonstrated in publications such as Huerlimann et al. (2008) and Fantazzini et al. (2011). Given that these studies have proven MRI's capability to non-invasively study sediment structure, advection and diffusion processes, molecular dynamics and chemical reactions, the technique is increasingly drawing attention from researchers involved in sedimentology and geomorphology. Recent examples include: monitoring porosities in geotechnical composites (Tyrologou et al. 2005); identifying sedimentary structures in seabed cores (Bortolotti et al. 2006); determining the permeability of rock fractures in aquifers (e.g. Nestle et al. 2003a); analysing the wetting of clays via diffusion (Vogt et al. 2002; Ohkubo & Yamaguchi 2007); visualising the mechanics of granular flows and fluidised

beds (e.g. Kawaguchi 2010); and assessing river bed structure (Kleinhans et al. 2008, Haynes et al. 2009). Whilst application of MRI to sediment research is recognised to be a science in its infancy, maturation of this technique may offer geomorphologists crucial quantitative insight into many of today's black-box sediment systems. This technical note therefore focuses on the current strengths and weaknesses of MRI, using examples directly relevant to geomorphology to highlight its capability and future potential.

Magnetic Resonance Imaging

The theory of magnetic resonance

Certain nuclei (^1H , ^{13}C , ^{23}Na , ^{31}P etc.) possess spin angular momentum, and hence a nuclear magnetic moment, or "spin". Though many nuclei can give an MR signal, only hydrogen nuclei (^1H) found in water (in its liquid form) provide sufficient signal for the practical use of MRI for sediments. When ^1H rich samples are placed in a static magnetic field (Figure 1), B_0 , they become polarized, resulting in a net magnetisation aligned (ie longitudinal) with the magnetic field. The net magnetisation exhibits precession about the static magnetic field at the Larmor Frequency, and will absorb and emit RF radiation at this resonant frequency.

By using an RF coil (Figure 1) tuned to resonate at the Larmor frequency, short pulses of RF radiation excite the nuclear spins, tipping the net magnetization into the plane transverse to B_0 . The precession of this transverse magnetization then induces an alternating current in the RF coil, giving the MR signal. Further, using magnetic field gradient coils (Figure 1) to linearly vary the magnetic field across the sample causes precession to occur at slightly different frequencies at different locations across the sample; this labels the spatial position of the nuclei and is the basis of MRI.

One important type of image in MRI is relaxation weighting, where the net magnetization returns to equilibrium following an RF pulse. This is described by the loss of transverse magnetisation (T_2 transverse relaxation) and the return of longitudinal magnetisation (T_1 longitudinal relaxation). T_1 and T_2 relaxation can result from the close proximity of fluid molecules to the pore

surface, thus the time of relaxation reflects the spatial scale of the pore space. At higher magnetic fields ($>10\text{MHz}$) T_2 relaxation is also affected by magnetic susceptibility broadening where fluid molecules diffuse through the internal magnetic field gradients (produced by magnetic susceptibility difference between the solid and fluid). These relaxation times can be shortened by paramagnetic contrast agents, thus enabling time-lapse imaging of fluid-related transport processes within porous media. An alternative MRI image for fluid transport analysis is a Pulse Field Gradient (PFG); this uses a pair of magnetic field gradients pulses to encode for molecular displacements, enabling the measurement of diffusion, dispersion and velocity imaging. For a more detailed explanation of the physics of these types of images and general theory of NMR, the reader is referred to Levitt (2002) or Callaghan (1993).

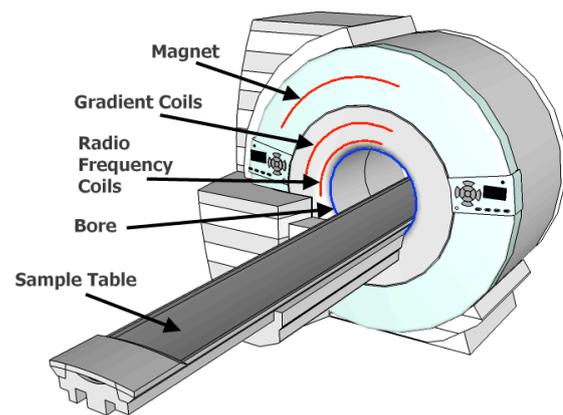


Figure 1: Schematic diagram of an MRI machine illustrating the concentric arrangement of coils (360°) and magnet.

Image Acquisition

The three gradient coils permit data acquisition in any orientation as 1D profiles, 2D slices or 3D volumes. The raw MRI dataset is a complex Cartesian grid with units of reciprocal space, which is termed k-space. For sediment-pore-fluid related research it is a volume which is of interest, hence the 3D k-space is inverse 3D Fourier transformed and the magnitude taken so as to produce a 3D image (MRI) which is spatially recognisable on an x, y, z co-ordinate grid of voxels (i.e. 3D pixels). Whilst areas of the image where nuclei are mobile (e.g. fluids) return a signal

and are observed as bright regions on a grey-scale spectrum, regions of solid fail to return a signal and appear black. Figure 2 illustrates this process, culminating in 3D data of the internal structure of the sample volume which can be quantitatively analysed using standard image processing software packages.

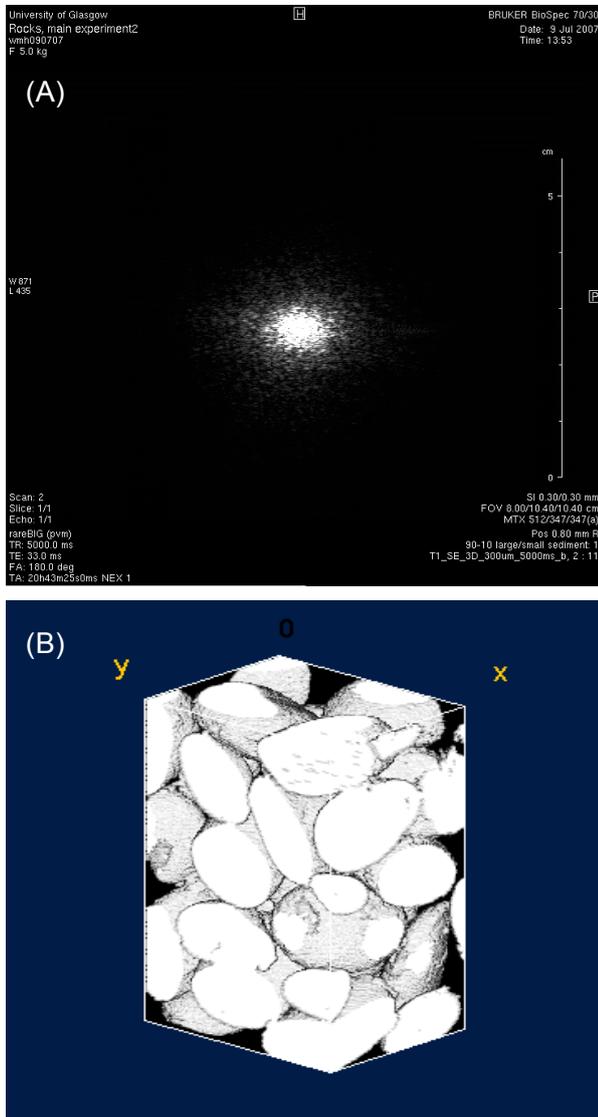


Figure 2: Example of image reconstruction, including (A) k -space signal, where the white signal indicates the presence of ^1H nuclei; (B) Fourier-transformed signal into spatial 3D volume of sediment immersed in water, as generated using ImageJ software.

Examples of use

To date, MRI has been used for a number of sedimentological analyses in a wide range of disciplines (see review papers of e.g. Mantle & Sederman 2003; Werth et al. 2010). Despite this, it is still considered an

'embryonic' technique for geomorphological investigation and three relevant exemplar topics are briefly explored below.

Porous media structure

Grain packing arrangements and pore size distributions are well studied using dynamic MRI (e.g. Baldwin et al. 1996; Baumann et al. 2000; Sederman & Gladden 2001; Sederman et al. 2004), including recent examples specific to geomorphology (e.g. Bortolotti et al. 2006; Kleinhans et al. 2008; Haynes et al. 2009; Haynes et al. 2012). River bed structure analysis is one such research arena where high strength MRI (3T–7T) has been used to yield 3D volumetric images (resolution 300–500 μm) of water-worked sediment patches or artificially-generated packed columns comprising sediments of 0.5–22.5mm diameter (Kleinhans et al. 2008; Haynes et al. 2009; Haynes et al. 2012). Image thresholding procedures, based on the signal intensity of each voxel, were applied in order to separate the grey-scale image into local regions of solid and fluid. Subsequent analysis included: (i) accurate measurement of grain axial dimensions, made possible if isotropic voxels are acquired such that the data set can subsequently be re-sliced in any orientation; (ii) porosity and void ratio measurements, taken as bulk averages of each horizontal slice of the volume space; (iii) description of fine sediment infiltration spatial patterns of sealing and siltation processes (Figures 3a and 3b); and (iv) porosity-based descriptors appropriate to resolving the surface-subsurface transition of river beds. These papers indicate that accuracy in measurements is dependent on the size of particles relative to that of the image resolution; typically <1% error can be achieved. Such outputs clearly highlight the particular benefit of visualising the sub-surface structure and illustrate the potential of MRI in fluvial sediment research, ranging from active layer processes and armour layer development to changes to hyporheic exchange processes.

Fluid processes

Single and multi-phase flows have been analysed over a range of scales using MRI, including research into rock fractures, sediments and simplified bead packs (e.g. Baumann et al. 2000; Sederman & Gladden

2001; Hermann et al. 2002; Mantle & Sederman 2003). Aquifer-related research undertaken by Li et al. (2009) mapped water flow velocities in cm-diameter cores of sand packs (0.2-0.8mm grain diameter) and natural carbonate limestone cores extracted from the field. Data clearly demonstrated preferential flow routes local to high permeability channels in the samples, with calibrated velocity data indicating 0.9mm/s in sand and 0.2mm/s in rock fractures. Similar research has been undertaken in coarser porous media (particle diameter ~5mm) where direct imaging of flow velocities up to 150mm/s have been measured in artificial bead pack arrangements (e.g. Sankey et al. 2009; Sains et al. 2005) placed within bespoke MR-compatible flow columns (details can be found in the respective papers). Here, 2D and 3D visualisation images of flow fields are presented over a range of scales from full samples (cm) to individual conduits (μm), clearly demonstrating complex flow structures such as high speed channels, stagnant zones, vortices at conduit confluences and back-flow. Such examples of MR sequences are continually being advanced to reduce image acquisition times towards real-time acquisition of 3D Cartesian component velocity (e.g. Bock et al. 1995; Sederman et al. 2004; Li et al. 2009) and shear stress images (e.g. Sederman & Gladden 2001; Swider et al. 2007). Yet, data on the porous structure can also be used to *indirectly* simulate the internal permeability (Figure 4) or flow field (e.g. Mantle et al. 2001) using numerical models superimposed onto MRI datasets.

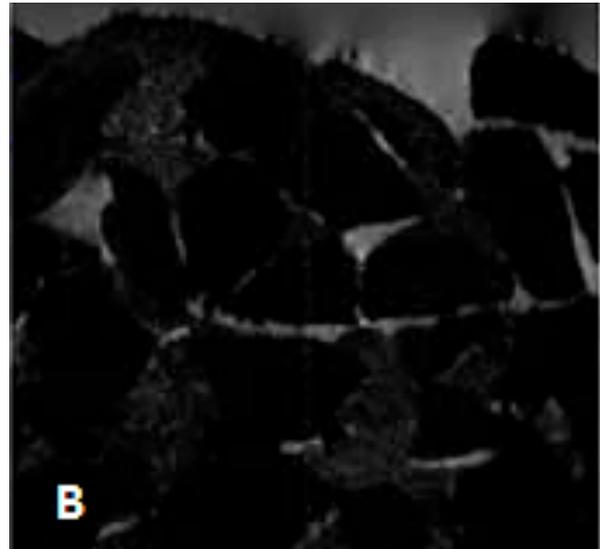


Figure 3a: Grey-scale (un-thresholded) MRI data slices (2D) of fine infiltration into a gravel framework following water-working. Image (A) shows sealing, where coarse sand deposits confined to the upper layers of the bed. Image (B) shows siltation by fine sand throughout the sample depth with isolated and interconnected voids present on the underside of gravel particles. Image adapted from Haynes et al. (2009).

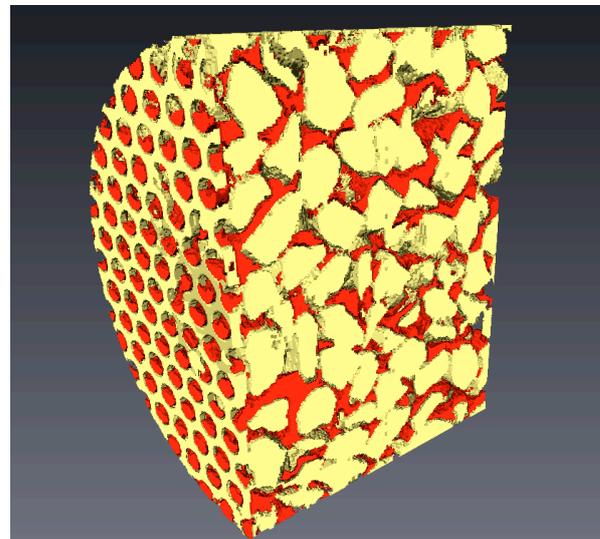


Figure 3b: Post-processed MRI data (3D volume) indicating kaolin deposition (red) within a 10mm gravel framework (yellow). Remaining pore spaces are shaded grey. This follows 10 days of clogging at 50ml/min and average sedimentation rates of 0.54g/hour. Pipe diameter was ~100mm with flow entering via a perforated plate (left of image). Images generated using Avizo software, courtesy of J. Minto, University of Glasgow.

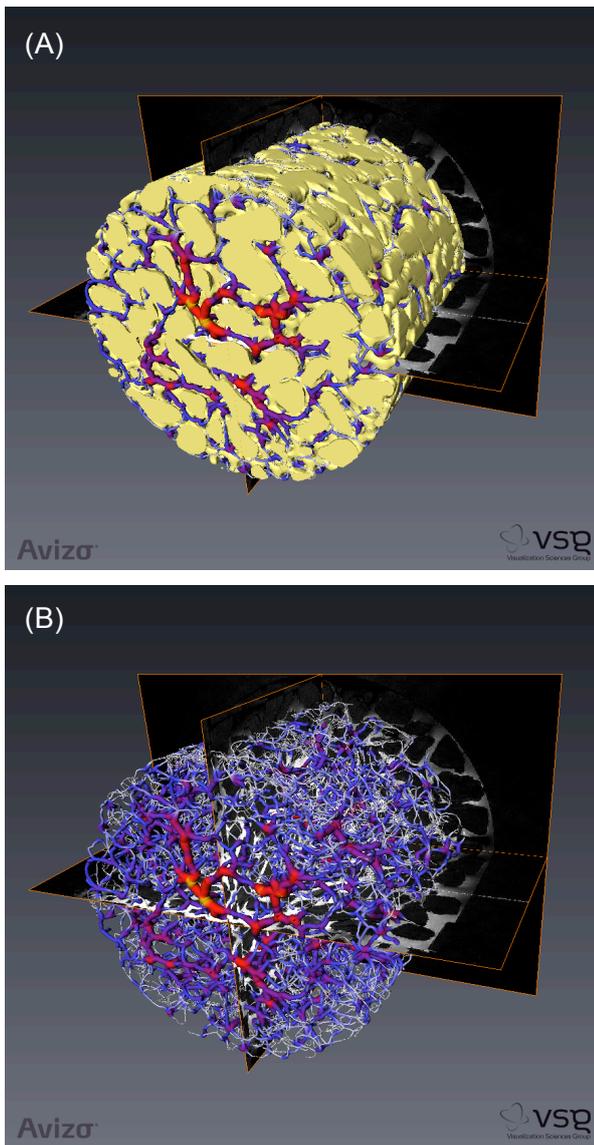


Figure 4: Permeability within a porous gravel media (field of view 86mm core diameter x 62mm length). Image (A) shows the 3D reconstruction of the MRI data with simulated permeability superimposed. Regions of sediment are shown in yellow. Image (B) shows permeability; blue represents low permeability and red indicates high permeability. Images generated using Avizo software, courtesy of J. Minto, University of Glasgow.

Contaminant & nutrient tracing

Due to its importance in hyporheic habitat, aquifer processes and environmental remediation, it is important to highlight that MRI studies have demonstrated the feasibility of using T_1 and T_2 relaxation times to analyse the mobilisation, transport and adsorption of paramagnetic ions (e.g. Gd^{3+} , Fe^{3+} , Cr^{3+} , Cu^{2+}) within the porewaters of saturated sediment matrices. Here, MRI is possible at

micromolar concentrations of heavy metal solution (e.g. Nestle et al. 2003a and b; Ramanan et al. 2012). The most recent example of this is analysis of the transport behaviour of different iron-oxide based nanoparticles (NP) within saturated heterogeneous gravels (3.5mm grain size). Using 7T field strength, Ramanan et al. (2012) used T_2 -weighted images to track the local concentration of NP and its transport through a closed-conduit flow column with a time-lapse interval of 5 minutes between captured images (Figure 5). This provided quantitative spatio-temporal data of NP transport inhibition caused adsorption of NPs onto gravel surfaces of opposite charge; this data was also employed in estimating and validating the coefficients of dispersion and retardation within a numerical convection-dispersion model. This indicates excellent potential for using MRI for wider contaminant tracing, leaching analysis, doping for sub-surface flow process studies, and flow model parameterization and validation.

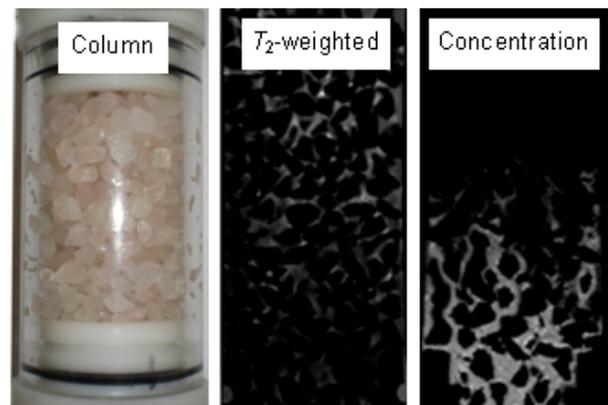


Figure 5: Packed column (100mm high x 45mm diameter) of 3.5mm diameter rose quartz grains, through which a solution of nanoparticles is flowed. A time series of T_2 -weighted images (taken at 5 minute intervals) was acquired and converted to grey-scale concentration maps (after Ramanan et al. 2012).

Wider potential for geomorphology

It is recognised that the examples included above are far from an exhaustive list of possible MRI applications relevant to geomorphology. Other measurements of sediment structure of either a surface or, of particular merit, the sub-surface volume include: pivoting angle; grain orientation; pore throat radii; and, strata depths (e.g. Haynes

et al. 2012). Similarly, time-resolved imaging shows potential for analysing flow related processes using bespoke image sequencing or paramagnetic contrast agents, *inter alia*: surface-subsurface flow interactions; biological colonisation and growth in pores; monitoring faunal movements in the benthic / hyporheic zone. Such fluid-based research is dependent on the development on appropriate MR-compatible flow chambers (e.g. Ramanan et al., 2012) constructed out of non-magnetic materials. Previous studies have employed closed conduits of vertical (e.g. Figure 5) or horizontal orientation fully-packed with material and subjected to pressurized flow. However, collaborative research funded by the Carnegie Trust for the Universities of Scotland is underway (in 2013) to investigate the viability of constructing bespoke MR-compatible open-channel flumes for flow-sediment research (Figure 6). The intention is to provide in-situ, dynamic and fully-coupled 3D flow-sediment data for the detailed description of bed evolution (surface and sub-surface) under a range of forcings.

Strengths and Weaknesses

Strengths of MRI

The examples above indicate that MRI is of particular benefit to geomorphologists where small-scale samples (such as representative cores of centimetre scale) require analysis for: (i) 3D sediment arrangement; (ii) sub-surface sediment structure/stratigraphy; (iii) or sub-surface flow. At present, MRI is complementary with core-based sampling approaches to field investigations in geomorphology; the structural integrity of the core is maintained during grain-scale analysis via MRI, thus minimising uncertainty of sample structure and permitting multiple or repeat tests (e.g. stratigraphic followed by permeability analysis) post-imaging.

Some preliminary research has been undertaken specifically comparing MRI to other commonly employed techniques for sediment investigation. Haynes et al. (2012) contrasts MRI to high-resolution laser displacement scanning over small-scale patches (cm-scale); this specifically focuses on the benefit of MRI data of sub-surface bed structure for porosity analysis and accurate modelling of surface topography.

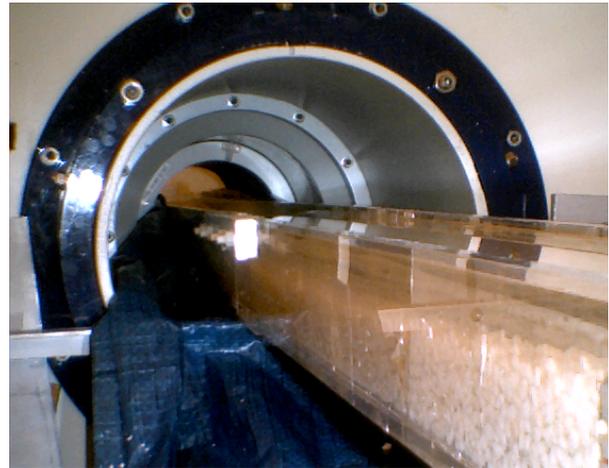


Figure 6: Trials of a bespoke MR-compatible open-channel flume (100mm rectangular cross-section) within the bore of the 7T MRI system at Glasgow's Experimental MRI Centre. All flume components and fixings within the MRI room are Perspex, made water-tight via rubber seals. The flow-recirculation pump is housed in a separate ante-room (due to metal components) with plastic pipes and connectors running through the wall between the pump and MRI rooms. The flume set-up shown includes a 60mm deep bed of 4mm dolomite sediments and a recirculating flow of water. At the time of press, this research is ongoing as a collaborative project between the Universities of Heriot-Watt, Glasgow, Dundee, Aberdeen & Strathclyde and funded by The Carnegie Trust for The Universities of Scotland (2013).

Similarly, it is prudent to compare MRI with 3D X-ray tomography (see Section 1.5.4 of Geomorphological Techniques) previously applied to porous media research. Here, the use of synchrotrons provides the highest energy X-rays capable of tuning photon energy to specific material types in the sample; hence Synchrotron X-ray Microtomography (SMT) is most comparable to MRI in terms of imaging capability, flexibility and resolution (μm -scale). Detailed information, including the science underpinning the SMT technique and its advantages, can be found in Werth et al. (2010); this indicates that that the relative advantages of MRI include: (i) flexibility in 1D, 2D or 3D acquisition and the scale/resolution of data; (ii) faster imaging sequences appropriate to fluid flow data capture at frequencies up to millisecond temporal resolution; (iii) greater capability to resolve

different chemical species; and, (iv) improved image contrast.

Difficulties with MRI

When MRI is considered, the following user-based set-up decisions are important. Implicitly, these can pose limitations to MRI use; this section highlights to what extent these can be minimised and, in doing so, what compromises will be made to other set-up parameters and variables:

Magnet strength and bore diameter affects the image resolution. The stronger the static magnetic field, the larger the NMR signal and the greater the potential for finer imaging resolution; however, the cost of superconducting magnets increases steeply with both field strength and bore diameter. The restriction of bore diameter may compromise sample size in terms of Representative Elementary Volume (REV). For example, whilst a 1.5T MRI machine offers bore diameter ~600mm it offers only mm-scale image resolution; conversely, a 7T MRI machine has a bore size only ~150mm (Figure 6) but image resolution of ~100µm. MRI is, therefore, compatible with field-based surface rock/sediment coring techniques which extract cm-scale samples (e.g. drilling, vacuum, freezing). Yet, future research is needed to develop methods appropriate to cutting larger samples into smaller sub-samples, compatible with the narrower MR bores of the higher strength magnets offering improved image resolution. These methods must account for possible disturbance at the cutting face in order to provide larger REV's and the potential for sub-sample images to be 'stitched' back together during post-processing. Similarly, recent technological advances towards human MRI "open bore" magnets may remove the bore-size constraint entirely, proving highly beneficial to future sedimentary research where larger samples require imaging. That said, it should be highlighted that this system is not yet designed to be portable nor intended to operate over very large spatial areas (such as that offered by surface-based Terrestrial Laser Scanning).

Image acquisition time is a function of the sample volume, required image resolution and relaxation time of the magnetic nuclei. For example, Haynes et al. (2009) state that

doubling the spatial resolution of the image matrix increases the number of voxels (3D pixels) by 2^3 and increases the scan time by a factor of 4. Thus, to achieve the same signal-to-noise ratio would require scan durations 32 times that of the original. High resolution scanning can prove expensive in terms of facility hire time, with even small cores (Ø100mm x 100mm) taking nearly a day to scan in their entirety at ~100µm resolution. In addition, the sediment framework needs to be immobile during this period, which precludes dynamic trials such as pore-clogging during the scan. Kleinhans et al. (2008) therefore illustrated how the signal intensity within each voxel may be used to increase the resolution of analysis, without increasing image acquisition times; detailed analysis of uncertainty is provided in their paper

Sediment geochemistry will dictate the magnetic susceptibility of the sample. Where materials contain a range of transition metal species, particularly iron, the relaxation times of the saturating fluid can be significantly affected by surface relaxation. In addition, at higher magnetic fields T_2 relaxation becomes dominated by magnetic susceptibility broadening, which can increase with the square of the magnetic field; this becomes more prominent the smaller the pore size and can make MRI extremely difficult (e.g. Haynes et al., 2009). As relaxation times depend greatly on sediment geochemistry in relation to field strength and pore size distribution, the reader is referred to Kleinberg et al. (1993, 1994) and Packer (1996) for detailed insight. Such relaxation effects are the reason 'clean' sediments (e.g. dolomite, limestone, sandstone and quartz) have been used in recent high resolution scanning (µm-scale); yet, it is important to note that mm-scale MRI is viable for 'dirty' sediments as demonstrated widely by the oil industry during the last two decades (Packer, 1996).

Access to facilities specific to non-medical MRI research is a notable challenge to the widespread use of this technique in geomorphology. Facility access is largely constrained to academic engineering and science schools, yet collaborative ventures here are broadly welcomed. This is important, as the complexity of applying this technique to non-medical research requires expert

knowledge. Typical costs are therefore in the region of £250-500 per scan.

Challenges of application to field samples

The examples and discussion provided above clearly demonstrate that MRI is an emerging technique of merit to geomorphology-related sedimentary science, in particular for grain-scale analysis of sub-surface structure and dynamic imaging of related flow-transport processes.

At present, MRI is restricted to a laboratory-based technique which is dependent upon, and appropriate to, small-scale core-based samples being extracted from the field and transported to the MRI. Whilst magnetically 'clean' materials may be scanned using any magnet strength, the use of low strength magnets (e.g. 1.5T or 3T) permits 3D imaging of natural sediment even where magnetic properties are present. The sensitivity of image quality and image resolution to the material type should be implicit to the choice of MR facility and associated set-up. A key challenge of the future is, therefore, to explore development of a toolbox of techniques specific to field sample analysis via MRI, including: methods of sample extraction/preservation as commensurate with solid-phase transport and fluid-state imaging; Representative Elementary Volumes in relation to the size of the bore size of the MRI facility; optimisation of magnet strength to both image resolution and sample geochemistry etc.

Whilst portable Nuclear MR (NMR) devices are mentioned in the literature (e.g. Blümich et al. 1995; Blümich 2007; Stork & Nestle 2007), modifications for MRI do not yet have evidence of field-based deployment for geomorphological research. Similarly, new open-bore MRI systems offering the potential to analyse larger samples (metre scale) have not been trialled on sediment/rock samples. Whilst combined portable and open-bore capability would be both idealistic and futuristic for field-based geomorphological purposes, this appears unlikely for a technology primarily motivated by medical application and funding. Thus, as the greatest benefits to field-based investigations are offered by portable systems, focus should be placed on how best develop future portable MRI systems in a manner appropriate to our

end-user requirements. This requires geomorphologists to both, ascertain the benefits and limits of existing MRI for sediment analysis and, engage in new collaborations with the MR scientists and manufacturers to refine imaging sequences and magnet designs. As MRI use for geomorphological application is at such an early stage, the opportunities for new scientific insight and methodological development provide an exciting research arena of the future.

Conclusions

MRI offers a flexible non-intrusive technique with which to visualize and quantitatively analyse 3D internal structure and processes within an opaque porous media, including the sub-surface. Whilst this is recognised as an 'emerging' methodology for geomorphological research, the handful of existing studies clearly indicate the significant benefits of MRI for high resolution spatio-temporal measurement of sediment packing arrangement, porosity, permeability, fluid flow and contaminant tracing. As this approach appears well suited to a wide range of porous media environments, there exists significant scope to develop MRI sequences to unlock the 'black-box processes' of many geomorphological research disciplines.

Acknowledgements

This research was funded by a research grant awarded from The Carnegie Trust for The Universities of Scotland into the use of MRI for analysing flow-sediment-biology processes in river systems. Data collection for the images used was provided by James Minto & Dr. Elisa Vignaga during PhD research at The University of Glasgow (2006 and 2012) as undertaken with supervision by the paper authors; here, co-supervision by Dr. Vernon Phoenix proved invaluable. The authors also sincerely thank the technical staff of GEMRIC, the School of Engineering at the University of Glasgow, and the School of the Built Environment at Heriot-Watt University for their assistance in this research.

References

- Baumann T, Petsch R, Niessner R. 2000. Direct 3-D measurement of the flow velocity in porous media using magnetic resonance tomography. *Environmental Science & Technology* **34** : 4242-4248
- Baldwin CA, Sederman AJ, Mantle MD, Alexander P, Gladden LF. 1996. Determination and characterization of the structure of a pore space from 3D volume images. *Journal of Colloidal and Interface Science* **181** : 79-92.
- Blümich B, Blümli P, Eidmann G, Guthausen A, Haken R, Schmitz U, Saito K, Zimmer G. 1998. The NMR MOUSE: Construction, Excitation, and Applications. *Magnetic Resonance Imaging* **16** : 479-84
- Blümich B. 2007. Mobile NMR: concepts and applications. <http://www.nmr-mouse.de/index.php>
- Bock M, Schad LR, Müller E, Lorenz WJ. 1995. Pulsewave velocity measurement using a new real-time method. *Magnetic Resonance Imaging* **13** : 21-29
- Bortolotti V, Gombia M, Cernich F, Michelozzi E, Fantazzini P. 2006. A study to apply nuclear magnetic resonance porosity measurements to seabed sediments. *Marine Geology* **230** : 21-27
- Callaghan PT. 1993. Principles of Nuclear Magnetic Resonance Microscopy. Oxford University Press : Oxford
- Fantazzini P, Bortolotti V, Kärger J, Galvosas P. (Eds.) 2011. Proceedings of the Tenth International Bologna Conference on Magnetic Resonance in Porous Media, 12 – 17 September 2010, Leipzig Germany, *Conference Proceedings Series American Institute of Physics* **1330**
- Haynes H, Vignaga E, Holmes WM. 2009. Using magnetic resonance imaging for experimental analysis of fine-sediment infiltration into gravel beds. *Sedimentology* **56** : 1961-1975
- Haynes H., Ockelford A-M, Vignaga E, Holmes WM. 2012. A new approach to define surface/sub-surface transition in gravel beds. *Acta Geophysica* **60** : 1589-1606
- Hermann K-H, Pohlmeier A, Gembris D, Vereecken H. 2002. Three-dimensional imaging of pore water diffusion and motion in porous media by nuclear magnetic resonance imaging. *Journal of Hydrology* **267** : 244-257
- Huerlimann M, Song Y-Q, Fantazzini P, Bortolotti V. (Eds.) 2008. Proceedings of the Ninth International Bologna Conference on Magnetic Resonance in Porous Media, 13-17 July 2008, Cambridge, USA, *Conference Proceedings Series American Institute of Physics* **1081**
- Kawaguchi T. 2010. MRI measurement of granular flows and fluid-particle flows. *Advanced Powder Technology* **21**: 235-241
- Kleinberg RL, Farooqui SA, Horsfield MA. 1993. T(1)/T(2) ratio and frequency-dependence of NMR relaxation in porous sedimentary rocks. *Journal of Colloid and Interface Science* **158** : 195-198
- Kleinberg RL, Kenyon WE, Mitra PP. 1994. Mechanism of NMR relaxation of fluids in rock. *Journal of Magnetic Resonance* **108** : 206-214
- Kleinhans MG, Jeukens CRLPN, Bakker CJG, Frings R. 2008. Magnetic resonance imaging of coarse sediment. *Sedimentary Geology* **208** : 69-78
- Levitt MH. 2002. Spin Dynamics. John Wiley & Sons : England
- Li L, Chen Q, Marble AE, Romero-Zerón L, Newling B. 2009. Flow imaging of fluids in porous media by magnetization prepared centric-scan SPRITE. *Journal of Magnetic Resonance* **197** : 1-8
- Mantle MD, Sederman, AJ. 2003. Dynamic MRI in chemical process and reaction engineering. *Progress in Nuclear Magnetic Resonance Spectroscopy* **43** : 3-60
- Mantle MD, Sederman AJ, Gladden LF. 2001. Single- and two-phase flow in fixed-bed reactors: MRI flow visualisation and lattice-Boltzmann simulations. *Chemical Engineering Science* **56** : 523-529.
- Nestle N, Baumann T, Wunderlich A, Niessner R. 2003a. MRI observation of heavy metal transport in aquifer matrices down to sub-mg quantities. *Magnetic Resonance Imaging* **21** : 345-349
- Nestle N, Wunderlich A, Niessner R, Baumann T. 2003b. Spatial and temporal observations of adsorption and remobilization of heavy metal ions in a sandy aquifer matrix using magnetic resonance imaging.

Environmental Science & Technology **37** : 3972-3977

Ohkubo T, Yamaguchi M. 2007. Time-dependent diffusion of water in clay gel-saturated porous media. *Magnetic Resonance Imaging* **25** : 577-8

Packer KJ. 1996. Oil Reservoir Rocks Examined by MRI. In: Encyclopedia of Magnetic Resonance. John Wiley & Sons. London : 3365-3376

Ramanan B, Holmes WM, Sloan WT, Phoenix VR. 2012. Investigation of nanoparticle transport inside coarse-grained geological media using magnetic resonance imaging. *Environmental Science Technology* **46** : 360-366.

Sains MC, El-Bachir SM, Sederman AJ, Gladden L. (2005) Rapid imaging of fluid flow patterns in a narrow packed bed using MRI. *Magnetic Resonance Imaging* **23** : 391-393

Sankey MH, Holland DJ, Sederman AJ, Gladden LF. 2009. Magnetic resonance velocity imaging of liquid and gas two-phase flow in packed beds. *Journal of Magnetic Resonance* **196** : 142-148

Sederman AJ, Gladden LF. 2001. Magnetic resonance visualisation of single- and two-phase flow in porous media. *Magnetic Resonance Imaging* **19** : 339-343

Sederman AJ, Mantle MD, Buckley C, Gladden LF. 2004. MRI technique for velocity measurement of velocity vectors, acceleration and autocorrelation functions in turbulent flow. *Journal of Magnetic Resonance* **166** : 182-189

Stork H, Nestle N. 2007. A low-cost, large volume, easy-access design for a portable NMR magnet. *Magnetic Resonance Imaging* **25** : 544-591

Swider P, Conroy M, Pedrono A, Ambard D, Mantell S, Søballe, K, Bechtold JE. 2007. Use of high-resolution MRI for investigation of fluid flow and global permeability in a material with interconnected porosity. *Journal of Biomechanics* **40** : 2112-2118

Tyrolougou P, William A, Dudeney L, Gattoni CA. 2005. Evolution of porosity in geotechnical composites. *Magnetic Resonance Imaging* **23** : 765-768

Vogt, C, Galvosas P, Klitzsch N, Stallmach F. 2002. Self-diffusion studies of pore fluids in

unconsolidated sediments by PFG NMR. *Journal of Applied Geophysics* **50** : 455-467

Werth CJ, Zhang, C, Brusseau ML, Ostrom, M, Baumann T. 2010. A review of non-invasive imaging methods and applications in contaminant hydrogeology research. *Journal of Contaminant Hydrology* **113** : 1-24

1.5.5. Ground Penetrating Radar

Martin Robinson¹, Charlie Bristow², Jennifer McKinley¹ and Alastair Ruffell¹

¹School of Geography, Archaeology and Palaeoecology, Queen's University Belfast

²Department of Earth and Planetary Sciences, Birkbeck University of London

(mrobinson34@qub.ac.uk)



ABSTRACT: Ground-penetrating radar (GPR) is an effective tool to visualise the structure of the shallow subsurface. The purpose of this article is to offer guidelines to non-specialist GPR users on the collection, processing and interpretation of GPR data in a range of environments. The discussion on survey design focuses on single fold, fixed-offset reflection profiling, the most common mode of GPR data collection, however the design factors can be applied to other survey types. Information on the visualisation of processed data, as well as the advantages and disadvantages of GPR, is provided. Possible applications of GPR in geomorphological research are presented, along with a case study outlining how GPR can be used to measure peat thickness.

KEYWORDS: Ground penetrating radar, survey design, processing, interpretation, applications

Introduction

Ground-penetrating radar (GPR) has become a popular tool in sedimentological studies, with Table 1 illustrating recent research that has utilised GPR in the analysis of environmental processes (Jol and Bristow, 2003). It is a non-invasive geophysical technique designed primarily for subsurface investigation (Neal, 2004; Comas *et al.*, 2004). A GPR system detects changes in the electrical properties of the shallow subsurface using discrete pulses of high frequency electromagnetic (EM) energy, usually in the 10-1000 MHz range (Neal, 2004). The technique has been successfully applied in a wide range of environmental studies however an understanding of the capabilities and limitations of GPR is vital when considering using the technique, with the quality of GPR results often being dependent on the surveyed environment (Daniels, 2004). The purpose of this article is to offer guidelines on the collection, processing and interpretation of GPR data.

This paper presents an overview of good practice material for non-specialist users working in the field of environmental

research. An operator of GPR equipment must have an understanding of the fundamental principles underlying the technique (Daniels, 2004). A training course or field survey with an experienced operator is highly recommended.

Principles

A GPR system transmits short pulses of high frequency EM energy (10-1000 MHz) from an antenna into the subsurface (Jol and Smith, 1991; Holden *et al.*, 2002). As an EM wave disseminates downwards, its velocity is altered due to encounters with materials of differing electrical properties (Neal, 2004). Abrupt changes in the dielectric constant results in a portion of the energy being reflected, with the receiving antenna of the GPR system detecting the reflected EM energy (Figure 1a). Scattering of the radar waves occurs when the radar signal travels through overburden. It is worth noting that most GPR antennas are not focussed and transmit energy into the air as well as the ground. As a result it is possible to get reflections from objects above ground such as walls, cars, fences or overhead cables. Some antennas are shielded to reduce

external noise and prevent the signals being transmitted through the air but this adds to the weight and bulk of the antennas which can be awkward in the field.

The time between transmission and reception, referred to as the two-way travel-time (TWT) and commonly measured in nanoseconds, is a function of reflector depth and the EM velocity of propagation (Neal, 2004; Jol and Smith, 1991). GPR provides a continuous profile of the subsurface, displaying horizontal survey distance against vertical TWT. Vertical TWT is converted to depth with knowledge of the propagation velocity, expressed as (Equation 1):

$$d = v \times t / 2 \quad (\text{Equation 1})$$

In which d is depth, v is velocity, and t is TWT. The propagation of EM energy through media is controlled by several material properties. Dielectric constant (dielectric permittivity), a property which is strongly dependent on the water content of a material, is the primary factor controlling the velocity of an EM wave. Reflections therefore can typically be related to interfaces where there is a considerable change in water content (Comas *et al.*, 2005). Electrical conductivity is a measure of charge transport, through a

medium, on application of an electric field (Powers, 1997). The most important electrical conduction losses, in relation to GPR performance, occur due to ionic charge transport in water and electrochemical processes associated with cation exchange. Olhoeft (1998) describes the importance of clay mineral cation exchange in studies of soil and sediment. The equation for the velocity of propagation is expressed as (Equation 2):

$$v = c / \sqrt{\epsilon_r} \quad (\text{Equation 2})$$

In which v is the velocity, c is the speed of light (300mm/ns) and ϵ_r is the relative dielectric constant. Table 2 provides typical dielectric constant and electrical conductivity values for common materials encountered using GPR.

Suitability of GPR

It is essential from a scientific view to clearly establish what data are required to test a particular hypothesis. This will influence the size of the survey, the depth of investigation and the resolution needed (Jol and Bristow, 2003). Before planning a survey it is important to determine if GPR will be effective.

Table 1. Recent environmental studies that use GPR

Depositional Setting	Recent Papers
Glacial	Appleby <i>et al.</i> (2010); Benediktsson <i>et al.</i> (2009); Degenhardt (2009); Gibbard <i>et al.</i> (2012); Hart <i>et al.</i> (2011); Irvine-Fynn <i>et al.</i> (2006); Kim <i>et al.</i> (2010); Langston <i>et al.</i> (2011); Leopold <i>et al.</i> (2011); Monnier <i>et al.</i> (2011); Murray and Booth (2010)
Fluvial	Ashworth <i>et al.</i> (2011); Johnson and Carpenter (In Press); Kostic and Aigner (2007); Lunt and Bridge (2004); Nobes <i>et al.</i> (2001); Rice <i>et al.</i> (2009); Słowik (2011)
Delta	Barnhardt and Sherrod (2006); Gibbard <i>et al.</i> (2012); Gutsell <i>et al.</i> (2004)
Coastal	Bennett <i>et al.</i> (2009); Bristow and Pucillo (2006); Buynevich <i>et al.</i> (2010); Clemmensen <i>et al.</i> (2012); Nielsen and Clemmensen (2009); Olson <i>et al.</i> (In Press); Pascucci <i>et al.</i> (2009); Tamura <i>et al.</i> (2011)
Aeolian	Bristow <i>et al.</i> (2007, 2010); Buynevich <i>et al.</i> (2010); Clemmensen <i>et al.</i> (2012); Tamura <i>et al.</i> (2011); Vriend <i>et al.</i> (2012)
Peatland	Comas <i>et al.</i> (2005, 2011); De Oliveira <i>et al.</i> (2012); Kettridge <i>et al.</i> (2008); Lowry <i>et al.</i> (2009); Plado <i>et al.</i> (2011); Rosa <i>et al.</i> (2009)
Faults	Bhosle <i>et al.</i> (2007); Christie <i>et al.</i> (2009); Malik <i>et al.</i> (2010)

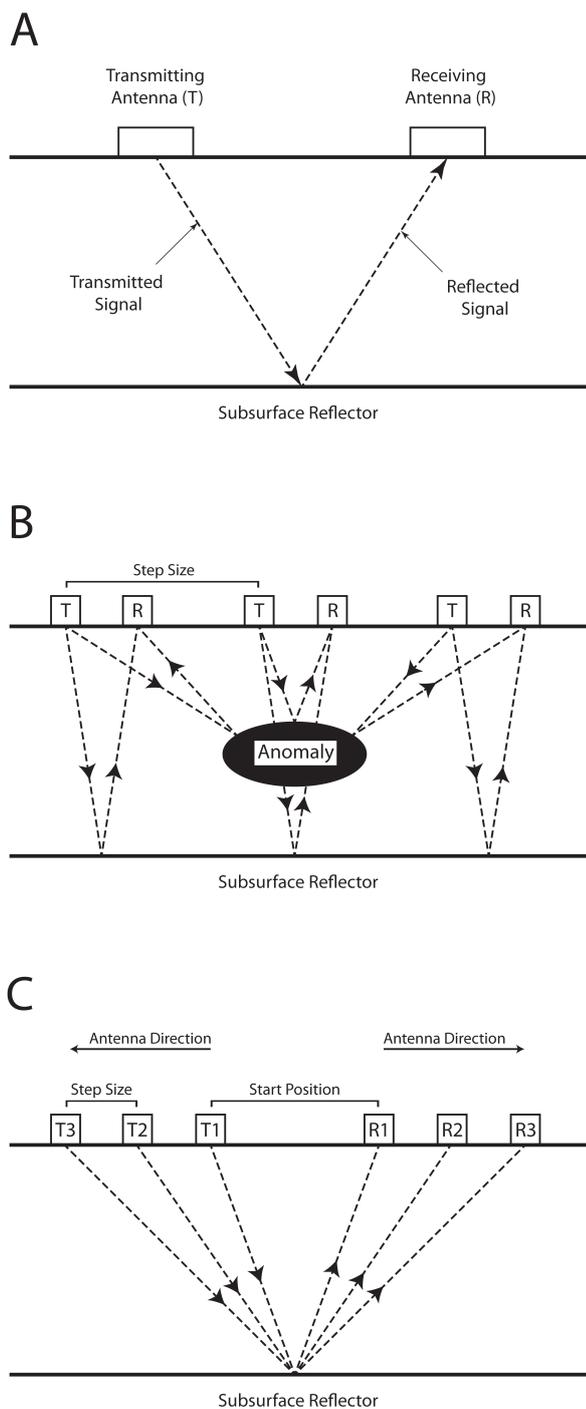


Figure 1. (a) The propagation of an EM wave in subsurface material (adapted from Neal, 2004); (b) Fixed offset profiling (adapted from Annan, 2001); (c) Common mid-point (CMP) sounding (adapted from Annan, 2001).

GPR can quickly be deemed an unsuitable technique if the target is beyond the depth range of the GPR system. For GPR to work effectively, the target must exhibit electric properties (dielectric constant and electrical conductivity) which contrast with the host subsurface. The strength of an EM reflection is proportional to the magnitude of this contrast, with the amount of energy reflected

being given by the reflection coefficient (R), expressed as (Equation 3):

$$R = \frac{\sqrt{v_2} - \sqrt{v_1}}{\sqrt{v_2} + \sqrt{v_1}} \quad (\text{Equation 3})$$

In which v_1 and v_2 are the velocities for layers 1 and 2 i.e the target and the host subsurface (Neal, 2004). In all cases, the value of R will be between +1 and -1, with values further from zero representing greater differences in electrical properties. Table 2 can be used as a guide to determine the suitability of GPR for a particular study, with velocity values for typical subsurface mediums being provided. The environment under investigation needs to be examined before performing a survey. Metallic objects and major structural features within the vicinity of the survey location can seriously affect GPR results (Annan and Cosway, 1994).

Survey Design

This discussion on survey design focuses on single fold, fixed-offset reflection profiling (Figure 1b), the most common mode of GPR data collection, however the design factors can be applied to other survey types (Jol and Bristow, 2003; Annan, 2005). The step-like procedure involves moving a signal transmitter and receiver, in a fixed antenna geometry, over the surface in repetitive steps. This mode is preferred in studies that require high spatial-horizontal resolution. Continuous data collection is not recommended for detailed sedimentary investigations as the movement during collection can smear the data, creating problems when locating a specific subsurface feature. Continuous data collection is more appropriate for reconnaissance surveys (Jol and Bristow, 2003).

Several parameters need to be defined when designing a single fold, fixed-offset reflection survey.

Radar Frequency: Vertical Resolution

Antenna frequency is a significant factor in survey design, determining the resolution and depth of penetration (Baker *et al.*, 2007). Low frequency waves tend to penetrate deeper into the subsurface as they are not as easily

*Table 2. Typical dielectric constant, electrical conductivity, velocity and attenuation values of common subsurface materials (Leckebusch, 2003; rows labelled * from Annan, 2005)*

Material	Dielectric Constant	Electrical Conductivity (mSm⁻¹)	Velocity (m ns⁻¹)	Attenuation (dB m⁻¹)
Air	1	0	0.3	0
Salt water	80	3000	0.033	600
Fresh water	80	0.5	0.033	0.1
Ice*	3-4	0.01	0.16	0.01
Granite, dry	5	0.01	0.13	0.01
Limestone*	4-8	0.5-2	0.12	0.4-1
Shales*	5-15	1-100	0.09	1-100
Sand, dry	5	0.01	0.13	0.01
Sand, wet*	20-30	0.1-1.0	0.06	0.03-0.3
Clay, wet	10	500	0.095	300
Soils:				
sandy, dry	2.6	1.4	0.19	1
sandy, wet	25	69	0.06	23
clayey, dry	2.5	2.7	0.19	3
clayey, wet	19	500	0.07	200
frozen	6	0.1	0.12	0.1

attenuated as high frequency waves. The ability of a GPR system to resolve fine subsurface features however deteriorates at lower frequencies, meaning the selection of the optimal antenna frequency is a trade-off between the desired resolution and the achievable penetration depth (Harari, 1996). As a rule of thumb, it is better to trade resolution for depth, with high resolution being useless if the target cannot be detected (Annan, 2005).

An antenna does not transmit EM waves at a single frequency but across a frequency spectrum, with the higher frequencies in the range being preferentially attenuated as they propagate through the subsurface. This results in the return centre frequency, the most common frequency detected by the receiving antenna, typically being lower than the nominal centre frequency of the transmitting antenna. More realistic vertical resolution estimates are consequently obtained using the return centre frequency (Neal, 2004). The preferential attenuation of high frequency radar waves also causes a decrease in resolution down the GPR profile even when no changes in the velocity exist.

The majority of sedimentary studies use antenna with frequencies between 50 and 500MHz (Jol and Bristow, 2003). Table 3, based on the assumption that the required spatial resolution is approximately 25% of the target depth, can be used as a simple guide to determine a suitable frequency. It is important to understand that the depth of penetration is highly dependent on the clay content of the medium, with clay significantly attenuating EM waves and limiting the investigation depth (Bristow, 2013). The information presented in Table 2 therefore should not be a substitute for thorough survey planning, with the depth of penetration being subject to the soil type (Annan, 2001).

Table 3. Guideline frequency values (Annan, 2001)

Depth (m)	Centre Frequency (MHz)
0.5	1000
1.0	500
2.0	200
5.0	100
10.0	50
30.0	25
50.0	10

Stratigraphy in sedimentary environments occurs at scales ranging from 10^{-3} to 10^1 m both horizontally and vertically (Jol and Bristow, 2003). Identifying a suitable resolution is often a vital component of many subsurface environmental studies. Vertical resolution is usually considered to be approximately one quarter of the radar signal wavelength in the subsurface, with wavelength being a function of antenna frequency and the velocity of the signal (Bristow, 2009). This is shown by Equation 4.

$$\lambda = \frac{V}{f} \quad (\text{Equation 4})$$

$$\begin{aligned} \lambda &= \text{Wavelength} \\ V &= \text{Velocity} \\ f &= \text{Frequency} \end{aligned}$$

Bristow (2009) provides theoretical values for GPR resolution in sands based on Eq.4 (Table 4). For example, a 100MHz transmitter used to investigate sand saturated with fresh water (velocity of 0.06m ns^{-1}) will produce a wavelength of 0.6m and a resolution of 0.15m (one-quarter of the wavelength value). The approximate velocity values presented in Table 2 can be used with a desired resolution value to identify a suitable GPR frequency for a particular study. Various antennas frequencies can be employed to effectively image the subsurface using GPR (Jol and Bristow, 2003). The use of multiple frequencies depends entirely on the research question, with the time and cost of swapping several different antennas, as well as processing data collected at different frequencies, needing to be considered.

Time Window

The required time window (W) is estimated using the following equation (Equation 5):

$$W = 1.3 \frac{2 \times \text{Depth}}{\text{Velocity}} \quad (\text{Equation 5})$$

Table 4. Theoretical values for GPR resolution in different sands (Bristow, 2009). Note: Velocity value for dry sand differs slightly from figure provided in Table 1

Antenna Central Frequency (MHz)	Theoretical Resolution (m) for Saturated Sand (0.06 m ns^{-1})	Theoretical Resolution (m) for Damp Sand (0.1m ns^{-1})	Theoretical Resolution (m) for Dry Sand (0.15m ns^{-1})
50	0.3	0.5	0.75
100	0.15	0.25	0.375
200	0.075	0.125	0.1875
400	0.0375	0.0625	0.09375

with the maximum depth and minimum velocity likely to be observed being used. The equation accounts for uncertainty in velocity and depth variations by increasing the estimated time by 30% (Annan, 2005). Table 2 can be used as a guide if no information on the electrical properties of the survey site is available.

Temporal Sampling Interval

The time interval between points on a recorded waveform is another parameter which needs to be considered when designing a survey. Annan (2005) indicates that as part of a good survey design, the sampling rate should be approximately six times the centre frequency of the antenna being used, with the following expression being utilised to calculate suitable sampling intervals (Equation 6):

$$t = \frac{1000}{6f} \quad (\text{Equation 6})$$

In which t is the maximum sampling interval (ns) and f is the centre frequency (MHz). Table 5 indicates the calculated sampling interval for a variety of antenna frequencies. The sampling values presented in the table should only be exceeded when data volume and acquisition speed are more important than the integrity of the data (Annan, 2001). The consequence of using small sample values is that random signals (noise) are given more influence on the trace. An increase in antenna frequency requires an increase in the sample rate due to the preferential attenuation at high frequencies. Increases in sample rate should be based on the Nyquist principle - the greatest vertical resolution that can be expected is one-quarter the wavelength (Jol and Bristow, 2003).

Table 5. Suitable sampling intervals and corresponding antenna frequencies (Annan, 2001)

Antenna Centre Frequency (MHz)	Maximum Sampling Interval (ns)
10	16.7
20	8.3
50	3.3
100	1.67
200	0.83
500	0.33
1000	0.17

Each trace in a GPR profile should be vertically stacked to enhance the return signals, with the running average of a number of radar transmissions being taken (Jol and Bristow, 2003). If a GPR system was programmed to collect 512 samples for each trace, with the system also being set up to stack 16 sequential traces into one record, then at least 8192 pulses (512 pulses multiplied by 16) would need to be transmitted for every recorded reflection trace (Conyers, 2013). This stacking process minimises random signals (noise) and emphasises persistent signals (reflections). An increase in trace stacking however will increase survey time (Jol and Bristow, 2003).

Step Size: Horizontal Resolution

Step size, the distance between each data collection point made along a GPR profile, is an important part of the survey design process (Jol and Bristow, 2003). A minimum step size for each antenna frequency is usually suggested by the manufacturer (Bristow, 2009). The extent of the target however outweighs the guidelines proposed by the manufacturer, with the object of interest needing to be resolved both horizontally and vertically. These values are based on the Nyquist sampling interval, one-quarter of the wavelength in the subsurface and expressed as (Equation 7):

$$n_x = \frac{75}{f\sqrt{K}} \quad (\text{Equation 7})$$

- n_x = Sampling Interval (m)
- f = Antenna Centre Frequency (MHz)
- K = Dielectric Constant

A maximum step size of one metre, often less (0.1-0.5m) depending on antenna frequency, should be used in sedimentary studies to provide detailed horizontal resolution of structures (Jol and Bristow, 2003). A typical survey performed with 100MHz antennas should have a step size of 0.25m, however a larger step size can be used when the subsurface comprises continuous horizontal layers. GPR data will not adequately define steeply dipping reflectors if the step size is too large (Annan, 2005).

Antenna Separation

The majority of GPR systems have separate antennas for transmitting and receiving (bistatic operation), with some antennas having a fixed separation while others can be varied (Annan, 2005; Jol and Bristow, 2003). Antenna separation should be as small as possible based on the needs of the survey and the wavelength of the antennas. The depth resolution of targets decreases as the distance between antennas increases, however this effect is not significant until the antenna separation nears half of the target depth (Jol and Bristow, 2003). A safe antenna separation, if there is very little site information available, is 20% of the target depth (Annan, 2005).

Line Spacing and Orientation

Line spacing is a compromise between the amount of detail required for the survey and the extent of the survey area. To avoid spatial aliasing, where the target is not correctly resolved, the line spacing is required to be approximately one quarter of the extent of the structure being investigated (Bristow, 2009). Survey lines should be arranged, whenever possible, in a grid, with lines running parallel and perpendicular to the expected dip direction of the structure (Jol and Bristow, 2003). Bristow (2009) indicates that images of sedimentary structures are improved if data is collected in the up-dip direction. Closely spaced lines give sufficiently dense coverage to produce 3D block models of the subsurface (Jol and Bristow, 2003). Grasmueck *et al.* (2005) recommend a quarter-wavelength spatial sampling (step size) as the minimum requirement for a full resolution 3D survey.

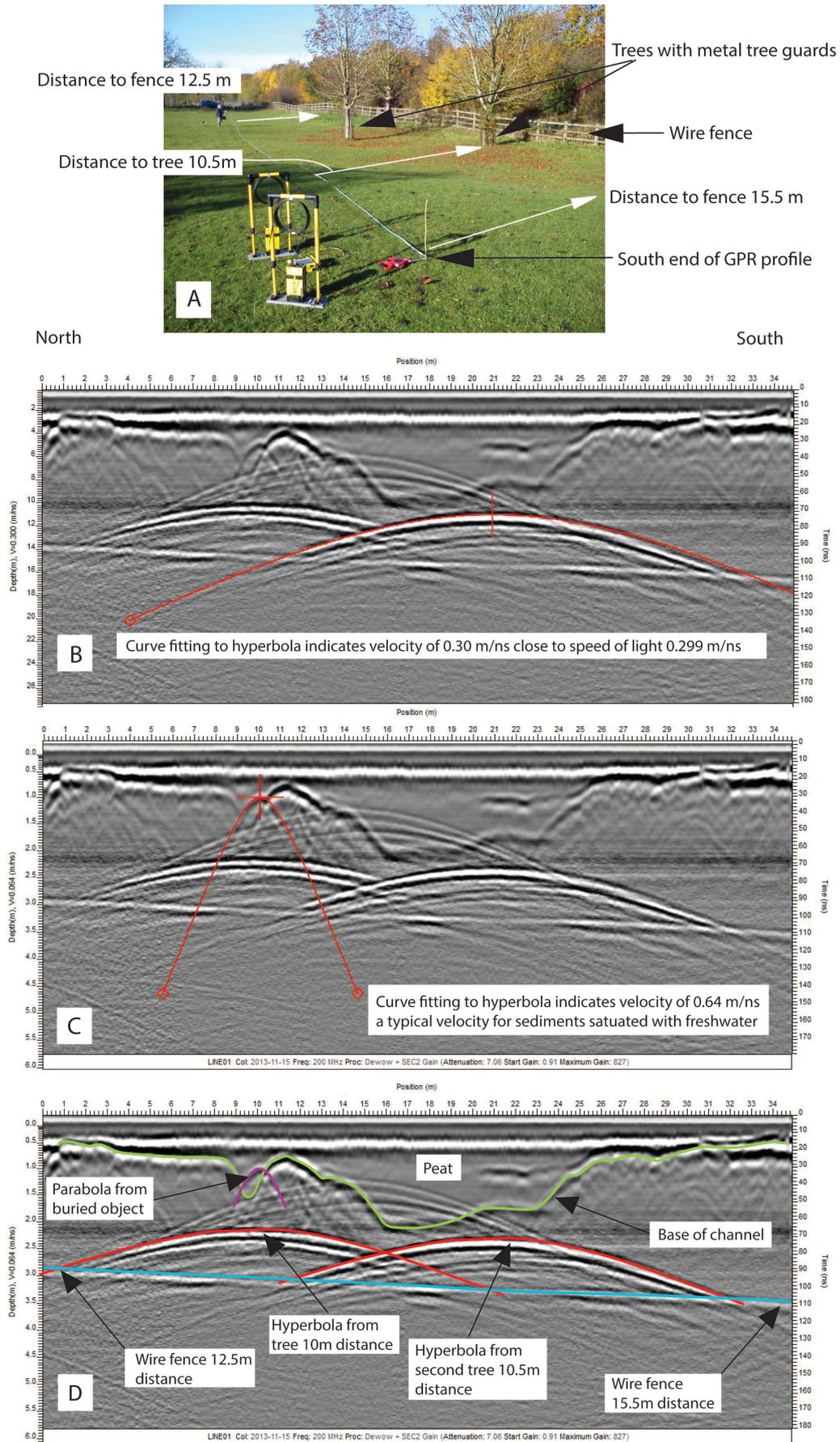


Figure 2. GPR profile using unshielded 200MHz antennas with 0.1m step size, dewow and SEC gain in a parkland setting (a) with trees and a fence, includes reflections from above (b) and beneath the ground (c) with interpretation (d) see text for details

Antenna Shielding

When a GPR antenna directs EM energy into the subsurface, a portion of the energy is lost to the air. These EM waves, as they propagate through the air, can encounter an object or planar surface with a different dielectric constant, resulting in part of the signal being reflected to the receiver (Neal, 2004). Thus, not all reflections on radar profiles necessarily originate from features within the subsurface which can be seen in (Figure 2). The GPR profile in Figure 2 includes reflections from objects above and beneath the ground. These can be readily distinguished by velocity analysis fitting hyperbolic curves to observed hyperbolic reflections. The hyperbolas from the trees with metal tree guards indicate a velocity around 0.3mns^{-1} (Figure 2b) which is consistent with the velocity of the GPR signal through the air at the speed of light (0.299mns^{-1}). The tighter hyperbola shown in Figure 2c indicates a velocity of 0.064mns^{-1} , a typical velocity for water saturated sediments. The gently inclined planar reflection picked out in blue in Figure 2d could be mistakenly interpreted as dipping bedrock strata but actually comes from a fence between 12.5 and 15.5 m from the GPR profile as shown in Figure 2a. Shielding can be used to reduce these unwanted surface reflections. However, shielded antenna never fully eliminate airwaves, therefore shielded antenna data should be treated with caution. Operators should always look for potential sources of airwave in the area of study. Profile lines - run towards and away from a possible source to evaluate if the object is scattering energy (Annan, 2005).

Velocity Sounding Design

The velocity of an EM wave needs to be determined to convert two-way travel-time into depth. This value can be obtained by the basic process of probing or excavating to determine depth to a known reflector, then using the depth and measured reflection time to calculate the velocity (Jol and Bristow, 2003). An example of this "ground truth" approach is shown in Figure 3 where a CMP survey on the top of the outcrop indicated a surface velocity of 0.13mns^{-1} (Jol *et al.*, 2003). However, correlation with a measured section of the outcrop indicates an average velocity at depths up to 30m is 0.1mns^{-1} .

Common mid-point (CMP) sounding (Figure 1c) and wide angle reflection and refraction (WARR) surveys estimate signal velocity by increasing the separation between the transmitter and receiver in steps at a fixed location and measuring the change of the two-way travel-time to reflections (Jol and Bristow, 2003; Annan, 2005).

CMP surveys are generally considered to be more precise than WARR surveys and should be the first survey completed on arrival at a site (Jol and Bristow, 2003). Jol and Bristow (2003) suggest to start a CMP survey with a minimum antenna separation then move each antenna out by 5 or 10cm until no signal is returned. This however is only appropriate when performed above a horizontal reflector in multiple directions. Some software programs can measure the angle of hyperbolic reflections as a proxy for the moveout encountered in CMP, calculating velocity by fitting curves to observed hyperbolas as illustrated in Figure 2b and Figure 2c.

Data Processing

GPR data processing can be a daunting task for new users. This section provides good practice guidelines to ensure consistent, efficient and realistic processing. A brief overview of several basic processing techniques is provided.

It is important to keep processing simple. The majority of GPR data collected in sedimentological environments only requires minimal processing to allow interpretation (Jol and Bristow, 2003). The temptation to overprocess should be avoided as more sophisticated processing methods are likely to introduce bias and potential artefacts into the data (Cassidy, 2009). The user should take a systematic and consistent approach to the processing sequence, accurately recording all processing steps and parameters. GPR utilises many aspects of seismic data analysis that have been developed to industry-standard. Although not all seismic processing methods can be applied to GPR data, the majority can be used directly (Annan, 2005).

Basic processing steps, usually directly applied to the raw data, typically take the form of trace editing, filtering or data

correction, introducing minimal user bias. Most, if not all, of these steps need to be used to make a basic interpretation. It must

be remembered that no amount of processing will save poor quality data (Cassidy, 2009).

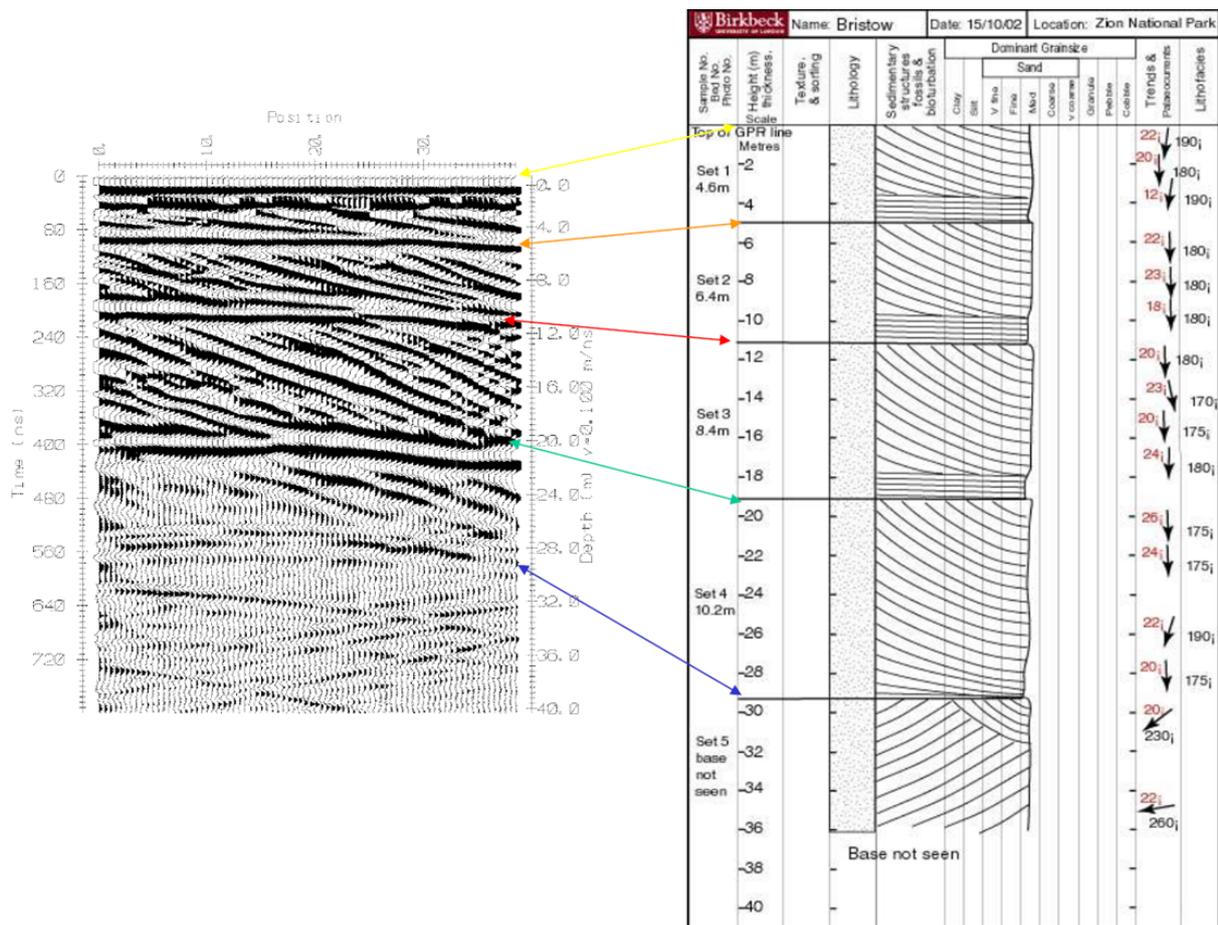


Figure 3. 50 MHz GPR profile of the Navajo Sandstone, a Jurassic cross-stratified aeolian sandstone at Zion National Park in Utah (adapted from Jol *et al.*, 2003). Reflections of cross-stratification and the bounding surfaces between them are clearly imaged by the GPR to a depth of almost 30m. A photograph of the outcrop is shown in Figure 8. Beneath 30m depth the arc shaped crossing reflections are reflections through the air from the canyon walls.

Data Editing

Data editing, often the most time consuming part of a processing sequence, involving tasks such as data reorganisation, data file merging and updating background information (Annan, 2005). The occurrence of inevitable errors during data acquisition means some traces may need to be reversed, merged or omitted. The maintenance of data is important when trying to obtain good quality interpretations (Cassidy, 2009).

Time-zero Correction

Traces need to be adjusted to a common time-zero position before processing can be

applied. This usually involves setting a particular rule such as the time-zero position is the airwave first break point or the first negative peak of the trace. Successful realignment causes all reflections beneath to become correctly aligned. Time-zero correction is therefore often the first data processing step performed (Neal, 2004). The processing software often performs this automatically (Cassidy, 2009).

Dewow Filtering

Dewowing is a basic processing step which uses temporal filtering to remove very low frequency components from the data (Annan, 2005). It reduces the data to a mean zero level, with the majority of modern GPR

systems applying dewow to each trace automatically (Cassidy, 2009). For advice on manual dewow correction see Cassidy (2009).

Topographic Correction

Topographic correction is required to place the GPR data within its correct spatial context (Cassidy, 2009). When the surface and subsurface stratigraphy are horizontal, elevation static corrections can be used to topographically correct data, repositioning the time zero in the vertical axis and adjusting reflections accordingly (Bristow, 2009). Dipping reflections can be restored to the correct dip by migrating the data, then applying the static correction for topography (Bristow, 2009). For information on advanced methods, such as migration, see Cassidy (2009).

Gain

The next basic processing step is to select a gain function for the data. Gains improve the visual form of the GPR sections, with most techniques altering the data structure in some form. It is therefore important to understand the effects of gain functions. Radar signals are rapidly attenuated as they propagate through the subsurface, making events from greater depths more difficult to discriminate (Annan, 2005). Gains enhance the appearance of later arrivals due to the effect of signal attenuation and geometrical spreading losses (Cassidy, 2009). There are several different types of gains including constant gain, exponential gain, exponential gain compensation (SEC) and automatic gain control (AGC), with each function having different characteristics. Gain functions can be easily changed, usually by altering the time window (a region of the trace in time), the gain function (linear, exponential, user-defined, etc) and the maximum gain allowed (Cassidy, 2009).

Constant, linear or exponential gains are systemically applied gain functions that have a specific mathematical operator that is defined by the user or system automatically (Cassidy, 2009).

Automatic gain control (AGC) and spherical and exponential gain compensation (SEC) are two of the most popular types of gain

function. AGC equalises the amplitudes down each trace, making it ideal for monitoring stratigraphic horizon continuity as well as the continuity of other reflections. This function however eliminates all amplitude information (Jol and Bristow, 2003).

SEC attempts to emulate the variation in signal amplitude as it travels through the subsurface (Annan, 2005). Unlike AGC, SEC retains the relative amplitude information, with the reflections representing the true strength of the returned signal (Cassidy, 2009). This however only really applies if you have equivalent gains at the same depth.

Filtering

Filters are applied to remove system or human-induced noise and improve the visual quality of the GPR data (Cassidy, 2009). There are many different types of filters, from simple band-pass filters to sophisticated domain and transform filters. Simple filters are usually very effective at removing high/low-frequency noise, while sophisticated functions are more appropriate for specific problems. Filters can be applied before or after gains however pre-gain filters operate on the data in its truest form (Cassidy, 2009).

Filters can average down the trace (temporal) or from trace to trace (spatial), smoothing the data and removing high frequency noise (Bristow, 2013). Temporal filters are used to remove noise at frequencies that are higher or lower than the main GPR signal, ultimately acting as clean-up filters which make the GPR section visually better (Jol and Bristow, 2003). Spatial filtering is applied to either suppress or emphasise specific features. These filters are often used to remove the strong air/wave response and ringing from GPR data (Cassidy, 2009). Low pass spatial filters are suitable when trying to identify sedimentological features of interest, such as bedding, which are usually spatially extensive and low-angled (Cassidy, 2009). A high pass spatial filter has the opposite effect in that it suppresses flat-lying reflections and emphasises dipping events (Annan, 2005).

High pass filters (frequency domain filters) allow higher frequency components to pass while removing low frequency components - low pass filters do the opposite. High-pass and low-pass filters can be combined in

band-pass filters, letting through frequencies either side of the peak frequency of the transmitted signal (Bristow, 2013). More information and advice on when to use certain filters can be found in Cassidy (2009).

Background Subtraction

Background removal is one of the most common processing steps applied to GPR data (Annan, 2005). It often takes the form of a high pass filter or an average trace removal (a form of spatial filtering). This step allows subtle weaker signals to become visible in the processed section (Annan, 2005). Background subtraction is usually not necessary, therefore if data always requires background removal the equipment used to collect it may be flawed (Annan, 2005).

For information on advanced processing methods see Cassidy (2009).

Visualising Processed Data

The processed GPR data only becomes useful when it is available in image format (Daniels *et al.*, 1988). It is traditionally presented as a radargram, a pseudo cross sectional image comprising a horizontal distance axis against a vertical TWT axis (Figure 4; Noon and Narayanan, 2002). These can be difficult to interpret, with several closely spaced point scatterers producing reflection hyperbolas which blur the radargram (Brunzell, 1999). The identification of multiple targets within a radargram unavoidably introduces subjectivity to the data analysis process (Daniels, 2004).

A selection of GPR profiles, collected from a range of depositional environments, has been provided.

Figure 5 shows an interpreted 200 MHz GPR profile across a normal fault at Piano di Pezza in the Italian Apennines. The processed GPR profile shows a strong sub-horizontal reflection that dips gently towards the north. At its southern end, this reflection is at a depth of approximately 0.5m below the surface. Towards the northern end of the profile, the deep purple reflection is at a depth of approximately 6m from the surface at its deepest part, approximately 7m south of the footwall scarp. The change in depth to the reflector is indicative of subsidence adjacent to the fault and the accumulation of sediment in the hanging wall of the fault. Breaks in the reflections where there is a vertical offset are interpreted as faults; the displacement of the reflection is of the order of 1m indicating a fault displacement of a similar magnitude. Between the northward dipping reflections there is a “wedge” of weakly defined reflections which thins towards the south, away from the main fault scarp. Offsets in the reflections indicate approximately 0.5m vertical displacement across faults. The faulting identified on the GPR profile is dominated by vertical offset across synthetic and antithetic normal faults within and parallel to the main fault zone and this is consistent with the published trench profile (Pantosi *et al.*, 1996). The main zone of deformation is localised within 8m south of the fault on the trench and GPR profiles. For additional information on this study see Jewell and Bristow (2004).

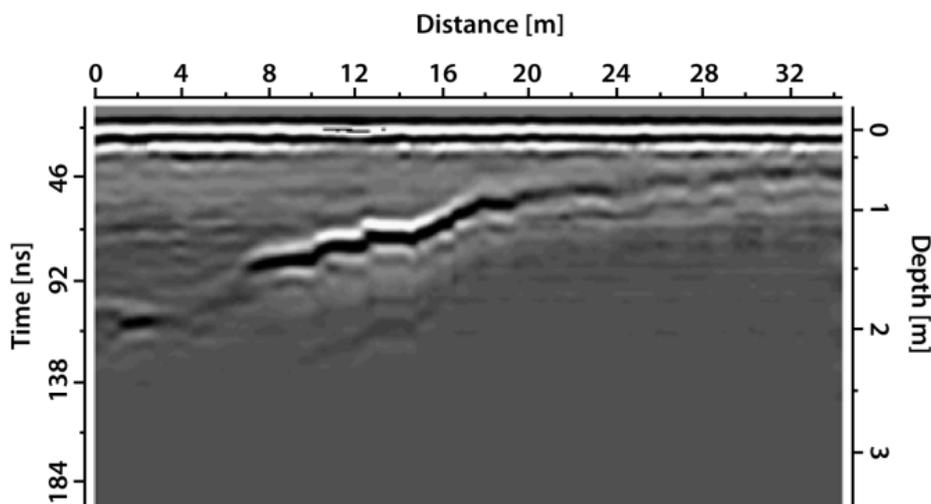


Figure 4. Radargram illustrating decrease in depth of peat

Figure 6 shows 200 MHz GPR profiles across sand dunes a linear dune in Namibia to investigate the internal structure of linear dunes from Bristow *et al.* (2000).

Figure 7 displays a 200 MHz GPR profile across a gravel beach ridge, Waitaki coast, South Island, New Zealand. The profile shows inclined reflections interpreted as sedimentary layers within the beach ridge. The reflections are divided into two radar facies shaded orange and pink. Seaward dipping reflections are interpreted to be formed by beach (foreshore) progradation (Orange), while landward dipping reflections are interpreted as backshore washover deposits (Pink). Radar sequence boundaries are picked at horizons where there are reflection terminations or truncations marking erosion of the foreshore during storms. Some of these surfaces can be correlated with surfaces within the washover deposits which are interpreted to have been formed during coeval storm events. For further information see Dickson *et al.* (2009).

As processing power has advanced, GPR data has been increasingly presented in the form of 3D models and time/depth slices (Annan, 2009). 3D surveying involves collecting data along closely spaced lines, with sufficiently dense data coverage allowing the generation of 3D block models (Jol and Bristow, 2003). These 3D models allow the user to interpret the depositional structure of stratigraphic units in greater detail than was possible using sparsely spaced 2D profiles (Jol and Bristow, 2003). Although 3D surveying can provide the best visualisation of the subsurface, the collection of 3D data is time consuming, with a considerable amount of time being required to accurately record the position and elevation of the data points (Serma and Setan, 2009; Bristow, 2009). This is why, in general, only small areas are surveyed.

In the case of Figure 8, a grid 10m x 7m was surveyed with 200MHz antennas which imaged to a depth of around 12m. This can be compared with the depth of penetration of almost 30m achieved on the same outcrop with 50MHz antennas (Figure 9). With a longer wavelength the 50MHz antennas penetrate much deeper but with an

associated loss of resolution. See figures on outcrop for scale and log on Figure 3 for depth and bedding thickness. Modified from Jol *et al.* (2003).

Interpretation

The interpretation of a GPR profile is inherently subjective, dependent on the knowledge, skill and experience of the user (Annan, 2005). GPR profiles should receive preliminary interpretation once collected to determine if resurveying is required (Jol and Bristow, 2003). After the data has been processed, an interpretation, independent of sedimentary models and other datasets such as cores, should be completed. Many interpretations based on models and other data can overlook key components that are visible in the GPR profile (Jol and Bristow, 2003). At the surface the earliest arrivals are the direct transmissions from the transmitter to the receiver through the air at the speed of light, the air wave, after this comes the direct signal through the ground or along the surface, the ground wave. These two direct signals form a pair of black and white lines along the top of all GPR profiles. They cover any reflections approximately equivalent to a depth of half the antenna spacing. Reflections from the subsurface appear beneath the air and ground waves. The GPR signal is not a simple sine wave but has a more complex waveform and as a result a single reflector can appear as a pair of reflections. This is sometimes overlooked resulting in over interpretation of GPR profiles.

Data should initially be printed out, with major reflections being identified and traced around. The next step is to identify and characterise (as downlap, onlap, offlap, etc) reflection terminations, constructing a chronology of sedimentary deposits (Jol and Bristow, 2003). Figures 7 and 9 show examples with bounding surfaces picked out in red and reflection terminations marked by small black arrows. The ability to identify significant radar reflection patterns allows the location of features such as bedding planes, bedsets and cross-stratification to be determined (Jol and Bristow, 2003).

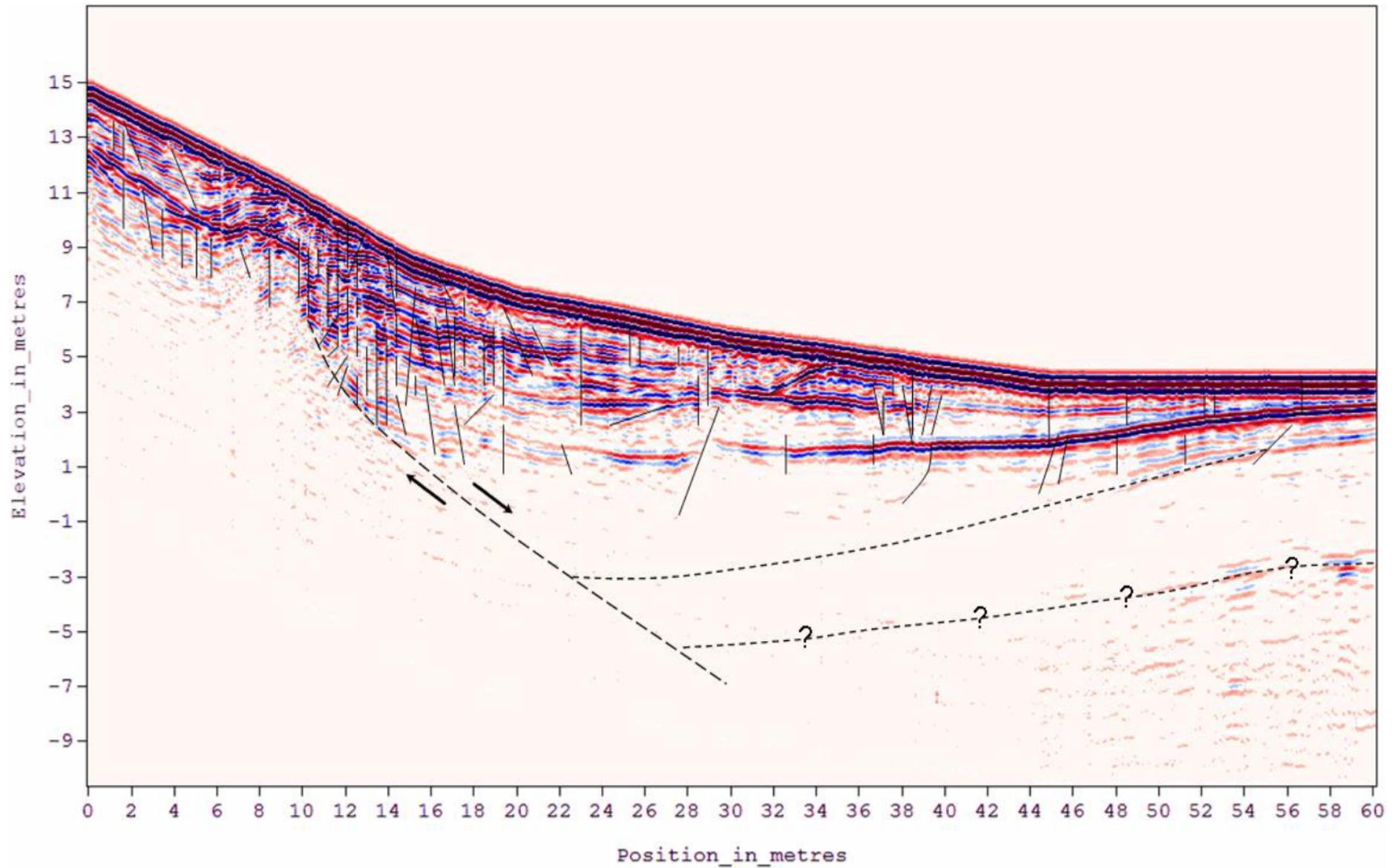


Figure 5. An interpreted 200 MHz GPR profile across a normal fault at Piano di Pezza in the Italian Apennines (modified from Jewel and Bristow, 2004).

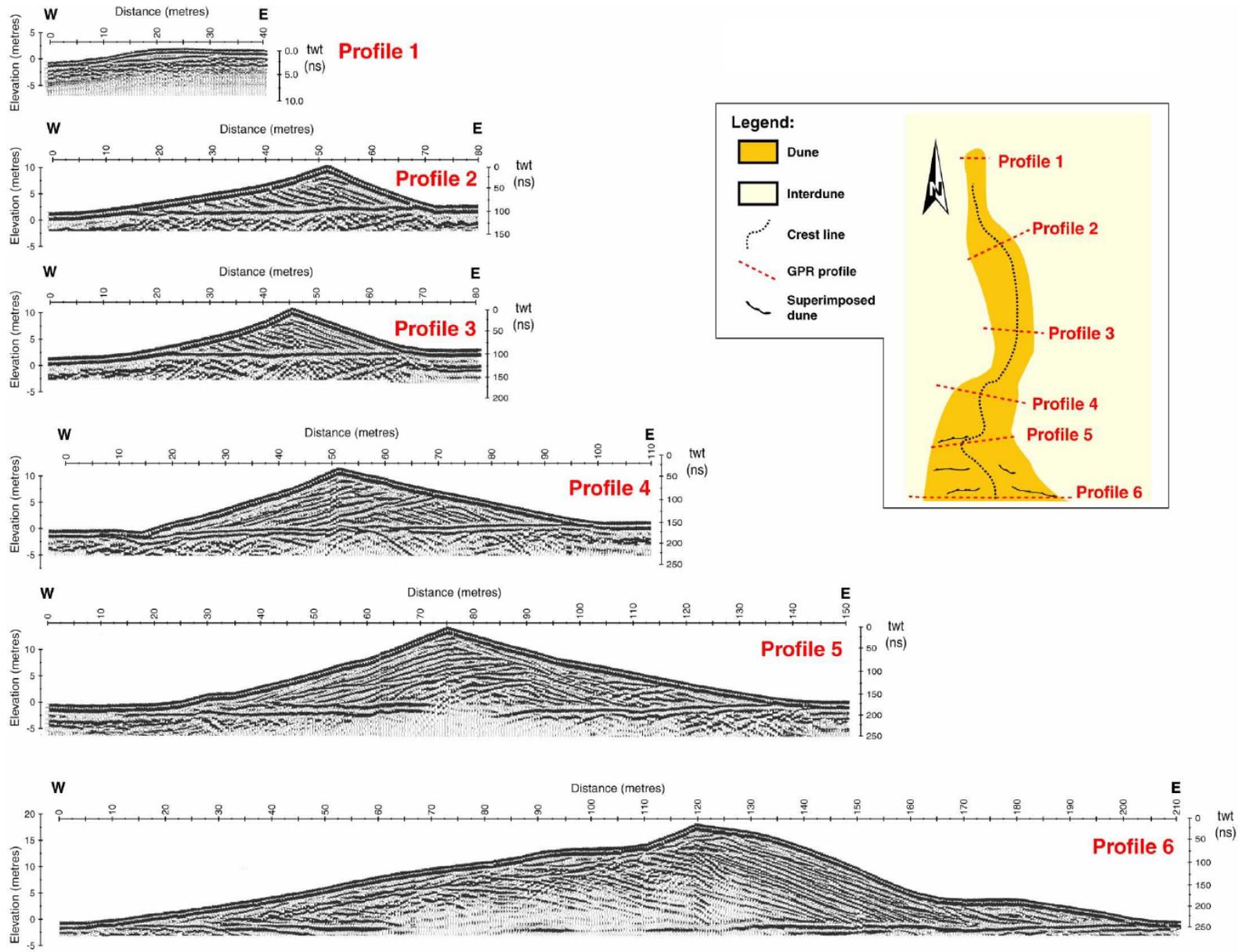


Figure 6. 200 MHz GPR profiles across sand dunes a linear dune in Namibia to investigate the internal structure of linear dunes (Bristow et al., 2005).

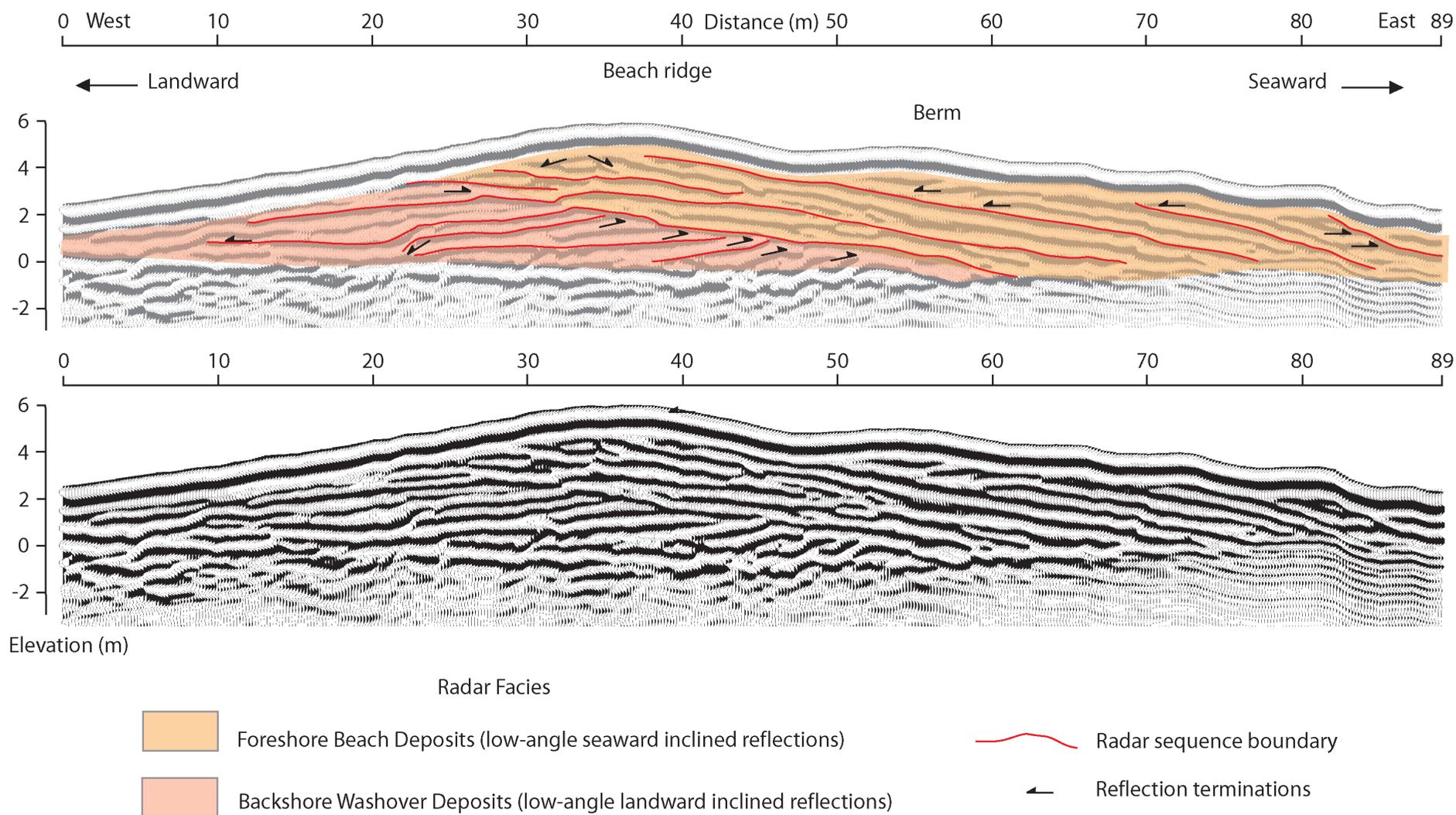


Figure 7. 200 MHz GPR profile across a gravel beach ridge, Waitaki coast, South Island, New Zealand (Dickson *et al.* 2009). Although saline sea-water will normally attenuate the GPR signal, this coarse grained, well drained gravel beach combined with high rainfall means that sea-water is flushed from the beach sediments permitting imaging of beach progrades and washovers.

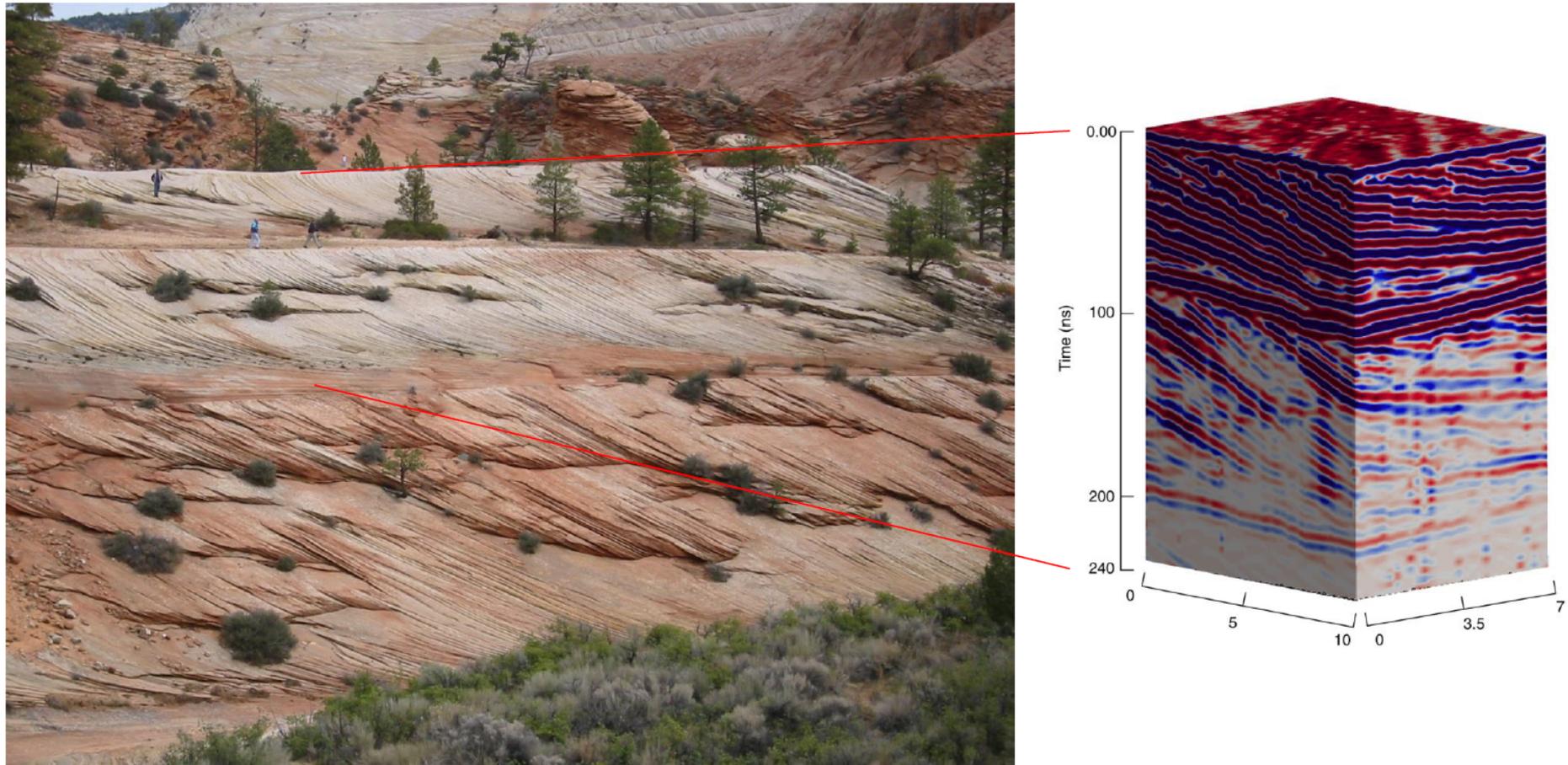


Figure 8. Site photograph accompanied by a 10m x 7m grid surveyed with 200MHz antennas, figures for scale (adapted from Jol et al., 2003). The depth of penetration (approx. 12m) and resolution of <math><0.2\text{m}</math> can be compared with the 50MHz profile of the same outcrop shown in Figure 3 where depth of penetration is approx. 30m with resolution approx. 1m.

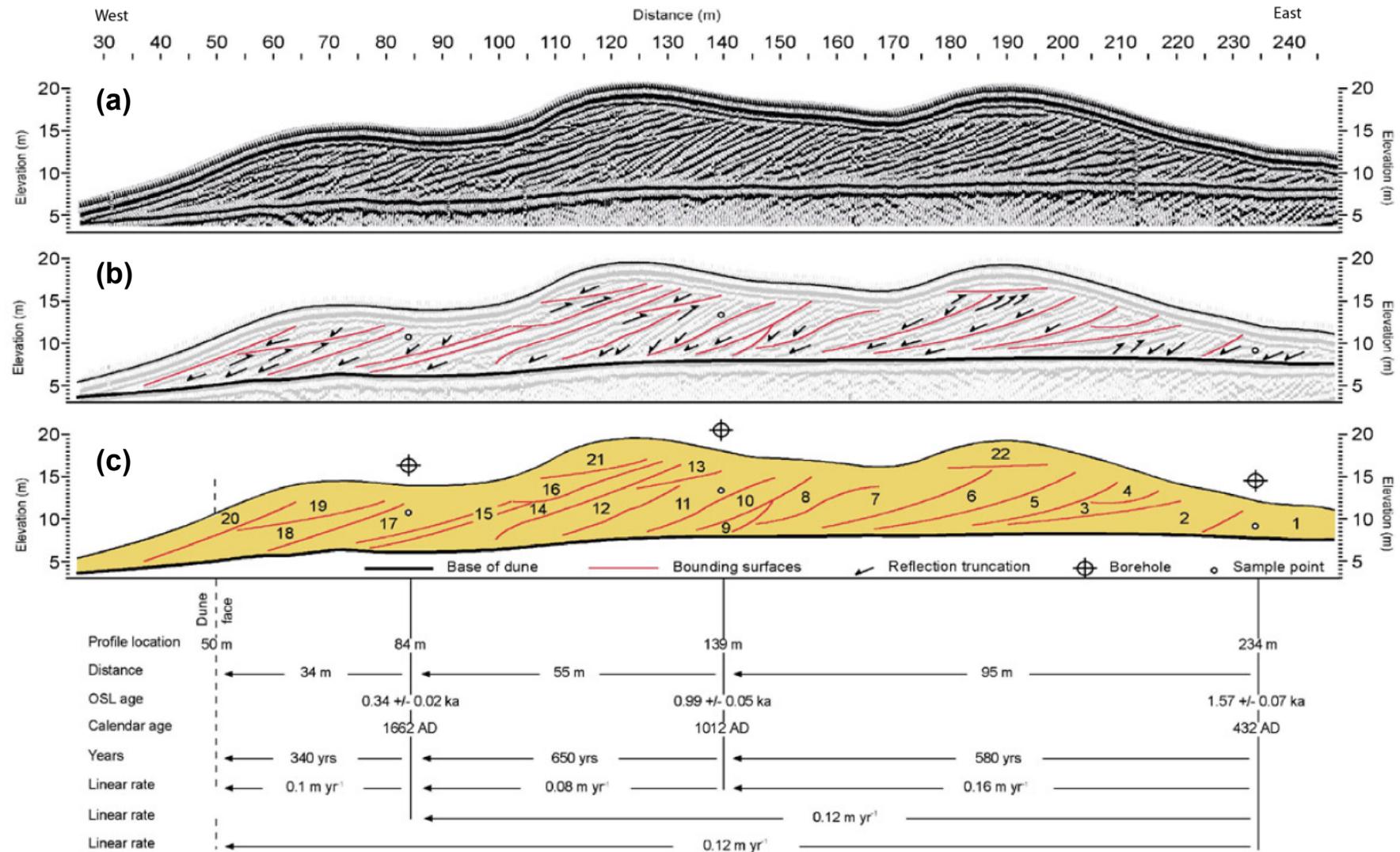


Figure 9. Image indicates how a GPR profile (a) can be interpreted to identify radar sequence boundaries (b) to derive a relative chronology (c) and select sites for optical dating to constrain rates of dune migration (Bristow *et al.*, 2005).

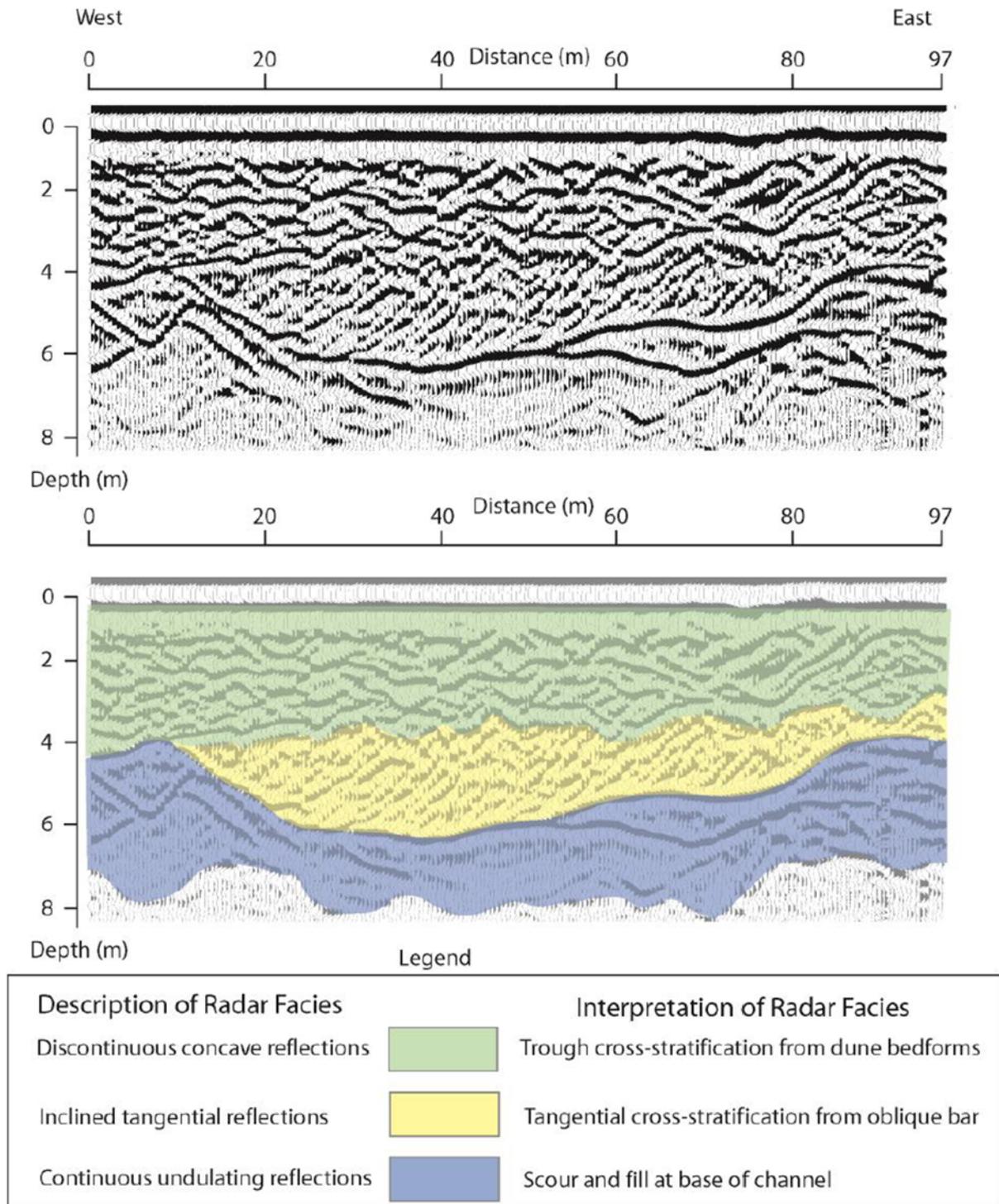


Figure 10. Annotated 100MHz GPR profile collected across part of a sand bar in the Jamuna / Brahmaputra River, Bangladesh (adapted from Best *et al.*, 2003).

Figure 9 shows a GPR profile across a sand dune in the Namib Sand Sea from Bristow *et al.* (2005). The top panel (Figure 9a) shows the GPR profile with topographic correction. The second panel (Figure 9b) shows reflection terminations marked by small arrows with truncation where arrows point up, and downlap where arrows point down. The radar sequence boundaries which are picked out in red mark breaks in deposition. On the

basis of superposition and cross-cutting relationships a relative chronology is derived numbers 1 to 22 in the yellow shaded section (Figure 9c). The interpreted GPR profile has been used to select sample points for dating in order to determine the rate of migration of the sand dune (Bristow *et al.*, 2005).

Another common interpretation strategy is to identify reflection patterns with similar

geometry which are termed radar facies (Jol and Bristow 2003). An example of radar facies interpretation is shown in Figure 10. Figure 10 shows a 100MHz GPR profile collected across part of a sand bar in the Jamuna/Brahmaputra River, Bangladesh. Three different reflection patterns have been identified and interpreted as radar facies, including: discontinuous concave reflections from sets of trough cross-stratification produced by dune bedforms migrating across the top of the sand bar; inclined tangential reflections from the lee-side slipface on the downstream margin of the bar migrating obliquely across the channel; continuous undulating reflections from scour and fill at the base of the channel (Best *et al.*, 2003).

The source of all signals should always be questioned. Novice users tend to incorrectly assume that all GPR responses originate from the subsurface. Airwave events, features associated with energy that propagates through the air and gets reflected back to the GPR receiver, can originate from above ground objects including fences, telephone poles and overhead wires (Annan, 2005). Such lateral reflections of the airwave can appear as hyperbolic features on a GPR profile (See Figure 2). Survey areas should be checked for potential sources of airwave events. If a potential source is found, profile lines should be run towards and away from the object to assess its effect on the GPR output. Users should always treat shielded antenna data with doubt as antenna shielding is never fully effective (Annan, 2005).

Case Study: Determining Peat Thickness using GPR

Aim of the Study

GPR was used to determine the thickness of peat at a site in Northern Ireland. The aim of the study was to rapidly acquire continuous peat thickness data for a profile located on an upland blanket bog (a peatland up to 6m deep) situated approximately 9km north-west of the town of Dungiven (Co.Derry). The site, a gently sloping hill facing the north and north-west, has an average topographical height of approximately 210mAOD. The underlying geology of the bog is psammite and pelite, with a superficial geology of sandy silt and sandy clay.

Existing studies (Table 1) suggest that GPR can be used effectively to identify the subsurface contact point between the base of the peat and drift material.

Survey Design

50 MHz antennas were used to ensure the maximum depth of the peat would be identifiable (~6m). Although the average peat thickness for the site was approximately 1.5m, blanket bogs can reach depths of 6m. The employment of an antenna frequency of 100MHz, reaching an approximate depth of 5m (Table 3), may have resulted in the target (peat and subsoil interface) not being detected.

The survey was completed using a MALÅ 50 MHz Rough Terrain Antenna (RTA). The flexible snake-like design of the antenna (comprising both a transmitter and receiver) provides optimum results in difficult environments, with the equipment allowing the operator to perform a survey without having to clear an access route. This system is therefore ideal for surveying peatlands affected by turf cutting (removal of peat for fuel or horticulture).

Continuous data collection was used in this survey, with a step size therefore not being specified. A time window of 405ns was employed, with the time trigger interval being every half second and traces being stacked by 16 - settings were selected based on the experience of GPR users at Queen's University Belfast. The antennas of the MALÅ system have a fixed separation meaning the EM wave velocity could not be determined using CMP sounding. Velocity was determined through the manual probing of the transect line, a process providing accurate information on the depth of the subsurface reflector. A series of connected probes were driven into the peat until a significant change in resistance was encountered. The peat probe was then removed and measured using a tape. Using the thickness of the peat and the basal reflection on the GPR profile, a velocity of 0.037mns^{-1} was calculated.

Due to the survey environment, it was not possible to collect data in a straight line. Differential global positioning system (dGPS)

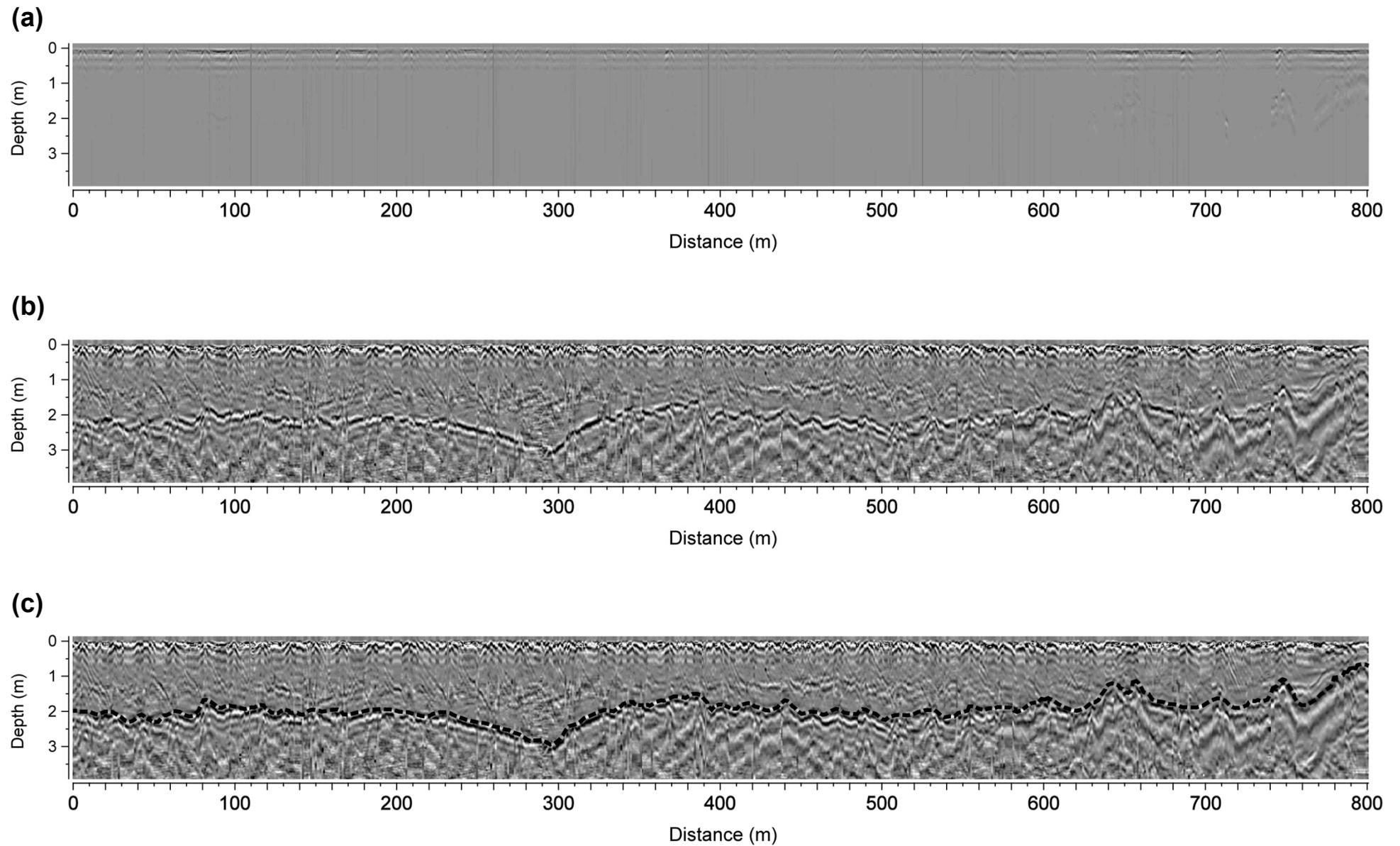


Figure 11. (a) Unprocessed radargram of case study transect; (b) Processed radargram of case study transect; (c) Processed radargram of case study transect with base of peat labelled.

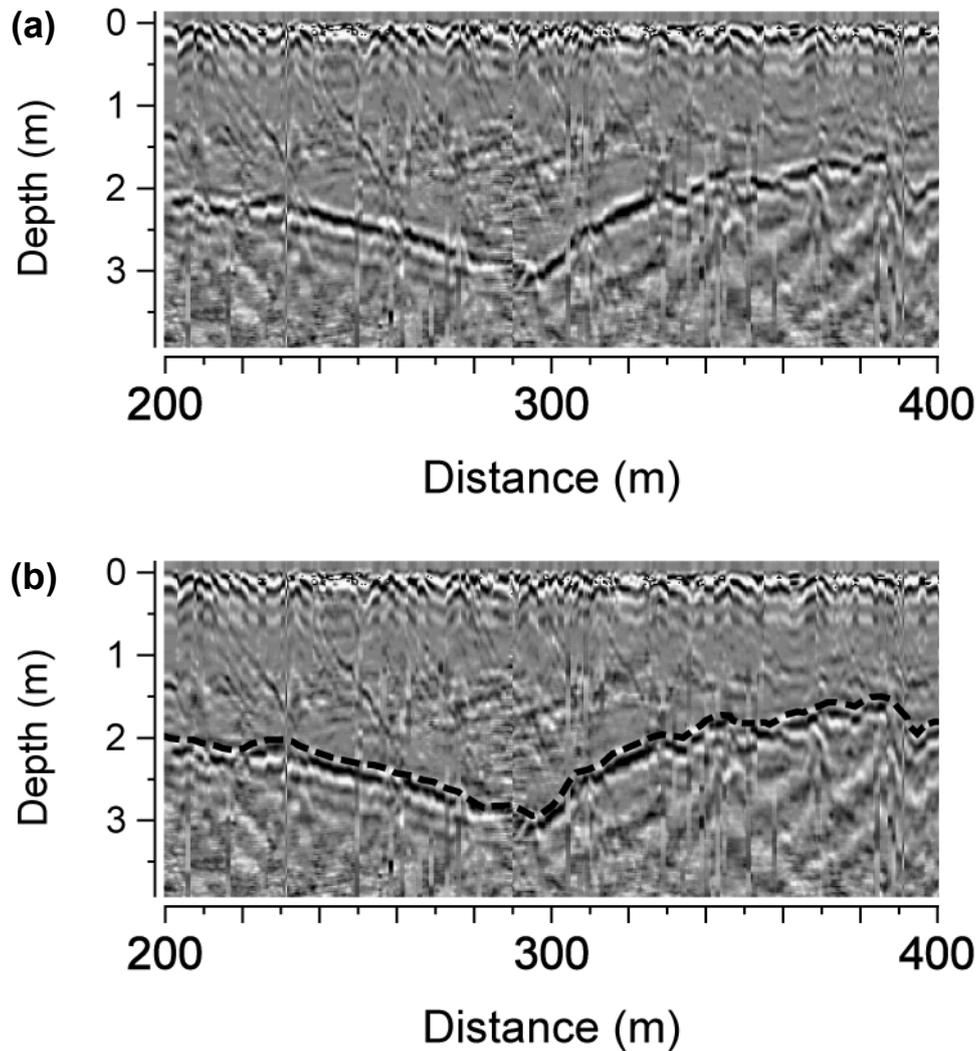


Figure 12. (a) Processed radargram of case study transect section (200m - 400m); (b) Processed radargram of case study transect section (200m - 400m) with base of peat labelled

points were collected approximately every 20m along the profile using a Leica SR530 rover system, with the degree of sinuosity in the survey line being investigated in ArcGIS. It was determined that each dGPS point, on average, deviated from the predetermined survey line by approximately 2.50m, with the greatest deviation being 8.07m. This degree of sinuosity in the profile was deemed acceptable, considering the extent of large-scale vegetation and safety issues related to historical peat cutting.

Data Processing

Minimal processing was applied to the collected data before interpretation. De-wowing and time-zero correction were automatically applied to the data. Automatic gain control (AGC) was employed,

accentuating stratigraphic horizon continuity in the radargram. Background subtraction, a common processing step, was then applied, removing noise from the data. Down-the-trace averaging (running average) was used to reduce random noise, acting as a low pass filter.

Visualising Processed Data

The data from the transect has been presented in the form of a radargram, with Figures 11a and 11b showing the unprocessed and processed data, respectively. Figure 11c shows the processed radargram, annotated with a broken line indicating the estimated base of the peat. Figures 12a and 12b show a subsection of the line, allowing reflections to be viewed in

greater detail. Depth scale was determined using manual probe measurements.

The radargram was printed, with the first continuous major reflection, the contact between the base of the peat and drift material, being traced. Although the data were not collected within close proximity of any metallic fencing or buildings, the profile was situated adjacent to a nearby windfarm, with the closest turbine being approximately 200m. The wind turbine, located approximately 150m along the profile, does not appear to have a noticeable impact on the radargram. This is due to the exponential decay of the emitted EM wave, with Table 3 clearly indicating that the footprint of a 50MHz antenna would not incorporate a wind turbine located over 200m from the survey line.

Summary

The brief case study provided shows how GPR can be successfully utilised in environmental studies, providing a continuous profile of varying peat thickness values for a selected case study site. Minimal processing was successfully used to indicate the boundary between overlying peat and the superficial geology, with the visualised profile being very useful in peatland studies.

Advantages and Limitations

A major advantage of GPR is that the scale of resolvable features can range from several centimetres to hundreds of metres using a selection of available antennas (Mellet, 1995). GPR systems have the ability to rapidly collect large amounts of continuous subsurface data (Hruska *et al.*, 1999; Doolittle and Collins, 1995). Most GPR techniques also allow acquired data to be reviewed on-site for quality control purposes (Mellet, 1995). The ability to use remote non-contacting transducers of radiated energy, as opposed to ground contacting types needed for seismic studies, is a significant advantage that GPR has over seismic techniques.

A significant limitation of GPR is that the antennas transmit and receive EM energy in a complex 3D cone, with reflections originating from anywhere on the radar wave front (Neal, 2004). The impression that a

reflection on a 2D radar profile is obtained from directly beneath the survey point is incorrect. Out-of-plane reflections from isolated point reflectors and reflector surfaces (known as "side swipes") are difficult to identify on reflection profiles (Neal, 2004). This can lead to errors in depth estimation and interpretation (Olhoeft, 1998) - see Figures 2 and 3 for examples of out of plane, airwave reflections.

A major limitation of GPR is that the performance of a system can be seriously affected by the environment it is surveying, with conductive soils strongly attenuating EM waves. Rapp and Hill (2006) states that a 100MHz GPR can penetrate approximately 15m of dry sandy soil or sediment compared to as little as 1m of wet clayey soil. The rapid attenuation of EM waves by conductive fine-grained sediments can produce high frequency ringing which can subsequently obscure primary reflections (Neal, 2004). Moorman *et al.* (2003) indicates that problems can occur when using GPR to survey permafrost environments, with the noise created by diffractions from small ice lenses resulting in deeper structures being masked. Another problem is that limitations that are dependent on the applied data collection configuration cannot be overcome once the basic GPR survey data has been collected (Neal, 2004).

Conclusion

GPR, a geophysical technique with multidisciplinary applications, has become a valuable tool in sedimentological studies. This article has provided good practice guidelines on the collection, processing and interpretation of GPR data. A variety of survey design factors were proposed including operating frequency, temporal sampling interval, step size and antenna separation. Basic data processing steps have been reviewed, with information also being provided on data visualisation and interpretation. Possible applications of GPR in geomorphological research have been presented, with a case study outlining how GPR has been used to measure peat thickness being provided.

Acknowledgements

I would like to thank Dr. Alastair Ruffell and Dr. Jennifer McKinley for their guidance during the preparation of this article. Special thanks go to Prof. Charlie Bristow for providing some excellent GPR images from a range of environmental settings - they are greatly appreciated.

References

- Annan AP. 2009. Electromagnetic principles of ground penetrating radar. In *Ground penetrating radar: theory and applications*, Jol HM (ed). Elsevier: Amsterdam; 3-40.
- Annan AP. 2005. Ground-penetrating radar. In *Near surface geophysics*, Butler DK (ed). Society of exploration geophysicists: Tulsa, Investigations in Geophysics **13**; 357-438.
- Annan AP. 2001. *Ground penetrating radar workshop notes*. Sensors and Software Inc: Mississauga.
- Annan AP and Cosway SW. 1994. GPR frequency selection. In *Proceedings of the fifth international conference on ground penetrating radar*, Redman JD (ed). Waterloo Centre for Groundwater Research: Kitchener; 747-760.
- Appleby JR, Brook MS, Vale SS and MacDonald-Creevey AM. 2010. Structural glaciology of a temperate maritime glacier: Lower Fox glacier, New Zealand. *Geografiska annaler: series a, physical geography* **92**: 451-467.
- Ashworth PJ, Sambrook Smith GH, Best JL, Bridge JS, Lane SN, Lunt IA, Reesink AJH, Simpson CJ and Thomas RE. 2011. Evolution and sedimentology of a channel fill in the sandy braided South Saskatchewan River and its comparison to the deposits of an adjacent compound bar. *Sedimentology* **58**: 1860-1883.
- Baker GS, Jordan TE and Pardy J. 2007. An introduction to ground penetrating radar (GPR). *The geological society of America special papers* **432**: 1-18.
- Barnhardt WA and Sherrod BL. 2006. Evolution of a Holocene delta driven by episodic sediment delivery and coseismic deformation, Puget Sound, Washington, USA. *Sedimentology* **53**: 1211-1228.
- Benediktsson ÍÖ, Ingólfsson Ó, Schomacker A and Kjær KH. 2009. Formation of submarginal and proglacial end moraines: implications of ice-flow mechanism during the 1963-64 surge of Bárjökull, Iceland. *Boreas* **38**: 440-457.
- Bennett MR, Cassidy NJ and Pile J. 2009. Internal structure of a barrier beach as revealed by ground penetrating radar (GPR): Chesil beach, UK. *Geomorphology* **104**: 218-229.
- Best JL, Ashworth PJ, Bristow CS and Roden J., 2003. Three dimensional sedimentary architecture of a large, mid-channel sand braid bar, Jamuna River, Bangladesh. *Journal of sedimentary research* **73**: 516-530.
- Bhosle B, Parkash B, Awasthi AK, Singh VN and Singh S. 2007. Remote sensing-GIS and GPR studies of two active faults, Western Gangetic Plains, India. *Journal of applied geophysics* **61**: 155-164.
- Bristow CS. 2013. Ground penetrating radar. In *Treatise on geomorphology, vol 14*, Shroder JF (ed). Academic Press: San Diego; 183-194.
- Bristow CS. 2009. Ground penetrating radar in aeolian dune sands. In *Ground penetrating radar: theory and applications*, Jol HM (ed). Elsevier: Amsterdam; 273-298.
- Bristow CS and Pucillo K. 2006. Quantifying rates of coastal progradation from sediment volume using GPR and OSL: the Holocene fill of Guichen Bay, south-east South Australia. *Sedimentology* **53**: 769-788.
- Bristow CS, Augustinus PC, Wallis IC, Jol HM and Rhodes EJ. 2010. Investigation of the age and migration of reversing dunes in Antarctica using GPR and OSL, with implications for GPR on Mars. *Earth and planetary science letters* **289**: 30-42. doi:10.1016/j.epsl.2009.10.026
- Bristow CS, Duller GAT and Lancaster N. 2007. Age and dynamics of linear dunes in the Namib Desert. *Geology* **35**:555-558.
- Bristow CS, Lancaster N and Duller GAT. 2005. Combining ground penetrating radar surveys and optical dating to determine dune migration in Namibia. *Journal of the geological society* **162(2)**: 315-322.
- Bristow CS, Bailey SD and Lancaster N. 2000. The sedimentary structure of linear sand dunes. *Nature* **406**: 56-59.

- Brunzell H. 1999. Detection of shallowly buried objects using impulse radar. *IEEE transactions on geoscience and remote sensing* **37**: 875-886.
- Buynevich IV, Filho PWMS and Asp NE. 2010. Dune advance into a coastal forest, equatorial Brazil: A subsurface perspective. *Aeolian research* **2**: 27-32.
- Cassidy NJ. 2009. Ground penetrating radar data processing, modelling and analysis. In *Ground penetrating radar: theory and applications*, Jol HM (ed). Elsevier: Amsterdam; 141-176.
- Christie M, Tsoflias GP, Stockli DF and Black R. 2009. Assessing fault displacement and off-fault deformation in an extensional tectonic setting using 3-D ground-penetrating radar imaging. *Journal of applied geophysics* **68**: 9-16.
- Clemmensen LB, Nielsen L, Bendixen M and Murray A. 2012. Morphology and sedimentary architecture of a beach-ridge system (Anholt, the Kattegat sea): a record of punctuated coastal progradation and sea-level change over the past ~1000 years. *Boreas* **41**: 422-434.
- Comas X, Slater L and Reeve A. 2011. Pool patterning in a northern peatland: geophysical evidence for the role of postglacial landforms. *Journal of hydrology* **399**: 173-184.
- Comas X, Slater L and Reeve A. 2005. Stratigraphic controls on pool formation in a domed bog inferred from ground penetrating radar (GPR). *Journal of hydrology* **315**: 40-51.
- Comas X, Slater L and Reeve A. 2004. Geophysical evidence for peat basin morphology and stratigraphic controls on vegetation observed in a northern peatland. *Journal of hydrology* **295**: 173-184.
- Conyers LB (2013) *Ground-penetrating radar for archaeology (3rd edition)*. AltaMira Press: Plymouth.
- Daniels DJ. 2004. *Ground penetrating radar (2nd edition)*. The Institution of Electrical Engineers: London.
- Daniels DJ, Gunton DJ and Scott HF. 1988. Introduction to subsurface radar. *IEE proceedings f: radar and signal processing* **135**: 278-320.
- De Oliveira MAT, Porsani JL, de Lima GL, Jeske-Pieruschka V and Behling H. 2012. Upper Pleistocene to Holocene peatland evolution in Southern Brazilian highlands as depicted by radar stratigraphy, sedimentology and palynology. *Quaternary research* **77**: 397-407.
- Degenhardt JJ. 2009. Development of tongue-shaped and multilobate rock glaciers in alpine environments – interpretations from ground penetrating radar surveys. *Geomorphology* **109**: 94-107.
- Dickson ME, Bristow CS, Hicks DM, Jol H, Stapleton J, Todd D. 2009. Beach volume on an eroding sand-gravel coast determined using ground penetrating radar. *Journal of coastal research* **25**: 1149-1159.
- Doolittle JA and Collins ME. 1995. Use of soil information to determine application of ground penetrating radar. *Journal of applied geophysics* **33**: 101-108.
- Gibbard PL, West RG, Boreham S and Rolfe CJ. 2012. Late Middle Pleistocene ice-marginal sedimentation in East Anglia, England. *Boreas* **41**: 319-336.
- Grasmueck M, Weger R and Horstmeyer H. 2005. Full resolution 3D GPR imaging. *Geophysics* **70**: k12-k19.
- Gutsell JE, Clague JJ, Best ME, Bobrowsky PT and Hutchinson I. 2004. Architecture and evolution of a fjord-head delta, western Vancouver Island, British Columbia. *Journal of Quaternary science* **19**: 497-511.
- Harari Z. 1996. Ground-penetrating radar (GPR) for imaging stratigraphic features and groundwater in sand dunes. *Journal of applied geophysics* **36**: 43-52.
- Hart JK, Rose KC and Martinez K. 2011. Temporal englacial water content variability associated with a rapidly retreating glacier. *Earth surface processes and landforms* **36**: 1230-1239.
- Holden J, Burt TP and Vilas M. 2002. Application of ground-penetrating radar to the identification of subsurface piping in blanket peat. *Earth surface processes and landforms* **27**: 235-249.
- Hruska J, Cermák J and Sustek S. 1999. Mapping tree root systems with ground-penetrating radar. *Tree physiology* **19**: 125-130.

- Irvine-Fynn TDL, Moorman BJ, Williams JLM and Walter FSA. 2006. Seasonal changes in ground-penetrating radar signature observed at a polythermal glacier, Bylot Island, Canada. *Earth surface processes and landforms* **31**: 892-909.
- Jewell C and Bristow CS. 2004. GPR studies in the Piano di Pezza area of the Ovindoli-Pezza Fault, Central Apennines, Italy: Extending palaeoseismic trench investigations with high resolution GPR. *Proceedings of the Tenth International Conference on Ground Penetrating Radar, Delft, The Netherlands*: 555-558 IEEEExplore.
- Johnson BA and Carpenter PJ. In Press. Geophysical response of slackwater and sandy terrace deposits near Savanna, Northwestern Illinois. *Environmental earth sciences*. DOI 10.1007/s12665-012-1798-5
- Jol HM and Bristow CS. 2003. GPR in sediments: advice on data collection, basic processing and interpretation, a good practice guide. In *Ground penetrating radar in sediments*, Bristow CS and Jol HM (eds). Geological Society: London, Special Publication **211**; 9-28.
- Jol HM and Smith DG. 1991. Ground penetrating radar of northern lacustrine deltas. *Canadian journal of earth sciences* **28**: 1939-1947.
- Jol HM, Bristow CS, Smith DG, Junck B and Putnam P. 2003. Stratigraphic imaging of the Navajo Sandstone using ground-penetrating radar. *The leading edge, September 2003*: 882-887.
- Kettridge N, Comas X, Baird A, Slater L, Strack M, Thompson D, Jol H and Binley A. 2008. Ecohydrologically important subsurface structures in peatlands revealed by ground-penetrating radar and complex conductivity surveys [on-line]. *Journal of geophysical research* **113**.
- Kim KY, Lee J, Hong MH, Hong JK, Jin YK and Shon H. 2010. Seismic and radar investigations of Fourcade Glacier on King George Island, Antarctica. *Polar research* **29**: 298-310.
- Kostic B and Aigner T. 2007. Sedimentary architecture and 3D ground-penetrating radar analysis of gravelly meandering river deposits (Neckar Valley, SW Germany). *Sedimentology* **54**: 789-808.
- Langston G, Bentley LR, Hayashi M, McClymont A and Pidlisecky A. 2011. Internal structure and hydrological functions of an alpine proglacial moraine. *Hydrological processes* **25**: 2967-2982.
- Leckebusch J. 2003. Ground-penetrating radar: a modern three-dimensional prospection method. *Archaeological Prospection* **10**: 213-240.
- Leopold M, Williams MW, Caine N, Völkel J and Dethier D. 2011. Internal structure of the Green Lake 5 rock glacier, Colorado Front Range, USA. *Permafrost and periglacial processes* **22**: 107-119.
- Lowry CS, Fratta D and Anderson MP. 2009. Ground penetrating radar and spring formation in a groundwater dominated peat wetland. *Journal of hydrology* **373**: 68-79.
- Lunt IA and Bridge JS. 2004. Evolution and deposits of a gravelly braid bar, Sagavanirktok River, Alaska. *Sedimentology* **51**: 415-432.
- Malik JN, Sahoo AK, Shah AA, Shinde DP, Juyal N and Singhvi AK. 2010. Paleoseismic evidence from trench investigation along Hajipur fault, Himalayan Frontal Thrust, NW Himalaya: implications of the faulting pattern on landscape evolution and seismic hazard. *Journal of structural geology* **32**: 350-361.
- Mellet JS. 1995. Ground penetrating radar applications in engineering, environmental management, and geology. *Journal of applied geophysics* **33**: 157-166.
- Monnier S, Camerlynck C, Rejiba F, Kinnard C, Feuillet T and Dhemaied A. 2011. Structure and genesis of the Tabor rock glacier (Northern French Alps) determined from morphological and ground-penetrating radar surveys. *Geomorphology* **134**: 269-279.
- Moorman BJ, Robinson SD and Burgess MM. 2003. Imaging periglacial conditions with ground-penetrating radar. *Permafrost and periglacial processes* **14**: 319-329.
- Murray T and Booth AD. 2010. Imaging glacial sediment inclusions in 3-D using ground-penetrating radar at Kongsvegen, Svalbard. *Journal of Quaternary science* **25**: 754-761.
- Neal A. 2004. Ground-penetrating radar and its use in sedimentology: principles, problems and progress. *Earth science reviews* **66**: 261-330.

- Nielsen L and Clemmensen LB. 2009. Sea-level markers identified in ground-penetrating radar data collected across a modern beach ridge system in a microtidal regime. *Terra nova* **21**: 474-479.
- Nobes DC, Ferguson RJ and Brierley GJ. 2001. Ground-penetrating radar and sedimentological analysis of Holocene floodplains: insight from the Tuross valley, New South Wales. *Australian journal of earth sciences* **48**: 347-355.
- Noon DA and Narayanan RM. 2002. Subsurface remote sensing. In *Review of radio science 1999-2002* Stone WR (ed). Wiley-IEEE: USA; 535-552.
- Olhoeft GR. 1998. Electrical, magnetic and geometric properties that determine ground penetrating radar performance. *Proceedings of GPR'98, seventh international conference on ground penetrating radar*. University of Kansas: Lawrence; 477-483.
- Olson D, Kennedy DM, Dawe I and Calder M. In Press. Decadal-scale gravel beach evolution on a tectonically-uplifting coast: Wellington, New Zealand. *Earth surface processes and landforms*. DOI: 10.1002/esp.3233
- Pantosti D, D'Addezio G and Cinti FR. 1996. Paleoseismicity of the Ovindoli-Pezza fault, central Apennines, Italy: a history including a large, previously unrecorded earthquake in the Middle Ages (860–1300 A.D.). *Journal of geophysical research: solid earth* **101(B3)**: 5937–5960.
- Pascucci V, Martini IP and Endres AL. 2009. Facies and ground-penetrating radar characteristics of coarse-grained beach deposits of the uppermost Pleistocene glacial Lake Algonquin, Ontario, Canada. *Sedimentology* **56**: 529-545.
- Plado J, Sibul I, Mustasaar M and Jõelet A. 2011. Ground-penetrating radar study of the Rahivere peat bog, eastern Estonia. *Estonian journal of earth sciences* **60**: 31-42.
- Powers MH. 1997. Modelling frequency-dependent GPR. *The leading edge* **16**: 1657-1662.
- Rapp G and Hill CL. 2006. *Geoarchaeology: the earth-science approach to archaeological interpretation (2nd edition)*. Yale University Press: London.
- Rice SP, Church M, Wooldridge CL and Hickin EJ. 2009. Morphology and evolution of bars in a wandering gravel-bed river; lower Fraser river, British Columbia, Canada. *Sedimentology* **56**: 709-736.
- Rosa E, Larocque M, Pellerin S, Gagné S and Fournier B. 2009. Determining the number of manual measurements required to improve peat thickness estimations by ground penetrating radar. *Earth surface processes and landforms* **34**: 377-383.
- Serma AIA and Setan H. 2009. Ground penetrating radar (GPR) for subsurface mapping: preliminary result. *Geoinformation science journal* **9**: 45-62.
- Słowik M. 2011. Reconstructing migration phases of meandering channels by means of ground-penetrating radar (GPR): the case of the Obra River, Poland. *Journal of soils and sediments* **11**: 1262-1278.
- Tamura T, Kodama Y, Bateman MD, Saitoh Y, Watanabe K, Matsumoto D and Yamaguchi N. 2011. Coastal barrier dune construction during sea-level highstands in MIS 3 and 5a on Tottori coast-line, Japan. *Palaeogeography, palaeoclimatology, palaeoecology* **308**: 492-501.
- Vriend NM, Hunt ML and Clayton RW. 2012. Sedimentary structure of large sand dunes: examples from Dumont and Eureka dunes, California. *Geophysical journal international* **190**: 981-992.

2.1.3. dGPS

Elizabeth J Young^{1,2}

¹ School of the Environment, University of Dundee (e.j.young@dundee.ac.uk)

² the James Hutton Institute, Invergowrie



ABSTRACT: Direct acquisition of elevation data is an important aspect of many geomorphological studies, particularly those which include mapping of features, or investigation of topographic change. Differential global positioning systems (dGPS) provide an accurate and efficient method of obtaining elevation data, and have been used in a range of environments. The basic methodology for conducting a dGPS survey is summarised, followed by discussion of a range of adaptations to the basic system to enhance productivity. Finally, the possibilities for integrating dGPS with other methods are considered. dGPS is an accurate and versatile geomorphological tool, and can be used successfully either as a stand-alone method for surveying, or in conjunction with other methods.

KEYWORDS: dGPS; surveying systems; surface morphology; terrain; topography

Introduction

Global positioning systems (GPS) calculate the transmission time of messages received from GPS satellites to provide information on their location, accurate to ~ 15 m. Improvements in technology and a requirement for greater accuracy led to the development of differential global positioning systems (dGPS) during the late 1980s (McCoy, 2005). The greater accuracy is due to use of a base station and roving receiver. The base station is positioned at a location with known coordinates, and continually records its position according to the satellite network. Corrections based on the base station's internal calculations are applied to the position data selectively recorded by the roving receiver, increasing accuracy to the cm level. For standard dGPS equipment these corrections are applied during the data processing stage, but real time kinematic dGPS stations are able to apply corrections as locations are recorded, providing an immediately available quality check.

In geomorphology, dGPS has been used for surveying and mapping since the 1990s (McCoy, 2005). Initially, it was mainly employed to provide control points for other

surveying techniques. However, since the mid-1990s dGPS has been used as a surveying method in its own right (Baptista *et al.*, 2008), offering advantages in the speed and density of data collection (Harley *et al.*, 2011). The high accuracy of dGPS made it possible to conduct short-timescale studies of small morphological changes (Mitasova *et al.*, 2003; Pardo-Pascual *et al.*, 2005; Harley *et al.*, 2011).

Since the early 2000s, dGPS has been adopted as one of the most common surveying tools in geomorphology. It has been used extensively in a wide range of environments (Table 1). The simplest application of dGPS involves recording the location of features or points of geomorphological interest, for example, the highest points of dune crests (e.g. Mitasova *et al.*, 2005). For small features a single dGPS reading may be sufficient, but for larger features it is also possible to take a series of points. The form of these is dependent on the feature being measured, but may be linear such as shorelines (e.g. Rocha *et al.*, 2009), or situated around the external edges of a feature, e.g. a pond. The end purpose of this application is normally to create new maps, or enhance existing ones

with the addition of extra features. Where the surface morphology of an area rather than a feature is of interest, dGPS can be used to conduct topographic surveys of the entire

area. dGPS surveys are capable of recording data in three dimensions (latitude, longitude, and elevation), and may be used for 2- or 3-dimensional mapping.

Table 1. Examples of studies using dGPS in a range of environments.

Environment	Example publications
Dune field	Mitasova <i>et al.</i> , 2005; Navarro <i>et al.</i> , 2011
Littoral	Mills <i>et al.</i> , 2005; Suursaar <i>et al.</i> , 2005; Baptista <i>et al.</i> , 2008; Rocha <i>et al.</i> , 2009; Yang 2010; Bertoni and Sarti, 2011
Agricultural land	Saha, 2003; Zhang <i>et al.</i> , 2011
Salt marsh	Chassereau <i>et al.</i> , 2011
Forest	Hartvich and Mentlik, 2010
Floodplain	Casas <i>et al.</i> , 2006
Volcanic	Gomez <i>et al.</i> , 2009; Onderdonk <i>et al.</i> , 2011
Fluvial	Ujvari <i>et al.</i> , 2009
Landslide	Malet <i>et al.</i> , 2002; Corsini <i>et al.</i> , 2005; Glen <i>et al.</i> ; 2006 Baldo <i>et al.</i> , 2009
Glacial	Quincey <i>et al.</i> , 2009; Serrano <i>et al.</i> , 2010; Zhang <i>et al.</i> , 2010

dGPS survey data is commonly imported to geographic information systems (GIS), where linear points can be converted to lines and survey points to digital terrain models (DTMs). This use of the data makes it possible to compare collected data to other datasets, which may have been collected by dGPS or other methods, e.g. photogrammetry, light detection and ranging (LiDAR), or digitised maps. Repeat surveys of the same areas or features can be compared, making dGPS a powerful tool for measuring morphological changes over short, medium and long timescales (e.g. Malet *et al.*, 2002; Baptista *et al.*, 2008; Ujvari *et al.*, 2009; Bertoni and Sarti, 2011; Harley *et al.*, 2011). The change measured is dependent on the data acquired. Linear data can be used to track the movement of features, e.g. Bertoni and Sarti (2011) conducted repeat surveys of shore-normal beach profiles, and tracked the distance of retreat of features such as berms. If survey data is interpolated to create a DTM, changes in volume can be calculated using raster subtraction between surfaces (e.g. Baptista *et al.*, 2008; Harley *et al.*, 2011).

A further use of dGPS in geomorphology is to determine the accuracy of other survey methods. Accuracy is dependent on the

equipment set-up used (Table 2), but can be reduced to 1-3 cm by increasing point occupation times and processing results using data from the GNSS (Global Navigation Satellite System) network (Landau *et al.*, 2008). This compares favourably with other survey methods, being similar to terrestrial laser scanning (TLS) accuracy (Dornbusch, 2010), and more accurate than airborne LiDAR (highest achievable accuracy ~0.1-0.2 m, e.g. Sallenger *et al.*, 2003). Additionally, dGPS accuracy is not significantly reduced by environmental factors. However, some survey methods, such as LiDAR can gather data more rapidly.

dGPS points may be used to determine the error associated with other techniques due to their high accuracy (e.g. Baldo *et al.*, 2009; Harley *et al.*, 2011; Zhang *et al.*, 2011). After comparing surveys of a salt marsh conducted using dGPS and LiDAR Chassereau *et al.* (2011) found LiDAR to be sufficiently accurate over most of their study area, but recommended the use of dGPS where microtopography or small morphological changes were important. For the above reasons, dGPS may be used in conjunction with other survey methods, as well as functioning as a stand-alone survey technique.

Table 2. Accuracy and productivity of different methods of conducting a dGPS survey. Note that values given are indicative of some of the highest achievable rates on beach terrain. Surveys on more complex terrain are likely to take longer to complete.

¹ Baptista et al., 2008; ² Pardo-Pascual et al., 2005; ³ Harley et al., 2011; ⁴ Baptista et al., 2011.

Method	Foot	Foot	Wheel	Vehicle (1 antenna)	Vehicle (2 antennae)
Mode	Static	Continuous	Continuous	Continuous	Continuous
Accuracy (m)	< 0.05 ³	~ 0.05 ^{2,3}	< 0.1 ¹	< 0.1 ¹	< 0.05 ⁴
Productivity	< 0.05 ³	0.5 ³	10 ¹	100 ¹	100 ¹

Method

Survey design

Before conducting a survey using dGPS it is important to determine what will be surveyed, the density of points required, and the method of surveying. The time and methods available, and the aim of the survey, will dictate the maximum point density and area of coverage. It is important to note that in certain environments the time allocated to a survey may be reduced by the natural characteristics of the site, e.g. tides may limit the time available to survey beaches and estuaries.

When the survey objective is to generate a surface grid, there are three main sources of error to consider when designing the survey, all of which are inter-linked. These are:

- i) Terrain morphology: for any given survey duration, it will be possible to map a less complex surface with less error than a more complex surface, i.e. a low-gradient, geomorphologically mature surface may be reliably mapped with less points than mountainous terrain.
- ii) Survey point density: Increasing survey point density will decrease error. The higher the number of survey points, the greater the ratio of known elevations to estimated elevations in the surface grid generated.
- iii) Interpolation technique used: Many different interpolation methods are available with most commercial GIS. Each one will produce a slightly different surface from the same set of survey points. Error is minimised by selecting an interpolation method appropriate to the survey area and objectives.

Aguilar *et al.* (2005) found that the terrain morphology was the largest source of error, followed by survey point density, with the choice of interpolation method having the smallest effect on surface grid error.

If the area surveyed has little variation in terrain morphology capturing elevation data at a series of points on a regular grid will provide the best results when data is interpolated. However, as variations in surface morphology, and variations in point density affect the error associated with the final product some degree of survey bias may be beneficial, particularly if time is limited and terrain morphology is variable across the survey area. For example, Baptista *et al.* (2008), surveying at a coastal site, examined the effects of varying point density on the surface grid produced. They found that a high point density was required to capture variation on the topographically complex dune surfaces, while the relatively planar beach surface could be accurately mapped with a much lower point density.

To create a surface grid or DTM from point data interpolation or geostatistical methods are used to estimate the area's topography from the measured elevation data. Several different methods are available to do this, including: kriging, natural neighbour, inverse distance weighting, spline, and triangulated irregular network (TIN). Each of these methods makes different statistical assumptions about the data being used. The error associated with an interpolation method can be established by removing some measured points from the analysis and comparing the difference between the estimated and measured values for these locations. A benefit of using kriging is that this process is carried out automatically, and

surface error maps can be produced in addition to a DTM. While choice of interpolation method has a relatively minor effect on surface accuracy (Aguilar *et al.*, 2005), Baptista *et al.* (2008) found that TIN, spline, and kriging produced results with lower RMS than the other methods in a coastal environment. Additionally, kriging and TIN are exact interpolators, which ensures that where elevation has been measured, this value is used in the resultant surface. Most interpolation methods require an approximately rectangular grid of points which may constrain the area surveyed. Alternatively, points falling outside a rectangular grid may be removed during processing (Dornbusch, 2010), or additional steps can be taken to fit irregularly spaced data to a grid (Harley *et al.*, 2011).

A final consideration in designing a survey with the purpose of creating a surface grid or DTM is the 'edge effect'. Edge effects are erroneous results occurring around the boundary of the interpolated area due to a lack of surrounding data to inform estimations. This can be avoided by extending the surveyed area beyond the area of interest, and then clipping the resultant surface grid to the area of interest.

Survey method

The dGPS system consists of a static base station and a movable rover (Figure 1). The rover can be configured in several ways which provide different methodologies for surveying. The simplest configurations involve mounting the roving antenna on a tribrach or pole. These are then positioned over the point to be measured, levelled, and a point is recorded manually (static mode). The time taken to measure points is normally 2-3 minutes when using a tribrach and ~10 seconds when using the pole. In windy conditions points may take longer to measure as the antenna must be held horizontal to minimise errors.

Several authors have proposed and used adaptations to this basic configuration to increase productivity, in some cases at the expense of accuracy (Table 2). These adaptations do not require the operator to manually decide which points to record; instead the rover measures points continually at a set frequency. The most basic

adaptation is simply placing the rover antenna in a back-pack and setting it to take continual measurements as the operator walks across the area to be surveyed (Pardo-Pascual *et al.*, 2005; Dornbusch, 2010). A further enhancement involves mounting the antenna on a pole attached to a wheel, which can be rolled across the survey area (Buckley and Mills, 2000). The most advanced method mounts one or two antennae onto a rig attached to a moving vehicle (normally a quad-bike) which drives backwards and forwards across the survey area continually collecting data (Baptista *et al.*, 2008; Baptista *et al.*, 2011). Each of these methods has advantages and disadvantages, which are summarised in Table 3.

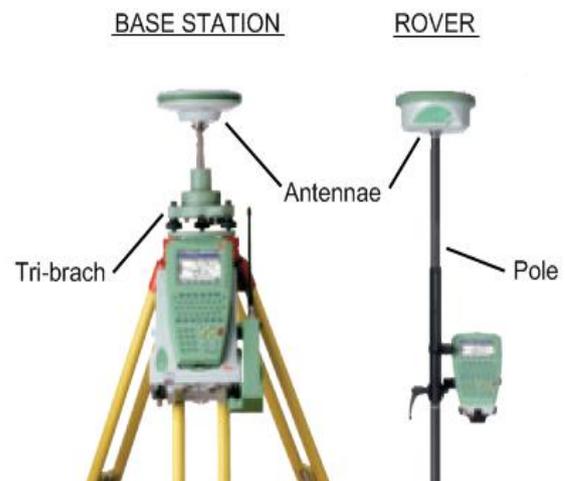


Figure 1. Leica dGPS 1200 system base station and rover with features referred to in the text annotated (adapted from Leica Geosystems, 2008).

Setting up the base station

Before points can be recorded the base station must be set up. The base station does not need to have a line of sight to the survey area, and depending on the dGPS system used, it may be up to a maximum distance of 2 – 10 km from the survey area. However, if surveying alone it is preferable to remain within view of the base station, or within easy access of it for security reasons.

The dGPS base must be level to obtain maximum accuracy. This is achieved by positioning the antenna on a tripod. This enables the base station to be set up on most terrain as the length of each leg can be

Table 3. Advantages and disadvantages of different dGPS survey methods.

Method	Advantages	Disadvantages
Foot (static)	Very high accuracy No modifications to equipment required Point density can be varied Suitable for all terrain	Very low productivity Accuracy and productivity may be reduced in high winds
Foot (continuous)	No modifications to equipment required Points taken automatically Suitable for all terrain	Error introduced as antenna not always vertical (especially on slopes)
Wheel	No expensive modifications required Points taken automatically Suitable for most terrain	Error introduced as antenna not always vertical Wheel may sink into very soft sediments
Vehicle (1 antenna)	Very high productivity Points taken automatically	Expensive modifications required Unsuitable for some terrains (e.g. fragile soils and plants, steep slopes) Vehicle may sink into soft sediments Requires vehicle access Error introduced as antenna not always vertical (especially on uneven terrain)
Vehicle (2 antennae)	Very high productivity High accuracy	Expensive modifications required Unsuitable for some terrains (e.g. fragile soils and plants, steep slopes) Vehicle may sink into soft sediments Requires vehicle access Error introduced as antenna not always vertical (especially on uneven terrain)

individually altered. While relatively stable, in windy conditions it is best to set the base station up on soil rather than concrete or tarmac for additional stability.

If repeat surveys of the same site are required it is preferable to 'monument' the position of the base station using a marker, e.g. paint, nails, screws, etc., to ensure that the same position can be returned to. If this is not possible control points at easily identifiable locations can be taken using the rover. These points can then be resurveyed on each subsequent visit, and used to correct for any differences in the position of the base station between surveys (McCoy, 2005).

Full instructions for setting up the base station are included in manuals supplied with the equipment, and are dependent on the make and model of the system. Leica and Trimble dGPS systems are commonly used by geomorphologists and their manuals are base station. During this time the rover can be set up. Whether using a vehicle, a pole or

available via their websites. The basic procedure common to all dGPS systems involves setting up the tripod on which the other components are mounted. These include the antenna, batteries, the dGPS unit, and the radio (optional) (see Figure 1). The antenna must be mounted on a tribrach to ensure that it is level. The radio is only required if the RTK function is desired.

Once the physical components of the dGPS system are set up, the screen on the dGPS unit is used to start a survey. The exact process depends on the make of the system and is available in the equipment manuals. Generally, this involves naming the survey, entering information about the required equipment configuration, and then setting the dGPS unit to acquire its location. This process takes ~20 minutes, during which time the base station continually receives location data from satellites, which are averaged to give an accurate position for the

a tribrach, the rover set-up will include an antenna, batteries, dGPS unit and a radio antenna if used.

Survey

To conduct a survey the rover's dGPS unit is turned on, and the survey name is entered. If continuous measurements are required the desired frequency of measurement is entered, and the operator simply moves across the survey area while points are automatically recorded. If static measurements are required each point must be individually 'occupied'. This involves pressing a button the dGPS rover unit and waiting for the point to be recorded before moving to the next point and repeating this procedure. Maximum accuracy is obtained by using the bull's eye spirit level connected to the antenna pole or tribrach to ensure the instrument is level while the point is occupied. Occupation time depends on whether the rover is mounted on a pole or tribrach, and the make of the system, taking on average 2-10 seconds, and between 30-180 seconds, respectively.

The path taken by the surveyor across the area of interest will be dependent on what the aims of the study are, what features are to be included in the survey, how much time is available, and the degree of coverage needed. To maximise efficiency Baptista *et al.* (2008) suggest planning the survey as a series of transects designed to capture areas of maximum variation. In their beach study most variation was cross-shore rather than long-shore, so transects could be widely spaced along the shore provided that points were densely spaced along each transect.

When sufficient points have been recorded the survey is stopped using the dGPS unit. Data is saved automatically onto CF cards in the unit.

Repeat surveys

For some studies it may be necessary to resurvey the same locations to measure changes in elevation, e.g. ground bulging in volcanic areas (e.g. Nishi *et al.*, 1999; Lagios *et al.*, 2005). Resurveying the same locations can be achieved by installing benchmarks or using the 'stake-out' function of the dGPS system. To use this function the

original survey coordinates must remain stored on the dGPS CF cards. The exact method for using stake-out depends on the make of the dGPS system, but involves the location of past survey points being displayed on a mini-map on the roving dGPS unit's screen. The map displays the position of the location to be resurveyed and the position of the roving receiver, making it possible to navigate to the point(s) of interest.

Processing

Discussion of the full data processing method is beyond the scope of an article on the direct acquisition of raw data. However, a few simple steps are required to transfer the survey data to a computer in a usable format. These involve copying the data from the CF cards to a computer's hard-drive. Once on the hard-drive the data can be processed using software provided by the manufacturer. The result is a table which includes information on each point acquired, such as the time recorded, coordinates, elevation, and quality. The table can be copied into a text file or a spreadsheet, ready for further processing.

Error

The accuracy associated with using different methods to conduct dGPS surveys has been discussed above, but there are some additional sources of error which should be considered. Some of these cannot be directly controlled, such as the number of satellites the dGPS can receive position data from. A minimum of four satellites are required to provide a vertical and horizontal dGPS position (Leica Geosystems, 1999), and a minimum of six satellites are required for RTK-dGPS (Pardo-Pascual *et al.*, 2005). The more satellites the dGPS can receive data from, the higher the accuracy of the position obtained. This is dependent on the position of satellites above the earth, and varies as they move through their orbits. If surveying in a built-up area, below cliffs or in a forest, the accuracy will be reduced as trees, buildings and cliffs will block some of the satellite signals (McCoy, 2005).

Preventable sources of error have already been mentioned and include deviation of the rover antenna from the vertical and failure to use control points for repeat surveys. Human

sources of error, and best practice to avoid them, are dependent on the survey method

and are summarised in Table 4.

Table 4. Sources of error associated with each dGPS method and suggestions to minimise them.
¹Baptista *et al.*, 2008; ²Dornbusch, 2010; ³Harley *et al.*, 2011.

Method	Sources of error	Best practice to prevent errors
Foot (static)	Pole or tribrach may sink in soft ground	Ensure pole/tribrach is held with point resting at surface level OR Place a mark a set distance from the base of the pole and ensure pole is always sunk into ground to level of mark
Foot (continuous)	Antenna not physically connected to ground ¹ Variation in height due to walking ² Change in antenna height when walking up slope or adjusting backpack ²	N/A Will cancel out over course of survey provided measurements not synchronous with steps ² Set up a series of control points and check these regularly throughout survey ²
Wheel	Wheel will change point of contact on beach when on slope ²	Use smallest wheel size possible ²
Vehicle (1 antenna)	Vehicle sinking into ground Mast with antenna deviates from vertical due to roll and pitch (particularly on slopes) ¹	N/A Restrict use to relatively flat surfaces OR use 2-antenna configuration ¹ OR on smooth surfaces error may be within accuracy of dGPS ¹
Vehicle (2 antennae)	Vehicle sinking into ground Mast with antenna deviates from vertical due to pitch ¹	N/A A 3rd antenna can be incorporated ¹

Integration of dGPS with other methods

Alternatives to dGPS for acquiring elevation data include total station, LiDAR, TLS, and stereophotogrammetry. The advantages and disadvantages of these techniques are detailed in Table 5. The most appropriate method for any survey will be reliant on a number of factors including: time available; accuracy required; survey area; cost; and, terrain.

The techniques listed in Table 5 need not be used in isolation. The variation in accuracy and productivity between techniques has led to a number of studies which use multiple techniques to obtain datasets covering a survey site at a range of scales and resolutions. This makes it possible to

investigate changes at small, medium and large scales, spatially and temporally.

Some of the techniques listed in Table 5 are relatively modern. To investigate geomorphic change over longer time periods than the last couple of decades, data obtained using older surveying techniques, or extracted from historical aerial photography and maps, may be combined with modern survey data to provide a more continuous record of change (e.g. Yang *et al.*, 2010; Harley *et al.*, 2011). Additionally, dGPS may be used to ground-truth the accuracy of lower accuracy, higher productivity techniques like airborne LiDAR (e.g. Saha, 2003; Baldo *et al.*, 2009; Hofle *et al.*, 2009; Chassereau *et al.*, 2011; Zhang *et al.*, 2011).

Table 5. Advantages and disadvantages of techniques for acquiring elevation data.
¹Pardo-Pascual *et al.*, 2005; ²Harley *et al.*, 2011; ³Dornbusch, 2010; ⁴Gregory and Goudie, 2011.

Method	Advantage	Disadvantage
dGPS	High accuracy Range of methods have been developed to suit different surveying requirements Line-of-sight not required ³	High cost ³ Some methods have low productivity Lock on 6+ satellites required ¹
LiDAR	High productivity ^{1,4} Can be used during the night ³ Airborne LiDAR can survey areas that are difficult to access ³ Not affected by vegetation cover ⁴	Very high cost ¹ Resolution may be insufficient to measure small changes ² Systematic errors on some terrains ³
TLS	High accuracy ³	Unable to capture all aspects of complex topographies (depending on equipment positioning) ³
Total station	Low cost Accurate	Line-of-sight required ³ Low productivity ^{1,3} Accuracy decreases with distance from base ³
Stereo-photogrammetry	High productivity Once set up, no operator required ² Continuous information can be captured ³	Low resolution ³ Equipment must be left in position for long periods of time (depending on survey) and may be vandalised or damaged Does not work in fog, mist etc ³

Once elevation data has been collected using dGPS, it is normally viewed and processed further in GIS. At its simplest, this use of the data makes it possible to view the survey area or features on a rotatable digital map (e.g. Lo Curzio and Magliulo, 2010), and to use interpolation methods to create DTMs (e.g. Casas *et al.*, 2006; Baptista *et al.*, 2008). Importing dGPS data to GIS also opens up other possibilities, which include combining the current survey data with data obtained on previous surveys, or by other researchers. It is also possible to use GIS functions to measure changes in volumes or the position of features between two or more surveys of the same area/feature conducted at different times (e.g. Bertoni *et al.*, 2011).

Conclusion

dGPS is a highly accurate method of directly acquiring elevation data. Considerable improvements in the capabilities of dGPS systems have been made since their first use in geomorphology in the 1990s. These improvements include technological advances which have led to more accurate and portable systems, as well as adaptations

of the standard dGPS equipment to enhance productivity. dGPS is not the only method for acquiring elevation data, but it is one of the most accurate and simplest to use methods currently available. In any study where acquisition of elevation data is required, it is important to consider the most appropriate technique, as each has different advantages and disadvantages. The selection of technique will be dependent on a number of factors, including the area to be surveyed, time and finances available, and the desired balance between productivity and accuracy. However it is not necessary to consider the selection of a surveying technique as an 'either/or' decision. Numerous studies have shown the benefits of using two or more survey methods to provide a greater understanding of geomorphological features and processes than would have been achievable using only one method. dGPS is an excellent tool when used in isolation for mapping geological features or areas of interest, and assessing changes in position or volume. Equally, it can make a valuable contribution to studies employing other techniques, either to contribute to the collection of survey data or to ground-truth the accuracy of other methods.

References

- Aguilar FJ, Aguera F, Agullar MA, Carvajai F. 2005. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogrammetric Engineering and Remote Sensing* **71**: 805-816.
- Baldo M, Biococchi C, Chiocchini U, Giordan U, Lollino G. 2009. LiDAR monitoring of mass wasting processes: The Radicofani landslide, Province of Siena, central Italy. *Geomorphology* **105**: 193-201.
- Baptista P, Bastos L, Bernardes C, Cunha T, Dias J. 2008. Monitoring sandy shores morphologies by DGPS – A practical tool to generate digital elevation models. *Journal of Coastal Research* **24**: 1516-1528.
- Bertoni G, Sarti D. 2011. On the profile evolution of three artificial pebble beaches at Marina di Pisa, Italy. *Geomorphology* **130**: 244-254.
- Buckley S, Mills J. 2000. GPS and the wheel – how integrating the world's greatest inventions is helping to monitor coastal erosion. *Surveying World* **9**: 41.
- Casas A, Benito G, Thorndycraft VR, Rico M. 2006. The topographic data source of digital terrain models as a key element in the accuracy of hydraulic flood modelling. *Earth Surface Processes and Landforms* **31**: 444-456.
- Chassereau JE, Bell JM, Torres R. 2011. A comparison of GPS and LiDAR salt marsh DEMs. *Earth Surface Processes and Landforms* **36**: 1770-1775.
- Corsini A, Pasuto A, Soldati M, Zannoni A. 2005. Field monitoring of the Corvara landslide (Dolomites, Italy) and its relevance for hazard assessment. *Geomorphology* **66**: 149-165.
- Dornbusch U. 2010. Ground survey methods for mixed sand and gravel beaches in intertidal environments: A comparison. *Journal of Coastal Research* **26**: 451-464.
- Glenn NF, Streutker DR, Chadwick DJ, Thackray GD, Dorsch SJ. 2006. Analysis of LiDAR-derived topographic information for characterising and differentiating landslide morphology and activity. *Geomorphology* **73**: 131-148.
- Gomez C, Lavigne F, Hadmoko DS, Lespinasse N, Wassmer P. 2009. Block-and-ash flow deposition: A conceptual model from a GPR survey on pyroclastic-flow deposits at Merapi Volcano, Indonesia. *Geomorphology* **110**: 118-127.
- Harley MD, Turner IL, Short AD, Ranasinghe R. 2011. Assessment and integration of conventional, RTK-GPS and image-derived beach survey methods for daily to decadal coastal monitoring. *Coastal Engineering* **58**: 194-205.
- Hartvich F, Mentlik P. 2010. Slope development reconstruction at two sites in the Bohemian Forest Mountains. *Earth Surface Processes and Landforms* **35**: 373-389.
- Hofle B, Vetter M, Pfeifer N, Mandlbürger G, Stotter J. 2009. Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms* **34**: 1635-1649.
- Lagios E, Sakkas V, Parcharidis I, Dietrich V. 2005. Ground deformations of Nisyros Volcano (Greece) for the period 1995-2002: results from DInSAR and DGPS observations. *Bulletin of Volcanology* **68**: 201-214.
- Landau H, Chen X, Klose S, Leandro R, Vollath U. 2008. Trimble's RTK and dGPS solutions in comparison with precise point positioning. *International Association of Geodesy Symposia* **133**: 709-718.
- Leica Geosystems. 1999. Introduction to GPS (Global Positioning System) version 1.0. *GPS Basics*. Leica Geosystems: Heerbrugg, Switzerland.
- Leica Geosystems. 2008. *Leica GPS1200+ Series: High Performance GNSS System*. Leica Geosystems: Heerbrugg, Switzerland.
- Lo Curzio S, Magliulo P. 2010. Soil erosion assessment using geomorphological remote sensing techniques: An example from southern Italy. *Earth Surface Processes and Landforms* **35**: 262-271.
- Malet JP, Maquaire O, Calais E. 2002. The use of global positioning system techniques for the continuous monitoring of landslides: Application to the Super-Sauze earthflow (Alpes-de-Haute-Provence, France). *Geomorphology* **43**: 33-54.
- McCoy RM. 2005. *Field Methods in Remote Sensing*. The Guilford Press: New York.

- Mills JP, Buckley SJ, Mitchell HL, Clarke PJ, Edwards SJ. 2005. A geomatics data integration technique for coastal change monitoring. *Earth Surface Processes and Landforms* **30**: 651-664.
- Mitasova H, Drake TG, Harmon RS, Bernstein D. 2004. Quantifying rapid changes in coastal topography, using modern mapping techniques and GIS. *Environmental and Engineering Geoscience* **10**: 1-11.
- Mitasova H, Overton M, Harmon RS. 2005. Geospatial analysis of a coastal sand dune field evolution: Jockey's Ridge, North Carolina. *Geomorphology* **72**: 204-221.
- Navarro M, Munoz-Perez JJ, Roman-Sierra J, Tsoar H, Rodriguez I, Gomez-Pina G. 2011. Assessment of highly active dune mobility in the medium, short and very short term. *Geomorphology* **129**: 14-28.
- Nishi K, Ono H, Mori H. 1999. Global position system measurements of ground deformation caused by magma intrusion and lava discharge: the 1990-1995 eruption at Unzendake Volcano, Kyushu, Japan. *Journal of Volcanology and Geothermal Research* **89**: 23-34.
- Onderdonk N, Mazzini A, Shafer L, Svensen H. 2011. Controls on the geomorphic expression and evolution of gryphons, pools, and caldera features at hydrothermal seeps in the Salton Sea geothermal field, southern California. *Geomorphology* **130**: 327-342.
- Pardo-Pascual JE, García-Asenjo L, Palomar-Vázquez J, Garrigues-Talens P. 2005. New methods and tools to analyze beach-dune system evolution using a real-time kinematic global positioning system and geographic information systems. *Journal of Coastal Research* **SI 49**: 34-39.
- Quincey DJ, Copland L, Mayer C, Bishop M, Luckman A, Belo M. 2009. Ice velocity and climate variations for Baltoro Glacier, Pakistan. *Journal of Glaciology* **55**: 1061-1071.
- Rocha CP, Araujo TCM, Mendonca FJB. 2009. Methodology for location of shorelines using 3D-GPS positioning: A case study at Sauacui Beach, northeast Brazil. *Journal of Coastal Research* **25**: 1052-1058.
- Saha SK. 2003. Water and wind induced soil erosion assessment and measurement using remote sensing and GIS. In Proceedings of the Training Workshop on Satellite Remote Sensing and GIS Applications in Agricultural Meteorology. World Meteorological Organisation: Dehra Dun, India. 315-330.
- Sallenger A, Krabill WB, Swift RN, Brock JC, List J, Hansen M, Holman RA, Manizade S, Sontag J, Meredith A, Morgan K, Yunkel JK, Frederick EB, Stockdon H. 2003. Evaluation of airborne topographic LIDAR for quantifying beach changes. *Journal of Coastal Research* **19**: 125-133.
- Serrano E, de Sanjose JJ, Gonzalez-Trueba JJ. 2010. Rock glacier dynamics in marginal periglacial environments. *Earth Surface Processes and Landforms* **35**: 1302-1314.
- Smith MJ, Pain CF. 2011. Geomorphological mapping. In *The SAGE Handbook of Geomorphology*, Gregory KJ, Goudie AS (eds.), pp. 142-153. SAGE Publications Ltd: London.
- Suursaar A, Tonisson H, Kullas T, Orviku K, Kont A, Rivis R, Otsmann M. 2005. A study of hydrodynamic and coastal geomorphic processes in Kudema Bay, the Baltic Sea. In Proceedings of the International Conference on Computer Modelling and Experimental Measurement of Seas and Coastal Regions. WIT Press: Algarve, Portugal. 187-196.
- Ujvari G, Mentés G, Banyai L, Kraft J, Gyimothy A, Kovacs J. 2009. Evolution of a bank failure along the River Danube at Dunaszekcső, Hungary. *Geomorphology* **109**: 197-209.
- Yang JY, Choi CU, Jeon SW, Hong SY, Kim YS. 2010. An analysis of coastal topography and land cover changes at Haeundae Beach, South Korea, by merging GIS and remote sensing data. *Acta Astronautica* **67**: 1280-1288.
- Zhang C-L, Yang S, Pan X-H, Zhang J-Q. 2011. Estimation of farmland soil wind erosion using RTK GPS measurements and the ¹³⁷Cs technique: A case study in Kangbao county, Hebei province, northern China. *Soil and Tillage Research* **112**: 140-148.
- Zhang Y, Fujita K, Liu SY, Liu QA, Wang X. 2010. Multi-decadal ice-velocity and elevation changes of a monsoonal maritime glacier: Hailuoguo Glacier, China. *Journal of Glaciology* **56**: 65-74.

Direct Acquisition of Data: Airborne laser scanning

Michal Gallay¹

¹ Institute of Geography, Faculty of Science, Pavol Jozef Šafárik University in Košice, Jesenná 5, 040 01, Slovakia, (michal.gallay@upjs.sk)



ABSTRACT: Airborne laser scanning (ALS) is an active remote sensing method capable of highly detailed land surface mapping even in conditions where other methods meet their limitations. It records distance measurements based on laser light detection and ranging (i.e. LiDAR) and tracks the position of each measurement using a global navigation satellite system (e.g. a GPS). The raw data acquired create a massive cloud of relatively uniformly distributed point heights, typically of 1-3 meters spacing accurate to within centimetres or decimetres. The level of accuracy is dependent on a number of factors (such as the flight altitude, slope angle, scanner parameters, inertial system and GPS signal quality); the vertical accuracy is generally two to three times higher than the horizontal accuracy. Earlier ALS systems relied on discrete recording of echoes; however, recent developments have enabled full-waveform recordings providing improved sampling of land cover and elevation. The main application of ALS is in producing detailed digital models of surfaces sampled by the laser beam, which can be reflected (returned) from several surface levels and recorded. Classification of the points can be achieved by filtering the point cloud. The points representing the bare ground (terrain) are the most relevant in geomorphology. ALS data provide an unprecedented level of detail and they are especially useful for studying earth surface processes and landforms under the vegetation canopy.

KEYWORDS: airborne LiDAR, high resolution, digital surface model, digital terrain model

Introduction

Airborne laser scanning is an active remote sensing technique that uses its own source of electromagnetic energy to ascertain the characteristics of the land surface. Therefore, the measurement is independent of external illumination by sunlight. The technology used is referred to as *light detection and ranging* (LiDAR). The term *LiDAR* or *lidar* also applies to the laser scanning device, in the same way that radar describes both the technology and the instrument. Topographic mapping based on LiDAR technology is referred to as *airborne laser scanning* (ALS), *laser altimetry* or *topographic LiDAR*. ALS will be used for consistency throughout the chapter.

The LiDAR technology uses the laser light to measure the distance between the laser scanner and the illuminated object. The theoretical framework for LASER (Light Amplification by Stimulated Emission of

Radiation) technology was developed by Albert Einstein in the late 1910s but the history of measuring distances with lasers dates back to the 1960s. Later improvements in the technology and computer hardware revolutionised the study of the earth surface. This has been achieved, in particular, by the ability to sample a higher density of terrain heights with greater accuracy than before, even through vegetation cover (Figure 1). As a result, a digital terrain model (DTM) depicting an unprecedented level of detail can now be generated (see Section 2.3.1 of *Geomorphological Techniques: Creation of DEMs from survey data*). The DTM can subsequently be used to analyse and model earth surface processes by the means of geomorphometry (Pike et al. 2009).

Pioneer works summarising the principles of ALS and its applications include Kraus and Pfeifer (1998), Baltsavias (1999), and Wehr and Lohr (1999). At that time, the costs

related to ALS were higher in comparison with other remote sensing methods. Progress in ALS technology in the last few years increased the cost-effectiveness of the method, making it more accessible to the wider research community and commercial enterprise. For these reasons, ALS has become a highly significant geomorphological technique. However, there are important aspects to consider in the collection, processing and application of ALS which are addressed this chapter.

ALS system

ALS integrates the use of various hardware and software components in a system. If a lidar is used as an isolated device it would not be possible to acquire geo-referenced laser measurements useful for geomorphological applications. Wehr (2009) differentiates the components of an ALS system into two groups: an airborne segment and a ground segment. The airborne segment comprises: (i) an airborne platform, (ii) a laser scanner, and (iii) a position and orientation system.

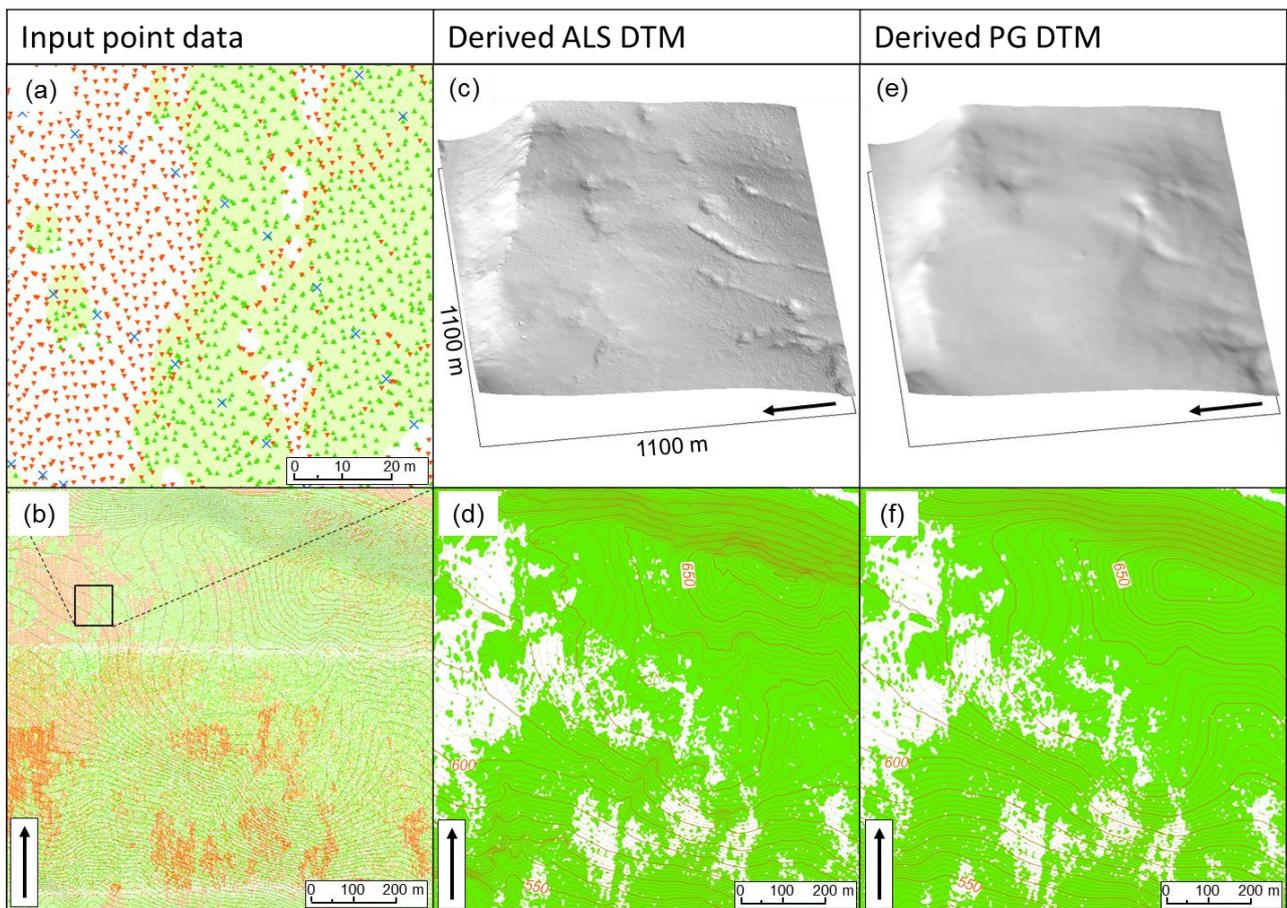


Figure 1: Terrain representation of a forested area in the Slovak Karst generated from point measurements (a, b) acquired with ALS and photogrammetry (PG). Derived DTMs (2 metres cell size) are visualized as 3-D surfaces (c, e) and 2-D contour lines (d, f). While (a), a close-up area taken from (b), shows all points are relatively uniformly distributed (37 points per 100 metres sq. on average), the spatial distribution of ALS ground points (red triangles) is heterogeneous (6 points per 100 metres sq. on average). Many more ALS ground points exist within the open land than in the forest, displayed as green areas in (a), (b), (d), and (f), where points classified as vegetation canopy (green inverted triangles) were mainly acquired. Yet still, the ALS ground points provide a more realistic terrain representation than the photogrammetric points (blue crosses in (a)) for which terrain altitude cannot be directly measured and it is usually estimated during the data capture. Especially under the forest canopy, shaded 3-D surface (c) and contours (d) clearly reveal a higher level of detail (ravines, dolines, or steeper slopes) than in (e) and (f). The area (b) contains 4,663 PG points and 463,053 ALS points of which 69,311 were classified as ground points. All ALS points would be stored in a 60 megabyte ESRI shapefile (x,y,z, class number).

The ground segment comprises: (i) Global Navigation Satellite System (GNSS) reference stations and (ii) processing hardware and software for synchronization and registration which is carried out off-line. Figure 2 shows the geometry and relationship between the ALS system components.

Typical ALS systems operate from fixed-wing airplanes or helicopters with flying heights H of 100-4000 metres above the ground. During the flight, the laser scanner samples line-of-sight slant ranges to the objects, R , referenced to the scanner coordinate system, the laser beam incidence angle θ , and the intensity of the reflected laser energy. The maximal value of θ defines the scanner field of view (FOV) angle. The travelling path of a laser beam within a swath is directed by a rotating or oscillating mirror and the aircraft forward motion. Hence, the footprint migrates on the earth surface creating a semi-regular pattern.

Flying height H and FOV control the width of the scanned surface, i.e. the swath width. The laser illuminates the scanned surface with a circular or elliptical footprint. Its shape depends on θ and laser beam divergence. The emitted laser energy has a very narrow beam divergence angles therefore the resulting footprints are small, typically 0.3 – 1 m at flying altitudes of about 1000 m (Wehr, 2009). The size of the footprint determines the penetration level of the pulse in vegetated areas and the level of detectable detail that can be recorded.

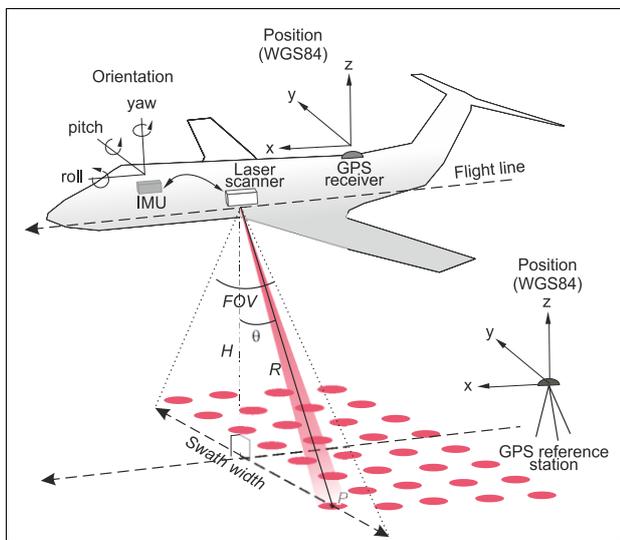


Figure 2. The geometry and components of an ALS system. Annotations are described in the main text.

The orientation of the scanner is monitored and stored by the inertial measuring unit (IMU) while a GNSS receiver records the geographic position of the scanner. For high laser pulse repetition frequency (Table 1) during scanning the position and orientation information must be measured at sufficiently high frequency to characterize the change in position and orientation of the laser scanner. IMU measurement frequencies of at least 100 Hz are usually needed for effective recording of aircraft orientation because it varies much more rapidly than position. A measurement frequency of 2 Hz or higher would enable the position of the scanner to be determined reliably (Devereux and Amable 2009). The laser scanner, IMU and GNSS sample data independently. At the same time, on-ground stable GNSS reference stations measure their own position for later post-processing of differential corrections of the airborne platform. The U.S. Global Positioning System (GPS) is the most widely used GNSS for positioning. Further information on GPS is provided in Section 2.1.3: *dGPS* of Geomorphological Techniques.

Integration of the differentially corrected GNSS and IMU data with a geometric model of the sensor scanning mechanism enables a vector to be calculated for each laser pulse with centimetre to decimetre total accuracy. Initial coordinates of the vector are referenced to WGS84 ellipsoid in case the GPS is used for positioning. In the final stage, the WGS84 coordinates are transformed into a local coordinate system such as the British National Grid. Even though the data are georeferenced instantaneously during the scanning process, they require further corrections and off-line processing. The accuracy of real-time data collection is increased in the process of registration and calibration in which ground control points located within the ALS swath are also used.

The data acquired during the mission provide a massive cloud of densely distributed point measurements (i.e. point cloud) containing altitude of multiple surfaces illuminated by the laser footprint. Sampling densities usually vary between 0.1 – 20 points per metre sq. The final post-processed data can be used to generate high-resolution digital elevation models (DEMs). The two most widely applicable types of DEMs are: (i) digital model of landscape canopy surface (DSM)

representing the elevation of the top parts of vegetation and buildings, and (ii) digital terrain model (DTM) which represents elevations of a bare ground.

Physical principles

The main principle of ALS is based on the theory that a proportion of laser energy will be reflected back to the scanner and time of flight information from the reflected laser energy may be recorded and used, together with instrument optical geometric information, to determine target object position in three-dimensional space. The laser light has constant wavelength and very narrow bandwidth (2-5 nanometres). In comparison with other light sources, it is emitted in a single direction in total phase correspondence and amplitude. Therefore, the laser light also has a very small beam divergence angle. Current ALS systems emit near-infrared light with typical wavelengths of

1064 nanometres. Bathymetric systems generate laser pulses at two wavelengths: an infrared (1064 nanometres) and a green one (532 nanometres). There have been two main physical approaches developed for measuring the range with a laser based on either *pulse* or *continuous* wave (CW) modulation of the laser light transmission (Baltsavias, 1999, Wehr and Lohr, 1999). For practical reasons, pulse laser scanning is preferred for topographic mapping.

With pulse modulation the transmitter generates energy pulses with a certain duration time (typically about 10 nanoseconds) between each pulse when it leaves the transmitter and is scattered back from an object and is detected by the laser receiver. The distance between the ranging unit and the object surface R is given by

$$R = c.t/2$$

Table 1. Technical specifications of selected currently available ALS systems

Manufacturer	Leica Geosystems	Optech	Airborne Hydrography AB	Riegl	TopoSys (Trimble since 2012)
Scanner type	ALS50-II	ALTM Gemini	Hawk Eye II	Riegl VQ-480	Falcon II
Weight of the laser scanner [kg]	36.1	23	95	13	41
Weight of the laser system [kg]	93.4	76	160	N/A	95
Wavelength [nm]	1,064	1,064	532 (bathy); 1,064 (topo)	1,550	1,560
Field of view [deg.]	75	50	50	60	14
Scanning method	oscillating mirror	oscillating mirror	2-axis gimbal mirror	rotating multi-facet mirror	fibre based
Intensity measurement [bit]	8	12	8 (bathy)	16	12
Beam divergence [mrad]	0.22	0.25 and 0.8	2-12	0.3	1
Pulse repetition frequency [kHz]	20 – 150	167	4 (bathy); 64 (topo)	50-200	83
Scanning rate [Hz]	≤ 90	≤ 70	≤ 13	≤ 100	N/A
Operating altitude (min./typical/max.) [m]	200/2500/7,000	150/2000/4,000	250/500	10/450/800	30/900/1,600
Number of echoes recorded [count]	≤ 4	≤ 4	full waveform (bathy); ≤ 4 (topo)	full waveform	≤ 9
Pulse duration [ns]	4-9	7	5	5	5
Vertical accuracy at 1σ [cm]	≤ 10	5 - 35	≤ 25 (bathy); ≤ 15 (topo)	≤ 15	≤ 15
Horizontal accuracy at 1σ [cm]	≤ 24	3 - 72	≤ 250 (bathy); ≤ 50 (topo)	≤ 10	≤ 10
Reference of application	Kereszturi et al. (2012)	Hladik and Alber (2012)	Chust et al. (2010)	Razak et al. (2011)	Morsdorf et al. (2008)

where t is the traveling time of a laser pulse and c is the speed of light (299,792,458 m.s⁻¹ in vacuum). The pulse duration determines the range resolution (Table 1). For example, when the pulse duration is 10 nanoseconds two objects must be more than 1.5 metre apart in slant range to be identified as separate targets. The range resolution can be significantly improved by quantizing the return pulse with a high sampling rate which is possible with full-waveform lidars.

Multiple-surface detection

An important benefit in ALS survey is collection of the altitude of several surface levels. Emitted laser light of the same pulse can be backscattered from multiple objects (e.g. tree leaves, roofs, bushes, ground) thus giving multiple echoes (Figure 3). The number of the recorded echoes (returns) depends on the amount of the laser energy which penetrated down through the ground and the technical parameters of the laser scanner. Early ALS systems only recorded the range to the first reflecting object or the first return. Later technological developments which increased storage capacity allowed for recording more returns, typically four (Table 1). This provides a means of distinguishing between the vegetation canopy and terrain surface. In general, the first returns are used for generating a DSM while a DTM is created from the last returns. However, the last returns can also represent off-terrain objects that are impermeable for the laser beam and filtering is required to separate these from the terrain heights. Subtraction of the DTM from

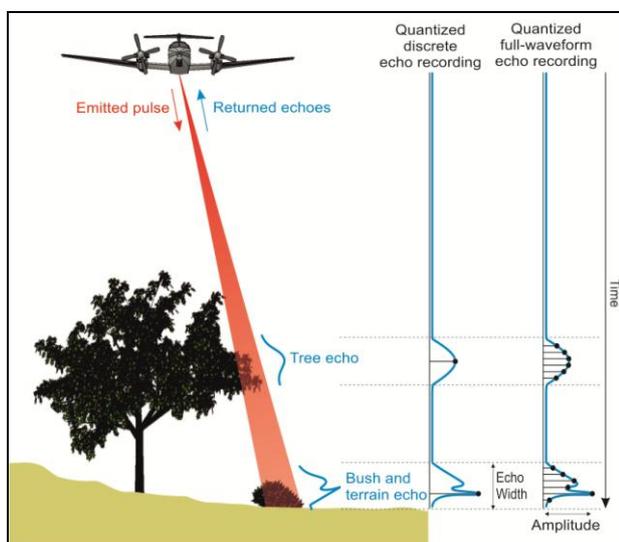


Figure 3. Discrete and full-waveform recording of multiple surfaces with ALS.

the DSM models yields a digital model of differences which is useful for parameterizing the vegetation or buildings (e.g. Waser et al., 2008). Further information can be found in *Section 2.3.2: DEMs of Difference of Geomorphological Techniques*.

The scanners which record several returned echoes are called discrete laser scanners. They monitor the return pulse recording time of flight information when an energy threshold is crossed. The threshold crossing depends on the amount of reflected energy received by the laser scanner. The amount of energy is influenced by the transmitted pulse energy, object reflectivity, illuminated area and atmospheric conditions. Recent developments in full waveform recording enable quantization of the transient of the return pulse with high sampling rates of about 1 gigahertz. Therefore different elevations can be resolved within the transient of one return pulse (Figure 3). This improves mapping in vegetated areas, especially in dense forest. Improved DTM generation is mainly based on the removal of low off-terrain objects (e.g. bushes) that can hardly be identified in 3-D point clouds acquired with discrete scanning.

Currently available ALS systems

The current ALS systems are classified as small footprint systems (Devereux and Amable, 2009). Standard commercial ALS systems employ discrete recording (Table 1). The most modern systems record the entire waveform, thus they are termed full waveform ALS systems (Mallet and Bretar, 2009, Lemmens 2009). Typical entry costs of an ALS system are in the order of £70,000 for hardware and up to £30,000 for the associated data gathering, handling and processing software (Heritage and Large, 2009). Table 1 lists parameters of several common ALS systems as stated by their manufacturers. The choice of a particular system is specific to the aims of the project concerned. For example, achieving a large spatial coverage is possible with systems flying at higher altitudes enabled by higher power scanner output and wider FOV (e.g. the Optech ALTM Gemini or Leica ALS50-II). Bathymetric ALS systems (e.g. the Hawkeye II) are suitable for simultaneous land and shallow water mapping up to 50 metres depth depending on water turbidity and sea bed

reflectance with a horizontal accuracy of 2.5 m and a vertical accuracy of 0.25 m (Milan and Heritage, 2012). The weight of an ALS system is an important factor if the platform payload is limited. This is the case of unmanned aerial vehicles (UAV) for which, for example, the Riegl VQ-480 scanner is suitable. The scanning mechanism can also be an important decision factor. Mirror-device scanners can cause a deflection error due to mechanics, acceleration and wear-out; therefore requiring regular calibration (Schnadt and Katzenbeisser, 2004). The Falcon II scanner by TopoSys uses stiff mounting of fibre optics on a rigid plate which ensures that the factory calibration remains valid over several years of operation.

Planning an ALS mission

There are often specific requirements in geomorphological research for which ALS data properties have to be tailored. Hence, it is useful to understand some of the important aspects of planning an ALS mission (Wehr 2009). Normally the customer defines the surveying area, the density of laser points on the ground and the accuracy. This determines the selection of the appropriate ALS system including the carrier platform, but the customer generally has limited options in platform selection. It depends on the contracted company and its equipment. Helicopters are used for lower altitude missions if high laser point densities are necessary and detailed surveys of small areas are essential (Wehr, 2009). Airplanes are preferred for surveying extensive areas at higher altitudes. For smaller areas and in rapid response situations (e.g. flooding, earthquake) ALS can be effectively undertaken with UAVs (Eck and Imbach, 2011). The UAV payload is relatively low but the development progresses in minimizing the weight of ALS systems and the UAV components.

Achieving a high spatial density of point measurements is one of the primary goals of ALS survey. This ensures that a dense set of terrain points will be obtained in the subsequent filtering of the off-terrain points. The point density defines the spatial resolution of the survey and it can be controlled by specifying flying height, field of view, speed of the aircraft (typically about 180 kilometres per hour), pulse repetition

frequency, scanning rate and footprint size. Table 1 reports the parameters for selected ALS systems.

Undertaking a successful mission requires selection of locations and areas for the GNSS reference stations and calibration. The stations should not be more than 20 km away from each surveying position of the platform to minimize the GNSS error component in the ALS data (FEMA 2003). To avoid incomplete ground coverage due to rolling of the aircraft and improper straight flight lines at least a 20% overlap is advisable (Wehr, 2009). Minimizing the number of aircraft turns reduces the costs, as a substantial part of the flight time is spent turning the aircraft around. The aircraft should also fly a series of manoeuvres over the GNSS reference station before and after the survey to ensure correct alignment of the IMU (Devereux and Amable, 2009). In addition, taking swaths crossing the swaths parallel to the flight line is needed to reduce systematic off-sets in the post-processing stage.

From point cloud to DTM

The creation of useable data from an ALS survey requires significant processing. Such a survey produces several data streams which require integration to generate a point cloud containing locations, elevations and intensities for each reflecting object (first return, last return etc.) returned for a laser shot (Devereux and Amable, 2009). Figure 4 displays the main stages of the entire data processing workflow. The entire process is usually performed in software tailored to the particular scanning device by its vendor but there are also freely available alternatives such as LAStools (Isenburg, 2013), FullAnalyze (IGN, 2011), or v.lidar in GRASS GIS (Brovelli et al., 2004). Usually, ALS measurements are distributed as ASCII point data for further processing by the user. The LAS file format is widely accepted as a public file format for post-processed LiDAR data distribution (ASPRS, 2012). Most of the geomorphological tasks are undertaken in a geographic information system (GIS). However, current GIS software and hardware still faces the challenge of processing the massive amount of ALS measurements so that the form of representation allows for further processing in GIS. DEM is the most important product of ALS either in the form of a DSM or a DTM.

Prior to generating a DTM the ALS points need to be classified by filtering because the points can represent the surface of different objects. The main filtering algorithms were compared in several studies (Sithole and Vosselman 2004, Kobler et al. 2007). All filtering procedures are based on comparing the points with their local neighbourhood. The main parameters used for distinguishing between terrain and off-terrain points are: height difference (Axelsson, 2000), slope angle (Vosselman 2000), surface curvature (Evans and Hudak, 2007), or signal intensity (Bao et al. 2008). If the value of the parameter for a given point exceeds the defined threshold for its neighbourhood, the point is classified as an off-terrain point. Wooded areas with steep terrain pose specific challenge to the filtering algorithms therefore its solution is currently an important research objective (e.g. Kobler et al. 2007, Bao et al. 2008).

Once the points are classified, the ground points can be used to generate a DTM (Figure 1). Further detailed information on this process can be found in *Geomorphological Techniques Section 2.3.1: Creation of DEMs from survey data*. Various methods can be used for creation of DEMs, in general. However, there are some specifics related to redundancy and spatial density of the ALS data. Grid-based representation of ALS DTMs is more popular for easier handling of raster data as opposed

to the vast amount of points in the vector-based TIN data structure. It was shown that the massive ALS point cloud could be decimated before generating a DTM while the plausibility of the resulting DTM remains sufficient for the particular tasks (Liu, 2008).

When generating a gridded DTM it is important to choose the appropriate method and cell size which control the DTM surface quality. If the density of points is homogeneous and at least one point occurs within a single cell of the resulting gridded DTM, the ALS points can be directly converted to a grid by vector-to-raster conversion. Input points often do not fall within every cell of the resulting DTM or the points are heterogeneously distributed. Hence, methods of spatial prediction need to be applied to interpolate elevations for grid cells. Various studies report that the choice of interpolation method is less influential when a surface is interpolated to coarser grids than the density of the input data (Rees, 2000; Smith et al., 2005). Therefore, simpler methods that are computationally less demanding are applicable (e.g. direct point-to-raster conversion using mean elevation value, IDW, or bilinear interpolation). If the cell size is much smaller than the average spacing between the points more complex methods are recommended such as splines (Mitasova et al. 2005) or kriging (Lloyd and Atkinson, 2006). Figure 5 demonstrates this effect.

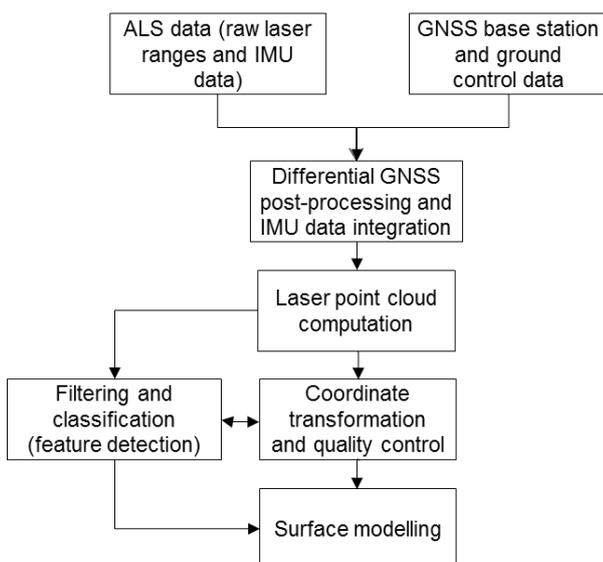


Figure 4: Schematic overview of main data processing stages after Devereux and Amable (2009).

Accuracy

The overall accuracy of ALS data sets depends on several internal, as well as external, factors. The majority of the total error budget in ALS data can be attributed to the accuracy of the GNSS/IMU system (Hodgson and Bresnahan, 2004). Other factors involve technical parameters and settings of the ALS system and also external factors such as object reflectance, atmospheric conditions, sunshine, terrain slope, or density of vegetation cover. Generally, the vertical accuracy is two to three times higher than the horizontal accuracy (Baltsavias, 1999). Table 1 reports the associated accuracy for the ALS systems listed. Additionally, vertical accuracy deteriorates with increasing terrain slope angle when horizontal error other than zero exists (Hodgson and Bresnahan, 2004). This aspect is important for geomorphological

applications in complex morphology (e.g. mountain catchments, badlands). Users are also frequently confronted with systematic errors in DTM/DSM data such as artefacts, mismatch of flight strips and distortions in the rendering of data. The systematic errors can be corrected by reducing the discrepancies between overlapping strips or ALS system calibration (Katzenbeisser 2003, Bang et al. 2009).

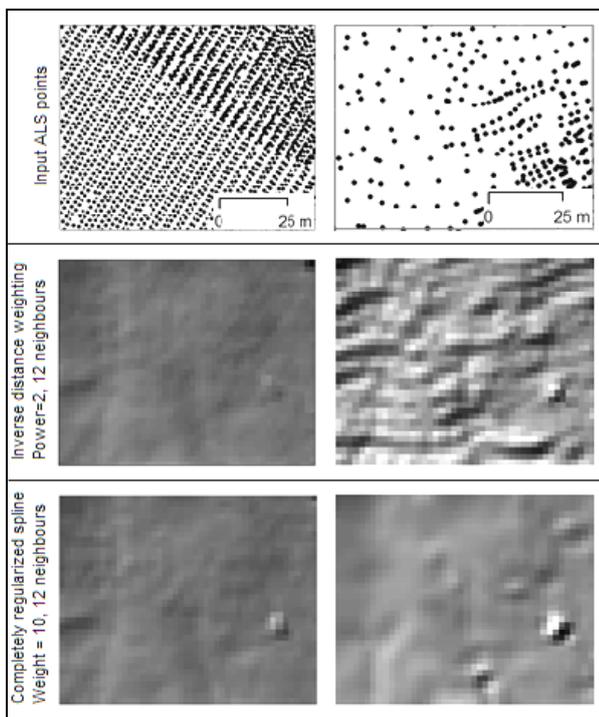


Figure 5: Hillshade surface models of an identical area interpolated from homogeneously and heterogeneously distributed ALS points into gridded DTMs of 2 metres cell size using IDW and spline interpolation.

Determining the accuracy of an ALS derived DSM or DTM is customarily based on in-situ check points that are surveyed with a more precise technique (e.g. dGPS) within the scanned area. There are different standards used for assessing DTMs and DSMs (FEMA 2003, Höhle and Potuckova, 2011) which assume normal distribution of errors and use the RMSE statistics as the main indicator. For assessing the accuracy of DTMs derived from ALS data robust statistics should be used (Höhle and Höhle, 2009). Several empirical studies revealed vertical accuracies of ALS DTMs between 0.08 and 0.33 m RMSE, which were subject to parameters of the platform and environmental conditions (Gallay et al., 2013). *Section 2.3.1: Creation*

of DEMs from Survey Data gives a detailed explanation of the error associated with DEMs.

Advantages and limitations

ALS has demonstrated its strengths where dense, accurate and rapid collection of height data is required. As an active remote sensing technique, ALS mapping is not dependent on daytime, weather or land cover. The best performance is achieved during a calm night, with cool and dry atmosphere (Baltsavias 1999). Compared to photogrammetry, sampling terrain heights with ALS is independent of diversity of image texture. Therefore ALS enables reliable 3-D terrain reconstruction even for areas with snow cover or sand.

The major advantage of ALS is in the ability to reconstruct the terrain under the vegetation canopy making it possible to derive more accurate DTMs even in densely forested areas (Kraus and Pfeifer, 1998). Despite the last return data clearly has the potential to penetrate vegetation cover, it is always likely that the last portion of a pulse will have not reached the ground and will have been reflected from higher in the canopy. As a consequence, these datasets often overestimate terrain (Devereux and Amable 2009). Current topographic ALS systems employ infrared lasers (Table 1), which is a very important limitation for mapping fluvial and coastal systems. The infrared pulse is adsorbed by water so that low or no reflection of the signal is recorded from the water surface. The lack of information on river or sea bathymetry has been recently overcome by the simultaneous use of infrared and green lasers. The measurement of water depth relies on the differential timing of laser pulses reflected from the water surface (infrared laser) and the underwater surface (green laser) to determine the water depth at the point where the laser pulses strike the water surface (Cavalli and Tarolli, 2011).

The use of ALS was limited by the high costs involved in the early years. Hence, ALS was preferentially utilized for mapping elongated or small areas of a maximum of several hundreds of kilometres squared. Nowadays, advances in the technology have decreased the costs and increased the amount and diversity of applications. Important coverage has been achieved across many European

countries, in the U.S, Canada and other regions. The amount of data acquired with ALS for a given area can be excessive for processing with conventional computers. This is an important issue to consider especially with full waveform mapping when data volumes can increase by 50 to 200 times in comparison with single return ALS techniques (Hug et al. 2004).

The high level of detail captured in ALS data can be regarded as a constraint in cases where processes or phenomena are to be observed or modelled on a larger scale (Wood 2009). Figure 6 illustrates this effect showing the spatial pattern of topographic

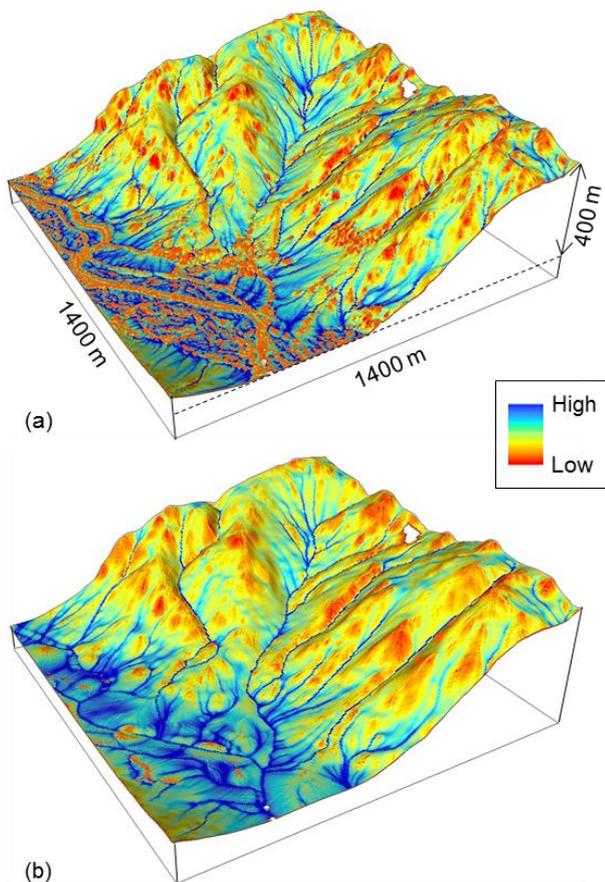


Figure 6: Topographic wetness index draped over the source DTM generated from (a) ALS data and (b) digitized contours showing the effect of differing spatial variation of elevations of a part of the Lake District, England. The spatial resolution of the DTMs is identical (5 metres cell size) but the level of detail contained in the input data differs being much higher for (a) and lower for (b) making the DTM surface of (b) smoother. (Source DTM data: (a) Environment Agency UK, (b) Crown Copyright).

wetness index (TWI) which is a common input parameter for hydrological modelling (Wise, 2007). The pattern is clearly different when comparing Figure 6a and Figure 6b, most markedly in the flat valley floor. This is due to differing spatial variation of elevation, i.e. level of detail and accuracy of the source DTM data. Using the two datasets in a hydrological model could yield different results. This example emphasizes the need for careful statement of the purpose to be addressed by ALS.

ALS in geomorphological research

The application of the ALS data in geomorphological research is very diverse. Höfle and Rutzinger (2011) reviewed the applications according to the main data products derived from ALS measurements comprising (i) raster-based digital elevation models, (ii) the point cloud, (iii) the integration of laser radiometric information, and (iv) the most recently available full-waveform ALS data. Table 2 summarises examples covering a large variety of application fields from a methodological and data processing point of view such as DTM enhancement, land cover classification, dating the landforms, general geomorphological mapping and input for modelling floods, debris flows as well as a thematic consideration such as mass movement processes and landforms, fluvial and coastal studies, and glacier and volcano related investigations.

Repeated ALS over a specific area can be used for studying earth surface dynamics when digital models of difference can be applied to the investigation of geomorphic change (Lane et al. 2003, Mitasova et al. 2009, Ventura et al. 2011). Recent technological developments of LiDAR sensors continue to explore the potential added value of LiDAR radiometric and full-waveform attributes (Alexander et al. 2010). The fusion of laser scanning data acquired from airborne and terrestrial platforms has also been explored (e.g., Bremer and Sass, 2012; Bell et al., 2013). The benefits of ALS compared to other sources of data used for terrain modelling such as synthetic aperture radar, electronic tachymetry, GPS surveys or topographic maps are outlined in Smith et al. (2006), Chiverrell et al. (2008).

Table 2. Examples of studies using LiDAR in a range of environments.

Application realm	Objective	Reference
Automated mapping and feature extraction	manual interpretation, semiautomatic object detection and interpretation automated geomorphological mapping extraction of 3-D breaklines detection of water bodies	Notebaert et al. (2009) Asselen and Seijmonsbergen (2006) Mandlburger and Briese (2009) Brzank et al. (2008), Höfle et al. (2009)
Tectonics	identification of tectonic structures	Cunningham et al. (2006), Székely et al. (2009)
Volcanology	mapping of volcanic deposits of eruptions	Csatho et al. (2008)
Landslides	identification and characterizing of landslides	McKean, J. and Roering, J. (2004), Ventura et al. (2011)
Glaciology	glacier landform mapping and monitoring calculating glacier mass balances estimating snow depth distribution glacier and periglacial surface interpretation	Smith et al. (2006) Geist et al. (2005) Hopkinson et al. (2004) Hopkinson and Demuth (2006)
Karst morphology	identification of landforms for archaeological research	Weishampel et al. (2011), Bernardini et al. (2013)
Fluvial morphology	river valley mapping river morphology quantifying riverbank erosion	Jones et al. (2007), Hohenthal et al. (2011) Charlton et al. (2003), Cavalli et al. (2008) Lane et al. (2003) Thoma et al. (2005)
Coastal morphology	extraction of the tidal channel networks morphodynamic changes of beaches sand dunes pattern storm surge effect, beach fill	Mason et al. (2006) Sallenger et al. (2003), Starek et al. (2012) Mitasova et al. (2009) Richter et al. (2013)
Seafloor morphology	bathymetric mapping	Finkl et al. (2005), Doneus et al. (2013)
Geochronology	dating of landforms	Glenn et al. (2006)
Process simulation	modelling floods modelling debris flows rockfall modelling avalanches modelling modelling river hydraulics	Cobby et al. (2003) Conway et al. (2010) Lan et al. (2010) McCollister and Comey (2009) French (2003), Straatsma and Baptist (2008)

Available sources of ALS data in the UK and other regions

The availability and accessibility of ALS data differs depending on the world region. Free data sets are available for researchers to download via servers listed in Table 3. The majority of the resources cover various parts

of the U.S., but there is also coverage for other parts of the world including Europe. From the European perspective, the ALS data coverage is contiguous on a national level but not for the whole of Europe. In most of the countries the data are distributed by the governmental mapping agencies, which contracted private vendors for data

acquisition. The Netherlands, Belgium, Denmark, Switzerland, Sweden and Finland have already mapped their entire territory or a significant part of it. Other countries, such as the UK, Germany, Poland, Italy, Spain, or Czech Republic systematically map their landscape to achieve full coverage in the near future.

In the UK, the areas with a frequent occurrence of flooding and large cities were preferentially scanned by the Environment Agency (Fig. 7). The Agency has used ALS data acquisition over England and Wales since 1998 achieving 68% coverage in 2012 (EAUK, 2013). The data are provided free of charge for the UK academic and research community, which can also make use of the Natural Environment Research Council's Airborne Research and Survey facility (NERC ARSF) (NERC 2012). Table 4 lists other providers of ALS data services. In some aspects, interferometric radar data provide an alternative to the ALS data (Intermap 2012).

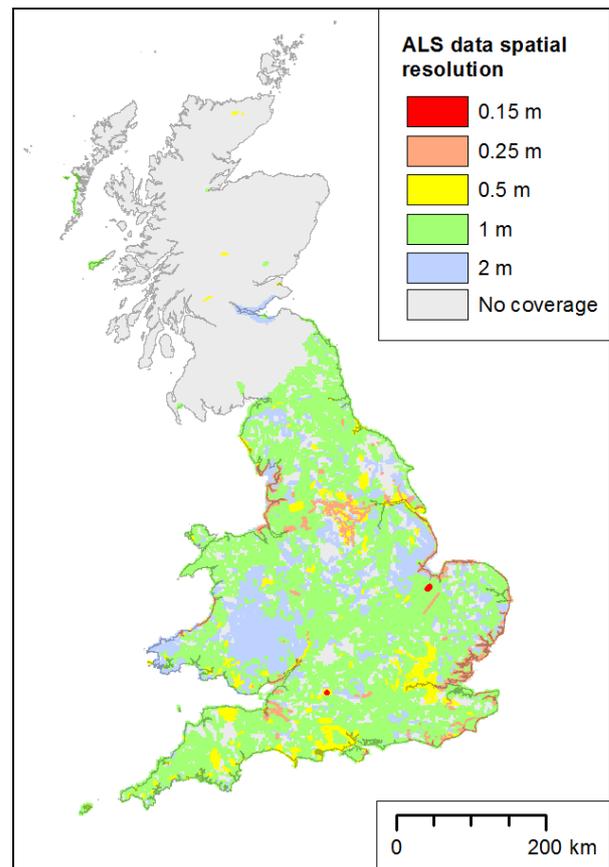


Figure 7: Spatial distribution of overlapping ALS data segments of various resolutions acquired by the Environment Agency UK for Great Britain in 2013. (Source: EAUK 2013).

Table 3. Web-based sources of freely accessible ALS data.

Data Source / Coverage	URL
Lidar On-line / World	http://www.lidar-online.com
Open Topography / USA	http://www.opentopography.org
USGS / USA	http://lidar.cr.usgs.gov
NOAA / USA	http://www.csc.noaa.gov/digitalcoast
Puget Sound Lidar Consortium / USA	http://pugetsoundlidar.ess.washington.edu
Louisiana State University / USA	http://atlas.lsu.edu/lidar/
Red River Basin Decision Information Network / USA	http://www.rrbdin.org/tools
North Carolina Floodplain Mapping Program / USA	http://www.ncfloodmaps.com/
Landmap Service / UK	http://www.landmap.ac.uk
GeoEuskadi / Basque Autonom. Community, Spain	http://www.geo.euskadi.net/s69-geonot/en/contenidos/noticia/derivados_lidar_2008/en_noticia/derivados_lidar_2008.html
National Land Survey of Finland / Finland	https://tiedostopalvelu.maanmittauslaitos.fi/tp/kartta?lang=en

Table 4. Companies providing ALS survey (S) or data distribution (D) in the UK and Ireland.

Data provider	Data pricing*	URL
Environment Agency (S, D)	125; 100 **	https://www.geomatics-group.co.uk
BLOM UK (S, D)	On demand	http://blomasa.com/
Fugro-BKS (S, D)	On demand	http://www.bks.co.uk
Astrium (incorporating Infoterra) (S, D)	75; 75	http://www.geostore.com/geostore4
BlueSky (S, D)	233; 186	http://www.bluesky-world.com
Merret Survey (S, D)	On demand	http://www.merrettsurvey.com/
CentreMaps (D)	125 to 199; 100	http://www.centremaps.co.uk
Scottish Environment Protection Agency (D)	On demand	http://www.sepa.org.uk
Geological Survey Ireland (D)	On demand	http://www.dcenr.gov.ie
Ordnance Survey Ireland (S, D)	On demand	http://www.osi.ie/Products/Professional-Mapping/Height-Data.aspx

* Price in GBP excluding VAT valid in 2013 per 1 km sq. DTM tile of 1m and 2 m cell size, respectively.

** 100% discount for non-commercial use.

Conclusion

The ALS technology has revolutionised the collection of DEM data. ALS obtains thousands of georeferenced point measurements per second, thus providing an unprecedented level of detail in terrain mapping where other methods are less efficient. These factors have stimulated ALS application in geomorphology. The technology has progressed significantly in the last five years by providing full-waveform recording for commercially available small-footprint systems, making ALS more relevant to an even larger field of applications. Combined integration of laser scanners, hyperspectral sensors, or other remote sensing devices appears a promising way of extracting even more high resolution information about the landscape. Future development in ALS is likely to focus on improving the accuracy of the on-board GNSS and IMU, the pulse repetition rate and the range resolution. This will continue to enhance ALS-derived digital elevation information and result in the technology becoming more affordable.

References

- Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J. 2010. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, **65**(5), 423-432.
- ASPRS 2012. LAS Specification, version 1.4–R 12 10 June 2012. American Society for Photogrammetry and Remote Sensing. http://asprs.org/a/society/committees/standards/LAS_1_4_r12.pdf Last accessed: 22nd March 2013.
- Asselen, S. van, Seijmonsbergen, A. C., 2006. Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM. *Geomorphology*, **78**(3–4): 309–320.
- Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, **33**(B4/1): 110-117.
- Baltsavias, E. P. 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2–3): 199–214.
- Bang, K.-I., Kersting, A. P., Habib, A., Lee, D.-C. 2009. LIDAR system calibration using point cloud coordinates in overlapping strips. *ASPRS 2009 Annual Conference, Baltimore, Maryland, March 9-13, 2009*. <http://www.asprs.org/a/publications/proceedings/baltimore09/0011.pdf>. Last accessed 28th June 2013.

- Bao, Y.F., Li, G.P., Cao, C.X., Li, X.W., Zhang, H., He, Q.S., Bai, L.Y., Chang, C.Y. 2008. Classification of LIDAR point cloud and generation of DTM from LIDAR height and intensity data in forested area. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **37** (Part B/3b): 313–318.
- Bell, A., McKinley, J., Hughes, D. 2013. Spatial and temporal analyses for multiscale monitoring of landslides: Examples from Northern Ireland. *Geophysical Research Abstracts*, **15** (EGU2013-4030-1): 1–1.
- Bernardini, F., Sgambati, A., Montagnari Kokelj, M., Zaccaria, C., Micheli, R., Fragiaco, A., Tiussi, C., Dreossi, D., Tuniz, C., De Min, A. 2013. Airborne LiDAR application to karstic areas: the example of Trieste province (north-eastern Italy) from prehistoric sites to Roman forts. *Journal of Archaeological Science*, **40**(4): 2152–2160.
- Bremer, M., Sass, O. 2012. Combining airborne and terrestrial laser scanning for quantifying erosion and deposition by a debris flow event, *Geomorphology*, **138**(1): 49–60.
- Brovelli M. A., Cannata M., Longoni U.M. 2004. LIDAR Data Filtering and DTM Interpolation Within GRASS. *Transactions in GIS*, **8**(2): 155–174.
- Brzank, A., Heipke, C., Goepfert, J. and Soergel, U. 2008. Aspects of generating precise digital terrain models in the Wadden Sea from lidar-water classification and structure line extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, **63**(5): 510–528.
- Cavalli, M., Tarolli, P. 2011. Application of lidar technology for river analysis. *Italian Journal of Engineering Geology and Environment*. Special Issue 2011, 33–44.
- Cavalli, M., Tarolli, P., Marchi, L., Giancarlo, Fontana, G. D. 2008. The effectiveness of airborne LiDAR data in the recognition of channel-bed morphology. *Catena*, **73**(3): 249–260.
- Charlton, M. E., Large, A. R. G. 2003. Application of airborne LiDAR in river environments: the River Coquet, Northumberland. *Earth Surface Processes and Landforms*, **28**(3): 299–306.
- Chiverrell, R.C., Thomas, G.S.P., Foster, G.C. 2008. Sediment-landform assemblages and digital elevation data: Testing an improved methodology for the assessment of sand and gravel aggregate resources in north-western Britain. *Engineering Geology*, **99**(1–2): 40–50.
- Chust, G., Grande, M., Galparsoro, I., Uriarte, A., Borja, A. 2010. Capabilities of the bathymetric Hawk Eye LiDAR for coastal habitat mapping: A case study within a Basque estuary. *Estuarine, Coastal and Shelf Science*, **89**(3), 200–213.
- Cobby, D.M., Mason, D.C., Horritt, M.S., Bates, P.D. 2003. Two-dimensional hydraulic flood modelling using a finite-element mesh decomposed according to vegetation and topographic features derived from airborne scanning laser altimetry. *Hydrological Processes*. **17**(10): 1979–2000.
- Conway, S.J., Decaulne, A., Balme, M.R., Murray, J.B., Towner, M.C. 2010. A new approach to estimating hazard posed by debris flows in the Westfjords of Iceland. *Geomorphology*, **114**(4): 556–572.
- Csatho, B., Schenk, T., Kyle, P., Wilson, T. and Krabill, W. B. 2008. Airborne laser swath mapping of the summit of Erebus volcano, Antarctica: Applications to geological mapping of a volcano. *Journal of Volcanology and Geothermal Research*, **177**(3): 531–548.
- Cunningham, D., Gregbby, S. and Tansey, K. 2006. Application of airborne LiDAR to mapping seismogenic faults in forested mountainous terrain, southeastern Alps, Slovenia. *Geophysical Research Letters*. **33**: L20308.
- Devereux, B., Amable, G. 2009. Airborne LiDAR: instrumentation, data acquisition and handling. In: Heritage, G.L., Large, A.R.G. *Laser Scanning for the Environment Sciences*. Wiley-Blackwell: pp. 49–66.
- Doneus, M., Doneus, N., Briese, C., Pregebauer, M., Mandlbauer, G., Verhoeven, G. 2013. Airborne laser bathymetry – detecting and recording submerged archaeological sites from the air. *Journal of Archaeological Science*, **40**(4), 2136–2151.
- EAUK 2013. Geomatics Group. Environment Agency. <https://www.geomatics-group.co.uk/GeoCMS/Products/LIDAR.aspx>. Last accessed 6th June 2013.
- Eck, C., Imbach, B. 2011. Aerial magnetic sensing with UAV helicopter. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*,

- Vol. XXXVIII-1/C22 UAV-g 2011, Conference on Unmanned Aerial Vehicle in Geomatics, Zurich, Switzerland.: <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-1-C22/81/2011/isprsarchives-XXXVIII-1-C22-81-2011.pdf> Last accessed: 4th April 2013.
- Evans, J.S., Hudak, A.T. 2007. A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, **45**(4): 1029-1038.
- FEMA 2003. *Guidelines and Specifications for Flood Hazards Mapping Partners: Appendix A*. <http://www.fema.gov/library/viewRecord.do?id=2206> Last accessed: 29th May 2013.
- Finkl, C.W., Benedet, L., Andrews, J.L. 2005. Interpretation of Seabed Geomorphology Based on Spatial Analysis of High-Density Airborne Laser Bathymetry. *Journal of Coastal Research*, **21**(3): 501-514.
- French, J.R. 2003. Airborne LiDAR in support of geomorphological and hydraulic modelling. *Earth Surface Processes and Landforms*, **28**(3): 321–335.
- Gallay, M., Lloyd, C.D., McKinley, J., Barry, L. 2013. Assessing modern ground survey methods and airborne laser scanning for digital terrain modelling: A case study from the Lake District, England. *Computers and Geosciences*, **51**: 216-227.
- Geist, T., Elvehøy, H., Jackson, M. and Stötter, J. 2005. Investigations on intra-annual elevation changes using multi-temporal airborne laser scanning data: case study Engabreen, Norway. *Annals of Glaciology*, **42**: 195–201.
- Glenn, N.F., Streutker, D.R., Chadwick, D.J., Thackray, G.D., Dorsch, S.J. 2006. Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. *Geomorphology*, **73**(1–2): 131–148.
- Pike, R.J., Evans, I.S., Hengl, T. 2009. In: Hengl, T., Reuter, H.I. (Eds.) *Geomorphometry — Concepts, Software, Applications: Developments in Soil Science*, Elsevier, pp 3-30.
- Heritage, G., Large, A. 2009. Principles of 3D Laser Scanning. In Heritage, G., Large, A. (Eds.) 2009. *Laser Scanning for the Environmental Sciences*. Blackwell Publishing Ltd. pp. 21-48.
- Hladik, C., Alber, M., 2012. Accuracy assessment and correction of a LiDAR-derived salt marsh digital elevation model. *Remote Sensing of Environment*, **121**: 224–235.
- Hodgson, M.E., Bresnahan, P. 2004. Accuracy of airborne LiDAR-derived elevation: empirical assessment and error budget. *Photogrammetric Engineering & Remote Sensing*, **70**(3): 331–339.
- Hohenthal, J., Alho, P., Hyyppä, J., Hyyppä, H. 2011. Laser scanning applications in fluvial studies. *Progress in Physical Geography*, **35** (6): 782-809.
- Höfle, B., Vetter, M., Pfeifer, N., Mandlbürger, G., Stötter, J. 2009. Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, **34**(12): 1635-1649.
- Höfle, B., Rutzinger, M. 2011. Topographic airborne LiDAR in geomorphology: A technological perspective. *Zeitschrift für Geomorphologie*, **55**(2): 1-29.
- Höhle, J., Höhle, M. 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**(4): pp. 398–406.
- Höhle and Potuckova, 2011. Assessment of the Quality of Digital Terrain Models, Official Publication no. 60. European Spatial Data Research (EuroSDR). <http://www.euroedr.net/publications/60.pdf> Last accessed: 10th April 2013.
- Hopkinson, C., Sitar, M., Chasmer, L., Treitz, P. 2004. Mapping snowpack depth beneath forest canopies using airborne lidar. *Photogrammetric Engineering and Remote Sensing*, **70**(3): 323–330.
- Hopkinson, C., Demuth, M. N. 2006. Using airborne lidar to assess the influence of glacier downwasting on water resources in the Canadian Rocky Mountains. *Canadian Journal of Remote Sensing* **32** (2): 212–222.
- Hug C, Ullrich A, Grimm A. 2004. Litemapper 5600 – A Waveform-digitizing LiDAR terrain and vegetation mapping system. *Proceedings of the ISPRS working group VIII/2, Freiberg, Germany, October 2004*.
- IGN 2011. Fullanalyze, open source lidar software. Institut Géographique National and

- Maison de la Télédétection, France.
<http://code.google.com/p/fullanalyze/>. Last accessed: 6th August 2013.
- Intermap 2012. Nextmap®.
<http://www.intermap.com/Portals/0/doc/Brochures/Architecture/NEXTMap%20EMEA%20Data%20Sheet.pdf>. Last accessed 4th June 2013.
- Isenburg, M. 2013. LAsTools - award-winning software for rapid LiDAR processing. Version 111216, <http://lastools.org>. Last accessed: 22nd March 2013.
- Jones, A. F., Brewer, P.A., Johnstone, E. and Macklin, M.G. 2007. High-resolution interpretative geomorphological mapping of river valley environments using airborne LiDAR data. *Earth Surface Processes and Landforms*, **32**(10): 1574–1592.
- Katzenbeisser, R. 2003. About the calibration of LiDAR sensors. *Proceedings ISPRS workshop on 3-D reconstruction from airborne laser-scanner and InSAR data, 8–10 October 2003, Dresden, Germany*. http://www.isprs.org/proceedings/XXXIV/3-W13/papers/Katzenbeisser_ALSDD2003.pdf Last accessed: 28th June 2013.
- Kereszturi, G., Procter, J., Cronin, S. J., Németh, K., Bebbington, M., Lindsay, J. 2012. LiDAR-based quantification of lava flow susceptibility in the City of Auckland (New Zealand). *Remote Sensing of Environment*, **125**, 198–213.
- Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Oštir, K. Džeroski, S. 2007. Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain, *Remote Sensing of Environment*, **108**(1): 9–23.
- Kraus, K., Pfeifer, N. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **53** (4): 193–203.
- Lan, H.X., Martin, C.D., Zhou, C.H., Lim, C.H. 2010. Rockfall hazard analysis using LiDAR and spatial modeling. *Geomorphology*, **118**(1–2): 213–223.
- Lane, N.S., Westaway, R.M., Murray Hicks, D. 2003. Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surface Processes and Landforms*, **28**(3): 249–271.
- Lemmens, M. 2009. Airborne Lidar Sensors. *GIM International*, **23**(2), 16–19. http://gim-international.com/files/productsurvey_v_pdfdocument_11.pdf. Last accessed: 22nd March 2012.
- Liu, X. 2008. Airborne LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*, **32**(1): 31–49.
- Lloyd, C. D., Atkinson, P. M. 2006. Deriving ground surface digital elevation models from LiDAR data with geostatistics. *International Journal of Geographical Information Science*, **20**: 535–563.
- Mallet, C., Bretar, F. 2009. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**(1): 1–16.
- Mandlbürger, G., Briese, C., 2009. Automatic derivation of natural and artificial lineaments from ALS point clouds in floodplains. *Geophysical Research Abstracts*, **11** (EGU2009-9820-1): 1–1.
- Mason, D.C., Scott, T.R. & Wang, H.-J. 2006. Extraction of tidal channel networks from airborne scanning laser altimetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, **61**(2): 67–83.
- McCullister, C.M., Comey, R.H. 2009. Using LiDAR (Light Distancing And Ranging) data to more accurately describe avalanche terrain. *Proceedings of the International Snow Science Workshop, Davos 2009*. <http://arc.lib.montana.edu/snow-science/objects/issw-2009-0463-0467.pdf>. Last accessed: 22nd March 2012.
- McKean, J., Roering, J. 2004. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry. *Geomorphology*, **57**(3–4): 331–351.
- Milan, D.J., Heritage, G.L. 2012. LiDAR and ADCP use in gravel bed rivers: Advances since GBR6. In Church, M., Biron, P., Roy, A. (Eds.). *Gravel-bed rivers: Processes, Tools, Environments*, John Wiley & Sons, Chichester: pp. 286–302.
- Mitasova, H., Overton, M.F., Recalde, J.J., Bernstein, D.J., Freeman, C.W. 2009. Raster-Based Analysis of Coastal Terrain Dynamics from Multitemporal Lidar Data. *Journal of Coastal Research* **25**(2): 507–514.
- Mitasova, H., Mitas, L., Harmon, R.S. 2005. Simultaneous spline approximation

- and topographic analysis for lidar elevation data in open source GIS. *IEEE Geoscience and Remote Sensing Letters*, **2**: 375-379.
- Morsdorf, F., Frey, O., Meier, E., Itten, K.I., Allgöwer, B. 2008. Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning. *International Journal of Remote Sensing*, **29** (5): 1387-1406.
- NERC 2012. The ARSF Instrument Suite. Natural Environment Research Council. <http://arsf.nerc.ac.uk/instruments>. Last accessed 4th June 2013.
- Notebaert, B., Verstraeten, G., Govers, G. and Poesen, J. 2009. Qualitative and quantitative applications of LiDAR imagery in fluvial geomorphology. *Earth Surface Processes and Landforms*, **34** (2): 217–231.
- Razak, K.A., Straatsma, M.W., van Westen, C.J., Malet, J.P., de Jong, S.M. 2011. Airborne laser scanning of forested landslide characterization: terrain model quality and visualization. *Geomorphology*, **126**(1-2): 186-200.
- Rees, W. G. 2000. The accuracy of Digital Elevation Models interpolated to higher resolutions. *International Journal of Remote Sensing*, **21**: 7-20.
- Richter, A., Faust, D., Maas, H.-G. 2013. Dune cliff erosion and beach width change at the northern and southern spits of Sylt detected with multi-temporal Lidar, *CATENA*, **103**: 103-111.
- Sallenger, A.H., Krabill, W.B., Swift, R.N., Brock, J., List, J., Hansen, M., Holman, R.A., Manizade, S., Sontag, J., Meredith, A., Morgan, K., Yunkel, J.K., Frederick, E.B. and Stockdon, H. 2003. Evaluation of airborne topographic lidar for quantifying beach changes. *Journal of Coastal Research*, **19**(1): 125–133.
- Schnadt, K., Katzenbeisser, R. 2004. Unique airborne fiber scanner technique for application-oriented LIDAR products. In Thies, M., Koch, B., Spiecker, H., Weinacker, H. (Eds.). *ISPRS Archives – Volume XXXVI-8/W2, WG VIII/2 Laser-Scanners for Forest and Landscape Assessment*, 03-06 October, 2004, Freiburg, Germany. <http://www.isprs.org/proceedings/XXXVI/8-W2/SCHNADT.pdf>. Last accessed: 26 June 2013.
- Sithole, G., Vosselman G. 2004. Experimental comparison of filter algorithms for bare—Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, **59**: 85–101.
- Smith, S. L., Holland, D. A., Longley, P. A. 2005. Quantifying Interpolation Errors in Urban Airborne Laser Scanning Models. *Geographical Analysis*, **37**: 200-224.
- Smith, M. J., Rose, J. and Booth, S. 2006. Geomorphological mapping of glacial landforms from remotely sensed data: An evaluation of the principal data sources and an assessment of their quality. *Geomorphology*, **76**: 148–165.
- Starek, M.J., Vemula, R., Slatton, K.C. 2012. Probabilistic detection of morphologic indicators for beach segmentation with multitemporal LiDAR measurements. *IEEE Transactions on Geoscience and Remote Sensing*, **50**(11 Part 2): 4759-4770.
- Straatsma, M.W., Baptist, M.J. 2008. Floodplain roughness parameterization using airborne laser scanning and spectral remote sensing. *Remote Sensing of Environment*, **112**(3): 1062–1080.
- Székely, B., Zámolyi, A., Draganits, E., Briese, C. 2009. Geomorphic expression of neotectonic activity in a low relief area in an Airborne Laser Scanning DTM: A case study of the Little Hungarian Plain (Pannonian Basin). *Tectonophysics*, **474** (1–2): 353–366.
- Thoma, D.P., Gupta, S.C., Bauer, M.E., Kirchoff, C.E. 2005. Airborne laser scanning for riverbank erosion assessment. *Remote Sensing of Environment*, **95**(4): 493–501.
- Ventura, G., Vilardo, G., Terranova, C. and Sessa, E.B. 2011. Tracking and evolution of complex active landslides by multi-temporal airborne LiDAR data: The Montaguto landslide (Southern Italy). *Remote Sensing of Environment*, **115**(12): 3237-3248.
- Vosselman, G. 2000. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **33**(B3/2): 935-942.
- Waser, L.T., Baltsavias, E., Ecker, K., Eisenbeiss, H., Feldmeyer-Christe, E., Ginzler, C., Kuchler, M., Zhang, L. 2008. Assessing changes of forest area and shrub encroachment in a mire ecosystem using digital surface models and CIR aerial images.

- Remote Sensing of Environment*, **112**(5): 1956-1968.
- Wehr, A. 2009. LiDAR Systems and Calibration. In Shan, J., Toth, C. K. (Eds.) *Topographic Laser Ranging and Scanning Principles and Processing*. pp. 129-172.
- Wehr, A., Lohr, U. 1999. Airborne laser scanning - an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**(2-3): 68-82.
- Weishampel, J.F., Hightower, J.N., Chase, A.F., Chase, D.Z., Patrick, R.A. 2011. Detection and morphologic analysis of potential below-canopy cave openings in the karst landscape around the Maya polity of Caracol using airborne laser scanning. *Journal of Cave and Karst Studies*, **73**(3): 187-196.
- Wise, S.M. 2007. Effect of differing DEM creation methods on the results from a hydrological model. *Computers and Geosciences*, **33**(10), 1351-1365.
- Wood, J. 2009. Chapter 14 Geomorphometry in LandSerf, In: Hengl, T., Reuter, H.I. (Eds.), *Geomorphometry — Concepts, Software, Applications, Developments in Soil Science*, Elsevier, **33**: 333-349.

2.1.7. Unmanned Aerial Vehicles (UAVs) and their application in geomorphic mapping

Christopher Hackney¹ and Alexander I Clayton¹

¹ Geography and Environment, University of Southampton, UK (C.R.Hackney@soton.ac.uk)



ABSTRACT: Detailed topographic surveys are a pre-requisite for many studies into Earth surface processes and dynamics. Often such surveys are required for large (>10 km²) areas and at a relatively high temporal resolution (sub-daily to daily) for use in hazard monitoring, monitoring ecological change, and detailed process studies. Techniques such as Terrestrial Laser Scanning, Total Stations and low-level aerial photography via chartered light aircraft flights may provide the spatial resolution required, but are often costly and time-intensive, making them less viable in obtaining the temporal resolution necessary. Further still, satellite imaging platforms often produce products whose image resolution is too coarse to resolve fine scale topography. Recent technological advances have seen the development of Unmanned Aerial Vehicles (UAVs) as a platform from which to acquire aerial photos over large spatial scales at high temporal resolution. These photos may then be combined as orthophotos for spectral analysis, or used to generate useful digital terrain models through Structure from Motion (SfM) photogrammetric techniques. The tandem development of low-cost, rapid deployment UAV platforms and SfM algorithms has seen the rapid growth in the application of UAVs for generating high-resolution topographic data. Here we detail some of the considerations needed before deployment of UAV systems, before showing how UAVs may be used to collect high resolution aerial photos to enable generation of pro-glacial topography in Iceland.

KEYWORDS: UAVs; fixed-wing platform, rotor-wing platform, aerial surveys; topographic models

Introduction

The acquisition of high resolution topographic data is key to many studies in Earth science. For mapping studies requiring data at high temporal (hourly, daily), and large spatial (>5 km²), scales traditional surveying methods are often costly and time intensive. Recent advances in technology have seen the advent of digital photogrammetry as a viable means of obtaining such high resolution topographic data (Smith *et al.*, 2009; Rosnell and Honavaara, 2012; Fonstad *et al.*, 2013). In tandem, the development and increased affordability of Unmanned Aerial Vehicles (UAVs) as a novel platform with which to collect the low-level aerial photography needed for such photogrammetry has seen a rapid increase in their usage in geomorphological studies (Lejot *et al.*, 2007; Hugenholtz *et al.*, 2013).

Previously the domain of the military, UAVs have seen an increase in civilian and academic use, partly driven by the improvements in affordable miniature GPS and Inertial Measurement Units (IMUs) which enable accurate operation of UAV systems. UAVs come in a range of designs. Large fixed-wing platforms have been adapted from military-grade platforms and are typically 5 m or more in wingspan and may carry payloads greater than 200 kg. These systems may have an extended range of ~500 km but require full aviation clearance and need a large ground operations team (Anderson and Gaston, 2013). Smaller UAV systems may come as either fixed-wing or multi-rotor systems. At this scale, many off-the-shelf designs and user-built kit systems are available. Small fixed-wing UAVs may be only a couple of meters wide. Small rotor-wing platforms may have up to eight rotors and may only weight one or two kilograms. At

the even smaller scale, Micro-drones may weigh less than kilogram, however they have limited flight durations (~10 mins) and payloads (<1 kg).

Of particular interest to the geomorphological community are the small, mini- and micro-UAV systems (Anderson and Gaston, 2013). The flexibility in operation and shorter response times afforded by small UAV systems means they enable rapid deployment and the collection of high spatial- and temporal-resolution datasets where traditional aerial techniques, including larger UAV systems, may not.

The use of UAVs in geomorphological mapping is often facilitated by the application of photogrammetric techniques such as Structure from Motion (SfM) (e.g. Harwin and Lucieer, 2012; Westoby *et al.*, 2012; Micheletti *et al.* 2015). SfM utilises overlapping imagery acquired from multiple viewpoints to reconstruct the camera position and camera geometry. From these reconstructed camera locations it is then possible to generate spatial relationships between common feature points and thereby generate a feature's structure (Westoby *et al.*, 2012; Fonstad *et al.* 2013; Micheletti *et al.* 2015). Given correct deployment and attainment of accurate ground control data, the horizontal accuracy and precision of resultant aerial imagery and Digital Elevation Models (DEMs) generated through SfM can be better than satellite imagery and aerial LIDAR (± 0.2 m; Fonstad *et al.*, 2013; Hugenholz *et al.*, 2013), whilst vertical accuracy is typically better than ± 0.1 m (Fonstad *et al.*, 2013).

After some years of development commercially available small fixed-wing and rotor-wing UAV systems now enable the low-cost acquisition of aerial photos over large areas at high temporal resolution. Coupled with the concurrent development of SfM techniques (Micheletti *et al.*, 2015) and a greater appreciation of the potential errors introduced by these methodologies (James and Robson, 2014; Nouwakpo *et al.*, 2014), low-cost, rapid, high resolution topographic data for use in geomorphological mapping collection is now becoming common place in the geosciences..

Applications of UAVs in geomorphology are wide ranging and include, for example, surveying fluvial bathymetry to map and monitor gravel bar location and change using a small remote controlled motorized vehicle (Lejot *et al.*, 2007). d'Oleire-Oltmanns *et al.* (2012) deployed a small fixed-wing platform to monitor rates of soil erosion over a 6km² area in Morocco, showing how small UAVs may be used to bridge the gap between field scale and satellite imagery. Repeat topographic surveys over a 2.5 km² gully allowed Grellier *et al.* (2012) to constrain rates of gully erosion and vegetation change allowing them to elucidate subsurface processes controlling gully evolution. Niethammer *et al.* (2012) deployed a quadcopter, rotary-wing platform to monitor landslides in the Southern French Alps, mapping failures, source and sink zones from orthophotos and digital terrain models (DTMs) generated from the UAV imagery. UAVs have also been deployed in glacial environments, with Whitehead *et al.* (2013), using a fixed-wing UAV system carrying a Lumix LX3 camera to monitor glacial ablation in consecutive ablation seasons on the Fountain Glacier, Canada.

Additionally, and as technology and capabilities evolve, the ability of UAVs to carry variable payloads will open up the possibility of using multispectral sensors to add value to mapping projects, for example it is already possible to detect the geomorphological controls on crop production from combined multi-spectral and traditional photogrammetric techniques using UAV systems (Dunford *et al.*, 2009; Martinez-Casasnovas *et al.*, 2013). More recent advances have seen the application of survey-grade Lidar equipment on small fixed-wing UAV platforms (Lin *et al.*, 2011). This advance will allow the acquisition of topographic data beneath vegetation which current SfM algorithms do not.

The rest of this article will focus specifically on the use of small UAVs (c. Anderson and Gaston, 2013) in obtaining aerial imagery for the purpose of geomorphological mapping. It will outline considerations when selecting a suitable UAV platform. It will then provide some background as to the legal framework within which such studies must be conducted in the EU. Finally, it will provide a case study

example of the deployment of a small fixed-wing UAV in a pro-glacial environment in Iceland, which details good practice workflows and site-specific considerations for operations in remote and topographically restrictive environments.

Considerations

Hardware

UAVs commonly used for geomorphological surveying are predominately built around two types of airframe; fixed-wing and rotor-wing. Both platforms are used in geomorphological applications working off of the same theoretical standpoint; the acquisition of overlapping, photographs which can later be used with SfM (Fonstad *et al.*, 2013; Micheletti *et al.*, 2015) algorithms to reconstruct topography. Recent work has shown how topographic datasets derived from UAV derived aerial photos are improved by having images captured from non-parallel viewing locations (James and Robson, 2014). That is to say, it may be beneficial to collect imagery from more unstable platforms which enable more photo acquisition from non-uniform camera locations. A summary of typical small fixed-wing and rotary-wing systems is provided in Table 1. These systems reflect those commonly used in geomorphological studies and do not represent the actual maximum values obtained by larger UAV systems (Anderson and Gaston, 2013). Note the specifications and details provided below are limited to

those which may be operated under CAA requirements (see below for more details).

Fixed-Wing Platforms

Fixed-wing platforms (Figure 1A) are perhaps the most common form of UAV. Wings and bodies are normally constructed from lightweight polystyrene with a wingspan typically <2m. The lift characteristics of a flying wing mean that, relative to rotor-wing platforms, their fuel efficiency is high. The limited number of moving parts and lightweight design also means that damage inflicted by hard landings is more limited than with rotor-wing designs (although it is less likely that hard landings will occur with rotor-wing designs). They are launched from the ground either by hand or, more commonly now, with use of a catapult. Whilst the apparatus is small, the necessity for a non-vertical climb to altitude means they require more space to take off than their rotor-wing counterparts. They are larger in size than their rotor-wing counterparts (Figure 1) and as such require more space to operate. They often also have more associated peripheries (additional laptop and launching gear, Figure 1A) required for their operation than rotor-wing UAVs. However, recent advances in fixed-wing platforms have seen an increase in their flight endurance such that they can now fly longer than rotor-wing platforms enabling greater areas to be covered with more ease. They are often controlled in-flight

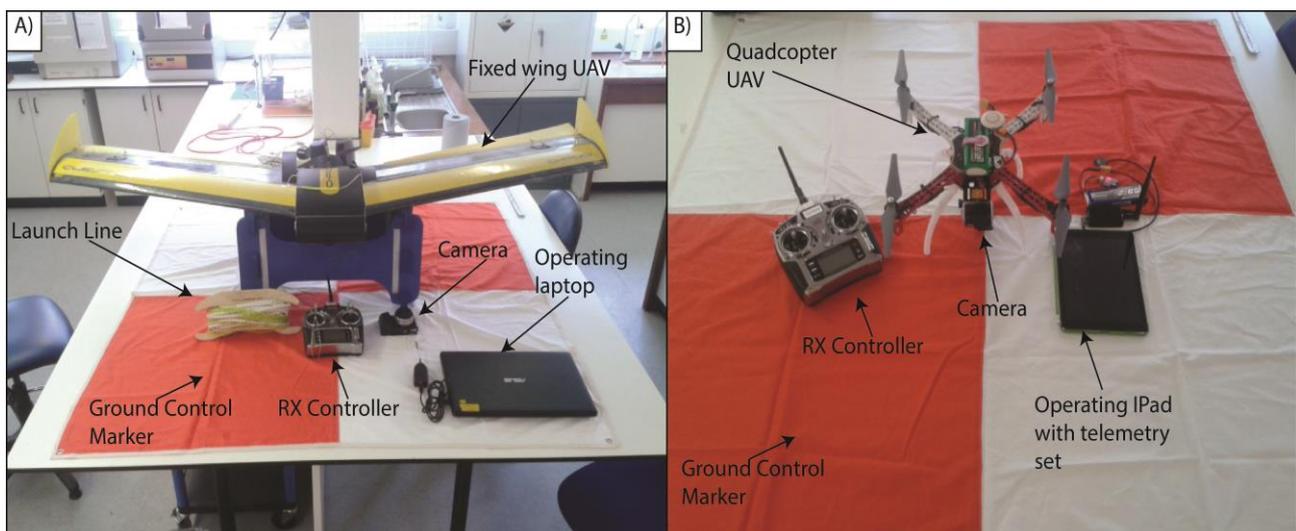


Figure 1: A) A fixed-wing UAV system with associated peripheries required for its operation. B) A quad-copter rotor-wing UAV system and associated operational peripheries. The same 1.5 m² ground control target has been used as a background to provide consistent scale between both A) and B).

Table 1: Technical specification comparison between typical small fixed-wing and multi-rotor UAV platforms which may be used in geomorphological studies. Values in italicised parentheses relate to CAA operating regulations as per CAP393 and CAP722. Values as reported in Hugenholtz et al. (2013), Mancini et al. (2013) and Anderson and Gaston (2013).

	Fixed-wing UAVs	Multi-rotor UAVs
Wingspan (m)	1 – 3	<1
Flight time (mins)	20 - 60	20
Max payload Weight (kg)	30 (7)	15 (7)
Max. Speed (km/h)	50 – 80 (130)	30 – 50 (130)
Operating Range (km)	1 - 5 (0.5)	1 – 2 (0.5)
Altitude Range (m)	> 2000 (121)	400 (121)

by built-in autopilots, with flight plans pre-programmed before deployment. This means they require less user interaction in flight, and are more stable, than rotor-wing platforms. However, this has implications for topographic data sets derived from photos obtained from fixed-wing platforms, as the increased stability of the platform and pre-programmed, often parallel flight lines may introduce errors into the topographic dataset (James and Robson, 2014).

Fixed-wing airframes inevitably require more space than rotor-wing options. Often a ~100 m strip is sensible to allow for overrun and variations in headwind strength. There is also a need for the survey area to be less constricted (e.g. from valley side walls) than is required with rotor-wing UAVs as they have a larger footprint and may be operated at a greater range from the operator.

Fixed-wing UAVs may be more suited to topographic surveys over larger spatial scales due to their longevity in flight. However, their application to locations bounded by terrain, trees or other obstructions may limit their successful operation. Likewise, their relative lack of maneuverability may provide challenges in certain situations.

Rotor-wing Platforms

Rotor-wing platforms comprise a suite of designs ranging from common helicopter designs to quad-copters (with four rotary

blades, Figure 1B), hexa-copters (with six rotary blades), and octa-copters (with eight rotary blades) which are becoming the standard for heavy lift photography work. Their footprint is typically smaller than fixed-wing platforms (~0.8 m). Vertical take-off and landing means that they do not require extensive unconstrained landing sites and can be deployed from relatively inaccessible areas. This favours rapid deployment and enables access to areas previously unfeasible with traditional survey techniques and fixed-wing UAV platforms.

Due to their smaller footprint, rotor-wing UAVs have limited flight endurance (typically less than 20 mins) when compared to fixed-wing platforms, requiring many battery packs and/or recharging units to accomplish the same spatial coverage as fixed-wing UAVs. This increases the likelihood of inclement weather impacting the survey. However, as rotor-wing platforms typically fly at slower speeds than fixed-wing UAVs, and often contain better gimbals, they are more stable at higher wind speeds than fixed-wing UAVs. This permits a trained pilot greater control of the UAV, and facilitates the collection of non-parallel survey lines under a wider range of wind conditions. It is therefore likely that topographic datasets derived from rotor-wing systems are likely to be of higher quality than equivalent data-sets produced from fixed-wing systems as their ability to be more flexible in their survey lines will permit the acquisition of a more non-uniform set of photos (James and Robson, 2014).

Additionally, unlike fixed-wing platforms, rotor-wing platforms are able to hover over objects and locations. This facilitates higher precision photogrammetry over features of interest and allows for complete 3D inspection of stationary objects, whilst also opening the possibility of at-a-point temporal sequences in measurements. Similarly, it enables users to survey the same feature at different altitudes, thereby assessing issues of pixel resolution and photo quality as a function of varying altitude.

Sensing Applications

Both fixed-wing and rotor-wing UAVs have payloads capable of carrying small to medium size digital cameras and video recorders (and are restricted to payloads of 7 kg under CAA regulations; see below for further details). These cameras can be simple RGB digital cameras suitable for the acquisition of high resolution aerial photos which can subsequently be used in the generations of georectified ortho-photos or digital elevation models (e.g. Fonstad *et al.*, 2013; Hugenholtz *et al.*, 2013). Alternatively, these sensing unit can be hyper-spectral cameras which can be used to enable an assessment of the local water stress (Zarco-Tejada *et al.*, 2012) or agricultural and forestry health (e.g. Saari *et al.*, 2011).

More recently, UAVs have been deployed with small Lidar sensors (Lin *et al.*, 2011; Wallace *et al.*, 2012). This growing area allows for high resolution topographic surveys from UAVs to be conducted without recourse to SfM software. It should be noted however that the current suite of Lidar sensors which may be deployed on UAVs are less high powered than traditional Lidars and have a higher signal-to-noise ratio. In addition, the development of smaller Lidar sensors capable of being lifted by UAVs could possibly enable more data to be obtained than tradition photogrammetric techniques, for example Lidar intensity return data can be used to identify moisture variability and surface geology variability more easily than would be visible from RGB camera images. Furthermore, one of the current limitations with SfM is that the cameras used are passive sensors, i.e. they cannot penetrate through vegetation. Having active Lidar, and phased-based sensors onboard UAVs would

enable greater detail of the surface to be captured.

Although these techniques would be less cost effective than photogrammetric techniques for obtaining topographic data, they will still require careful planning of flight lines, acquisition of a detailed ground control network and will still have to operate within the legal restrictions.

Ground Control

For the successful registration and alignment of photos collected from any UAV platform and sensor, an accurate and precisely located network of Ground Control Points (GCPs) must be acquired. To ensure that no lens warping, or 'doming' (c. James and Robson, 2014) is present in the final product, it is vital the GCPs are distributed throughout the study area. Of particular importance is the placement of GCPs close to the edges of the survey area, where potential doming may be exacerbated. The number of GCPs deployed depends on the overall aim of the survey. For example, for geomorphic mapping it is more important to distribute GCPs across the survey area to ensure doming and registration errors are constrained. However, if the survey is designed for obtain detailed topography, then it is important that subtle variations in topography are accurately resolved. As such a denser network of GCPs in locations of interest may be necessary to ensure such fine-scale topography is captured. It is worth noting, that it is still important to maintain a good distribution of GCPs in this case, to avoid doming of the final DEMs.

GCPs normally consist of brightly coloured targets which will be visible within the aerial photographs, laid out by the operator evenly across the survey area. The GCPs should be larger than the pixel resolution of the sensor onboard the UAV such that they are clearly visible within the photographs. As pixel resolution varies in size with survey elevation, surveys conducted at higher altitudes require larger GCPs. The pixel size (P_R) for a sensor with a given focal length (L_F) and pixel dimension (P_D), at a given altitude (A) can be calculated using the following calculation

$$P_R = P_D(A/L_F) \quad (1)$$

Equation 1 can therefore be used to determine the required size of GCPs for a survey of any given altitude. For example,

Smith *et al.* (2009) deployed 30 GCPs across a 40 x 50 m survey area. Their GCPs consisted of 20 cm laminated sheets containing a black and white target design. Smith *et al.* (2009) flew their survey at 50 m altitude. For surveys flown at higher altitudes (> 50m) the size of GCPs deployed are recommended to be larger than 1 m² to be captured in the aerial photos. Alternatively, it is possible to use unique ground features as GCPs, for example, Fonstad *et al.* (2013) identified features on the ground to use as GCPs which provide fixed features in the survey photos (e.g. field corners, lake edges, road junctions).

Once a network of GCPs has been constructed, the position of each GCP must be obtained. This can either be a real world coordinate of the GCP obtained through the use of a differential Global Positioning System (dGPS) or a relative position to an arbitrary coordinate grid obtained through a total station or similar surveying equipment. Either technique should result in positional accuracies of approximately ± 0.05 m. The technique used depends upon the ultimate goal of the survey being conducted. If the data is to be used in conjunction with other data sets and registered within a real-world context then global coordinates will be necessary. However, if this is not the case then it is possible to use an arbitrary, relative, coordinate system.

Legal Limits

The legal requirements for UAV flight vary between countries. It is necessary to research country specific regulations before any survey work is undertaken to ensure you meet the requirements.

Many legal considerations exist when planning a survey with a UAV. Within the EU flights undertaken with a UAV must adhere to central legislation (and the reader is guided towards the Civil Aviation Authority (the governing body within the UK) protocols CAP393 and CAP722 for full details of the legislation which is applicable across the entire EU). Here (for brevity) we will cover a few key points arising from this legislation which require consideration when planning geomorphological surveys. It is stressed that you check the legislation before deployment at each new site to ensure you adhere to the rules.

EU law states that UAVs operated in any EU country airspace must be kept within the visual line of sight. This equates to approximately 500 m in the horizontal and 400 feet in the vertical, although this is heavily dependent upon local weather conditions and terrain. This has practical implications for survey and flight line design, limiting the operational space achievable with each flight. As such, it is likely that multiple launches will need to be made if your area of interest covers a substantial area. Further, this increases the importance of a dense ground control network (see above) to aid stitching of photos and georectification during post-processing.

Additionally, UAVs may not be flown within 50 m of a member of the public (with the exception of the operating crew), thus limiting their use within public spaces and over tourist locations. Similarly, they may not be flown over or within 150 m of any organised open-air assembly of more than 1,000 people.

Case Study

Science questions and aims

Proglacial zones are highly dynamic regions which are subject to seasonal variations in energy regime and thus geomorphological activity. These regions are often inaccessible and remote. The foreland of Skalafellsjokull, Iceland (Figure 2) includes a series of well-preserved push moraines, the spacing of which relate to local climatic conditions (Boulton, 1986; Bennett, 2001). Yet, our understanding of how these features respond to climatic variations at Skalafellsjokull is poorly understood. In order to map these features and determine accurately the inter-seasonal spacing, a high resolution DEM was required (see Chandler *et al.*, In Prep, for details). These features are located in a topographically constrained region which does not easily facilitate high resolution mapping with terrestrial laser scanning, Lidar or satellite radar mapping. As such, the use of a UAV presented a lower cost, more time effective option than standard aerial surveys and due to the size of the site (2 km²) a terrestrial approach would not have been suitable. In addition, satellite imagery is of too coarse a resolution to resolve the spatial scales of the push moraines.

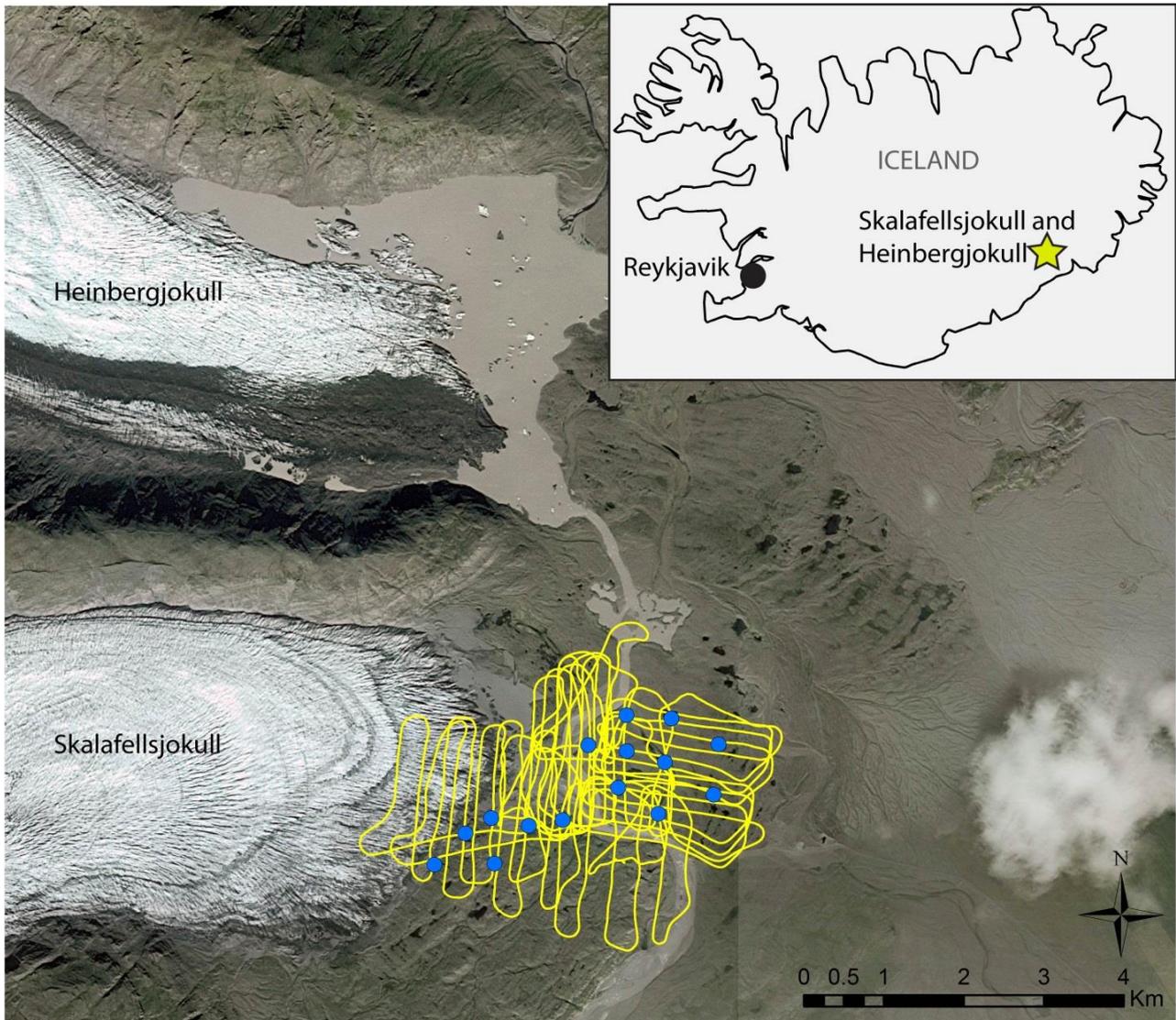


Figure 2: Google Earth image of Heinbergjokull and Skallafellsjokull in Iceland. The UAV flight lines are depicted in yellow and blue dots represent the GCPs used to register the aerial photos.

Methodology and Results

To that end, during 2013 a Quest 200 fixed-wing UAV carrying a Panasonic Lumix LX5 off-the-shelf, point-and-shoot camera (see Table 2 for details) was used to survey proglacial and ice marginal geomorphology at Skallafellsjokull and Heinbergjokull in Iceland (Figures 2 and 3). The aim of the surveys was to obtain low level aerial photography which could subsequently be processed with SfM software to produce high resolution topographic surveys of the study area (Figure 3). The UAV was used in conjunction with a Leica dGPS deployed in Real-Time Kinematic (RTK) mode to limit distortion in the final photogrammetric product. RTK corrections were accurate to ± 0.01 m and

± 0.05 m. in the horizontal and vertical, respectively.

Fifteen ground control point targets were created in the field from orange plastic material and measured 2 m². Tape was used to indicate the centre of the target. The use of tape larger than the pixel resolution enabled a precise determination of target centres. The targets were deployed in a grid network spaced approximately 0.5 km apart and a density of 1.2 points/km²; this results in ~ 0.03 GCPs per image (Figure 2). GCPs were deployed before the UAV was launched to ensure they would be visible in the survey.

The resulting orthophoto (Figure 3A) and DEM (Figure 3B) have spatial resolutions of 0.05 and 0.1 m respectively, and are of suitable resolution to be able to determine

sub-seasonal variations in push moraine location (Chandler *et al.*, In Prep). It is noted that the resolution of the orthophoto and DEM vary with topography, particularly as this site displays large relief (~100 m).

Pre-departure checks

Prior to departing to the field, the flight regulations for Iceland were checked and accounted for in all pre-flight planning. The legal framework in Iceland is different to those in place within the EU (see above for EU restrictions). The Icelandic CAA stated that as long as the UAV was <5 kg there were no specific requirements on altitude and range. This relatively relaxed regulatory stance is a function of the sparsely populated area and allowed for more freedom in survey planning than would otherwise have been available within the EU. However, during the field season surveys were flown within the EU limitations (see above) to avoid any issues which may arise with operating

beyond known limits, experience and ability in the field.

Table 2: Attributes of the Skalafellsjokul and Heinbergjokull surveys conducted in 2013.

Attribute	Value
UAV Platform	Quest UAV 200 Fixed-wing
Sensor	Panasonic Lumix LX5
Survey altitude (m)	100
Photo endlap (%)	80
Photo sidelap (%)	60
No. of images captured	1980
Image resolution (m)	0.05
DEM resolution (m)	0.1

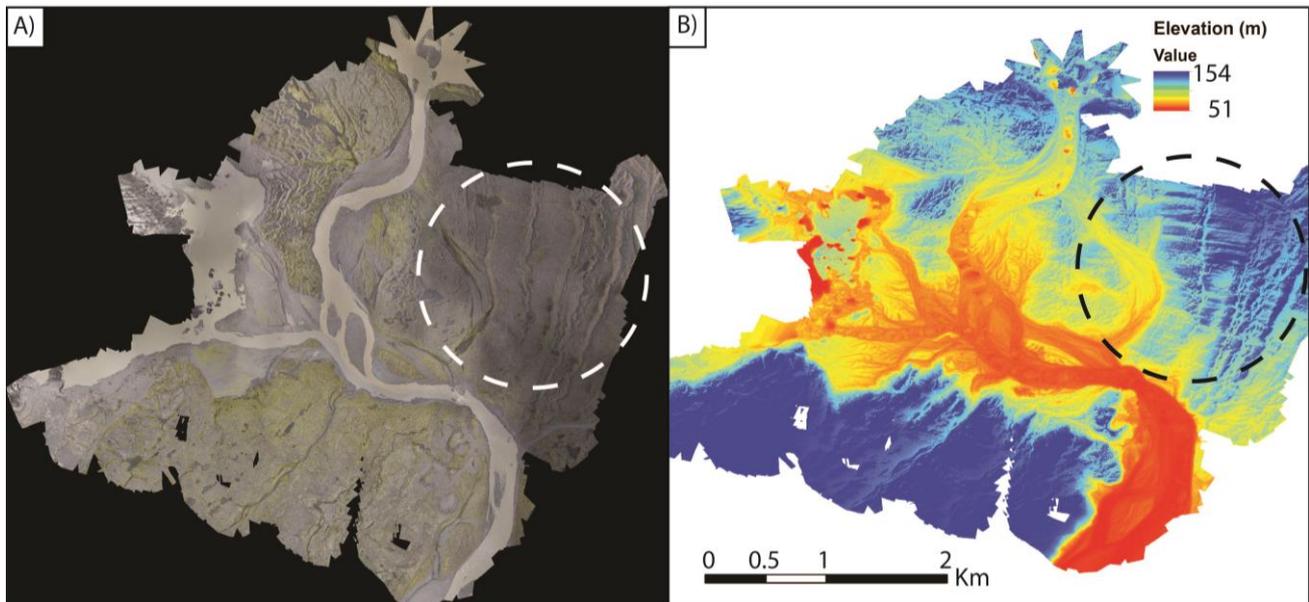


Figure 3: A) Final registered orthophoto of Skallafellsjokull generated from aerial photos captured using the UAV. B) DEM of the Skallafellsjokull proglacial foreland derived from Agisoft Photoscan. The push moraines are identified by the dashed circles in both A and B.

On-site considerations

Access, logistics and changeable weather limited the survey time available, as such the deployment methodology was designed to be flexible and quick. As in many locations worldwide, surveying conditions are highly

changeable, with the possibility of unpredictable wind and precipitation. Accordingly, the survey was carried out as series of small surveys to avoid the prospect of having to cut a longer survey short, thus potentially compromising the acquisition of data. This is facilitated by the rapid

deployment of the fixed-wing UAV, and is one of the advantages of using such technology in these challenging areas. Each survey lasted approximately 20 minutes and covered an area of 2 km². In total an area of 15 km² was surveyed over a total of 7 days. After each survey was conducted the ground control point targets were registered with the dGPS and collected. The above process was repeated over a total of 9 surveys, to ensure the entire survey area was captured.

The study site (Figure 2) is characterised by rugged landscapes and highly changeable weather conditions. Accordingly, the selection of a suitable operating base which permits accessibility, safe take-off and landing sites and enough space to permit safe operating practices is necessary. Similarly it places more emphasis on all pre-flight checks including obtaining accurate weather forecasts the day before deployment, as well as detail on foot (or by vehicle) reconnaissance of the field site(s) to identify access points and areas of shelter and safety should the weather turn all of a sudden.

As with any UAV survey conducted in a maritime climate at high latitudes, weather was the primary concern. Despite being in Eastern Iceland frequent rain storms were possible and high winds probable. There is little that can be done to mitigate these factors, but test flights in the area were an essential part of our preparation for the actual surveys. These helped establish a rough guide to the vertical profile of wind speeds and give an indication as to safe flying conditions.

However, whilst wind speeds provided a good indication of flying conditions and incoming precipitation was simple to spot, the turbulence caused by mixing air masses proved difficult to predict and was a major issue to the operational safety of the UAV. Whilst this is probably common in many valley confined environments, the mixing of the katabatic winds off Vatnajökull with air masses blowing off the Atlantic frequently resulted in unpredictable UAV behavior in apparently stable conditions. Telemetry recorded on the UAV logger revealed that the time of day, local valley topography and valley floor wind direction all appeared to correlate with turbulence at height. However, the lack of capability to monitor the upper air

turbulence meant that it was impossible to remove this issue from the survey operation. Therefore, to limit any potential damage to the UAV and operation team, a sensible emergency rally point for the aircraft was set on the on board auto-pilot and constant visual on the UAV was kept during the survey. If it was felt the atmospheric conditions were exceeding the pilot's operational abilities, the UAV was recalled to its emergency recall point and the survey was halted until conditions became more favourable. At all times vigilance on behalf of the operating crew was essential.

Terrain, in any mountainous environment, presents difficulty to a UAV operator. In Iceland the expansive unpopulated glacial valleys made judging distances difficult, safe landing spots are rare and maintaining line of sight to the aircraft can require limiting surveys. Limiting the impacts these issues requires careful survey planning. Initially, reconnaissance trips to the field site were conducted prior to any deployment to familiarise the operational team with the survey location and surround topography. Prior to departing on the field trip, an accurate and recently geo-referenced image was obtained such that in situations when communication with the UAV is lost, and the operator is reliant on visual cues, they have a detailed map of the surrounding area. It was also decided to identify numerous landing sites which would enable the UAV to be landed when conditions became operationally difficult. As it is beneficial for the pilot to circle into a landing, topographically unconstrained sites of different orientation were selected to cover a range of possible wind directions.

The most difficult landings occurred when the survey area moved onto the glacier foreslope and the UAV was landing on the bare ice of Skalafellsjökull. Despite flying from the flattest area there remained approximately 0.5 m of relief in the ice topography. Whilst landing on snow proved simple, it was difficult to mitigate the damaging landings on the hard ice. In low wind conditions, even by flaring the aircraft almost into a stall before landing, the impact was hard and caused some damage to the airframe. This issue could be avoided if a parachute option (now built in to some UAV systems) was available.

As well as an increased chance of inclement weather, the higher latitudes also presented issues with the quality of the photos obtained from the surveys; namely light angle and quality. However, the low angle of the sun at high latitudes resulted in a very flat lighting angle combined with high levels of shading due to the confined valley environments. Clearly this presents an issue when image quality, and specifically detail, are of paramount importance in a photogrammetric survey.

The available options were limited as weather conditions were the primary restraint on surveying. To that end it is highly worth investing in a camera with good low light sensitivity for flat light situations. Ideally one should be used that has the option to programme a range of apertures, exposures and ISO settings in order to best capture the image with the minimal amount of distortion but best detail. This functionality should enable the user to optimise the camera settings for a range of lighting conditions and facilitate a much broader range of deployment conditions with the same result in image quality.

Conclusion

Unmanned aerial vehicles (UAVs) provide a low-cost, rapid deployment method of obtaining high-resolution aerial photography over areas of varying size. Whether fixed-wing or rotor-wing, UAVs provide a viable alternative to traditional surveying techniques which can be deployed in a range of situations and locations. Although strict restrictions apply with regards to their use and deployment in many locations, their application in a wide range of geomorphological environments (glacial, fluvial, hillslope, coastal) means UAVs are becoming more and more popular in geomorphological research.

Here, we outline some of the considerations and regulations which must be adhered to when operating UAVs in many situations. It is vital that weather conditions are researched and that the operating team have scouted the study site prior to deployment. We use the example of an aerial survey of pro-glacial push moraines in Iceland to detail a suggested best practice when operating UAVs in challenging and remote locations.

Although set in a remote, constricted location, the methodology and work-flow adopted in Iceland can be applied to the majority of geomorphological settings UAVs are likely to be deployed within.

Acknowledgements

Christopher Hackney was supported by award NE/JO21970/1 to the University of Southampton from NERC. Alexander Clayton gratefully acknowledges PhD studentship funding from NERC. We wish to extend our thanks to Tom Bishop and James O'Dwyer for their assistance in the field. The comments of two reviewers and the editor, Lucy Clarke, have greatly improved the quality of this paper.

References

- Anderson K, Gaston KJ. 2013 Lightweight unmanned aerial vehicles will revolutionize spatial ecology, *Frontiers in Ecology and the Environment* **11**: 138-146. DOI: 10.1890/120150.
- Bennett MR. 2001 The morphology, structural evolution and significance of push moraines. *Earth-Science Reviews* **53**(3-4): 197-236. DOI: 10.1016/S0012-8252(00)00039-8.
- Boulton GS. 1986 Push-moraines and glacier-contact fans in marine and terrestrial environments, *Sedimentology* **33**(5): 677-698. DOI: 10.1111/j.1365-3091.1986.tb01969.x
- Chandler BMP, Ewertowski MW, Evans DJ, Roberts DH, Clayton A. (In Prep) Glacial geomorphological mapping of "annual" moraines on the foreland of Skjalafellsjökull, S.E. Iceland, *Journal of Maps*.
- d'Oleire-Oltmanns S, Marzloff I, Peter KD, Ries JB. 2012. Unmanned aerial vehicles (UAVs) for monitoring soil erosion in Morocco. *Remote Sensing* **4**: 3390-3416. DOI: 10.3390/rs4113390.
- Dunford R, Michel K, Gagnage M, Piegay H, Tremelo ML. 2009. Potential and constraints of unmanned aerial vehicle technology for characterization of Mediterranean riparian forest. *International Journal of Remote Sensing* **30**(19): 4915-4935.
- Fonstad MA, Dietrich JT, Courville BC, Jensen JL, Carbonneau PE. 2013. Topographic structure from motion: a new

- development in photogrammetric measurement. *Earth Surface Processes and Landforms* **38**: 421-430. DOI: 10.1002/esp.3366
- Grellier S, Kemp J, Janeau JJ, Florsch N, Ward D, Barot S, Podwojewski P, Lorentz S, Valentin C. 2012. The indirect impact of encroaching trees on gully extension: A 64 year study in a sub-humid grassland of South Africa. *Catena* **98**: 110-119. DOI: 10.1016/j.catena.2012.07.002.
- Harwin S, Lucieer A. 2012. Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. *Remote Sensing* **4**(6): 1573-1599. DOI: 10.3390/rs4061573.
- Hugenholtz CH, Whitehead K, Brown OW, Barchyn TE, Moorman BJ, LeClair A, Riddell K, Hamilton T. 2013. Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model. *Geomorphology* **194**: 16-24. DOI: 10.1016/j.geomorph.2013.03.023.
- James M, Robson S. 2014 Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms* **39**(10): 1413-1420. DOI: 10.1002/esp.3609.
- Lejot J, Delacourt C, Piegay H, Fourier T, Tremelo ML, Allemand P. 2007. Very high spatial resolution imagery for channel bathymetry and topography from an unmanned mapping controlled platform. *Earth Surface Processes and Landforms* **32**: 1705-1725. DOI: 10.1002/esp.1595.
- Lin Y, Hyypä J, Jaakkola A. 2011. Mini-UAV-Borne LIDAR for fine-scale mapping. *IEEE Geoscience and Remote Sensing Letters* **8**(3): 426-430.
- Mancini F, Dubbini M, Gattelli M, Stecchi F, Fabbri S, Gabbianelli G. 2013. Using unmanned aerial vehicles (UAV) for high-resolution reconstructions of topography: the structure from motion approach on coastal environments. *Remote Sensing* **5**(12): 6880-6898. DOI: 10.3390/rs5126880.
- Martinez-Casasnovas J, Ramos M, Balasch C. 2013. Precision analysis of the effect of ephemeral gully erosion on vine vigour using NDVI images. In Stafford, J. (Eds.) *Precision Agriculture '13*. Wageningen Academic Publishers, Wageningen, The Netherlands.
- Micheletti N, Chandler JH, Lane SN, 2015. Section 2.2: Structure from Motion (SfM) Photogrammetry. In: Clarke LE, Nield JM. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology; London, UK. ISSN: 2047-0371.
- Niethammer U, James MR, Rothmund S, Travelletti J, Joswig M. 2012. UAV-based remote sensing of the Super-Sauze landslide: Evaluation and results. *Engineering Geology* **128**: 2-11. DOI: 10.1016/j.enggeo.2011.03.012.
- Nouwakpo SK, James M, Weltz MA, Huang C, Chagas I, Lima L. 2014 Evaluation of structure from motion for soil microtopography measurement. *Photogrammetric record* **147**: 297-316. DOI: 10.1111/phor.12072.
- Rosnell T, Honkavaara E., 2012. Point cloud generation from aerial image data acquisition by a quadcopter type micro unmanned aerial vehicle and a digital still camera. *Sensors* **12**: 453-480. DOI: 10.3390/s120100453.
- Saari H, Pellikka I, Pesonen L, Tuominen S, Heikkilä J, Holmlund C, Mäkyen J, Ojala K, Anttila T. 2011. *Unmanned aerial vehicle (UAV) operated spectral camera system for forest and agriculture applications*. Proceedings SPIE 8174, Remote sensing for agriculture, ecosystems and hydrology XIII. DOI: 10.1117/12.897585.
- Smith M, Chandler J, Rose J. 2009. High spatial resolution data acquisition for the geosciences: kite aerial photography. *Earth Surface Processes and Landforms* **34**: 161-255. DOI:10.1002/esp.1702.
- Wallace L, Lucieer A, Watson C, Turner D. 2012. Development of a UAV-LIDAR system with application to forest inventory. *Remote Sensing* **4**: 1519-1543.
- Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 2012. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **179**: 300-314, DOI: 10.1016/j.geomorph.2012.08.021.
- Whitehead K, Moorman BJ, Hugenholtz CH. 2013. Brief communication: Low-cost, on-demand aerial photogrammetry for

glaciological measurement. *The Cryosphere* **7**: 1879-1884. DOI: 10.5194/tc-7-1879-2013.

Zarco-Tejada PJ, Gonzalez-Dugo V, Berni JAJ. 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and thermal camera. *Remote sensing and the Environment* **117**: 322-337.

2.2.2. Structure from Motion (SfM) Photogrammetry

Natan Micheletti¹, Jim H Chandler², Stuart N Lane¹

¹ Institute of Earth Surface Dynamics, University of Lausanne (natan.micheletti@unil.ch)

² School of Civil and Building Engineering, Loughborough University



ABSTRACT: Topographic data measurement is a fundamental aspect of many geomorphic research applications, particularly those including landform monitoring and investigation of changes in topography. However, most surveying techniques require relatively expensive technologies or specialized user supervision. Structure from Motion (SfM) photogrammetric technology reduces both these constraints by allowing the use of consumer grade digital cameras and highly automated data processing, which can be free to use. SfM photogrammetry therefore offers the possibility of fast, automated and low-cost acquisition of 3-D data, which has inevitably created great interest amongst the geomorphological community. In this contribution, the basic concepts of SfM photogrammetry are presented, whilst recognising its heritage. A few examples are employed to illustrate the potential of SfM applications for geomorphological research. In particular, SfM photogrammetry offers to geomorphologists a tool for high-resolution characterisation of 3-D forms at a range of scales and for change detection purposes. The high level of automation of SfM data processing creates both opportunities and threats, particularly because user control tends to focus upon visualisation of the final product rather than upon inherent data quality. Accordingly, this contribution seeks to guide potential new users in successfully applying SfM for a range of geomorphic studies.

KEYWORDS: Structure from Motion, close-range photogrammetry, smartphone technology, surveying systems, surface morphology

Introduction

Geomorphological approaches for the acquisition of topographic data are experiencing a remarkable technological leap nowadays, with both a substantial increase in the possible frequency of three-dimensional terrain surveying and the ease in which associated methods can be applied. Traditionally, topographic research focused upon constructing digital elevation models (DEMs) using photogrammetric (e.g. Lane *et al.*, 1994; Barker *et al.*, 1997; Chandler, 1999; Lane, 2000; Westaway *et al.*, 2000; Bennett *et al.*, 2012) and differential global positioning system (dGPS) (e.g. Fix and Burt, 1995; Brasington *et al.*, 2000; Young, 2012) data. More recently, both airborne and terrestrial laser scanner have been widely employed to collect very high-quality and high resolution data (Heritage and Hetherington, 2007; Alho *et al.*, 2009; Hodge *et al.*, 2009a, 2009b;

Schaefer and Inkpen, 2010). However, most of these techniques still require expensive equipment and specialized user expertise to process data and improve its quality. In contrast, the development of Structure from Motion (SfM) methods provides the opportunity for very low-cost three-dimensional data acquisition with strongly reduced user supervision and required expertise. The ability to extract high resolution and accurate spatial data using cheap consumer grade digital cameras appears truly remarkable and SfM photogrammetry could answer a range of new research questions.

As in traditional photogrammetry, SfM photogrammetry employs overlapping images acquired from multiple viewpoints. However, SfM photogrammetry differs from traditional photogrammetric approaches by determining internal camera geometry and camera

position and orientation automatically and without the need for a pre-defined set of “ground control”, visible points at known three-dimensional positions (Westoby *et al.*, 2012). The need for a high degree of overlap to cover the full geometry of the object or scene of interest, gives rise to the name: structure derived from a moving sensor.

Whilst the exact implementation of SfM may vary with how it is coded, the general approach has been outlined by other authors (Westoby *et al.*, 2012; James and Robson, 2012; Fonstad *et al.*, 2013; Micheletti *et al.*, 2014) and only a brief explanation is required here. In essence, multiple views of an object are captured with a digital camera from a range of different positions. A scale invariant feature transform (SIFT) then identifies common feature points across the image set, sufficient to establish the spatial relationships between the original image locations in an arbitrary 3-D coordinate system. A sparse bundle adjustment (e.g. Snavely *et al.*, 2008), needed to transform measured image coordinates into 3-D points covering the area of interest, is used in this process. The result is three-dimensional locations of the feature points in the form of a sparse point cloud in the same local 3-D co-ordinate system. Accurate key point correspondence requires the availability of visually distinct texture appearing in the imagery, which can present a problem with some objects and/or lighting conditions. The sparse point cloud is then intensified using Multi View Stereo (MVS) techniques (e.g. Furukawa and Ponce, 2010; Rothmel *et al.*, 2012). It is the ability of these techniques to generate very high resolution datasets, whilst isolating and removing gross errors, which is now allowing such visually impressive 3-D models to be generated so easily when compared to traditional stereo based DEM generation methods involving “*stereomatching*” (Remondino *et al.*, 2014). Effectively, because of the ease with which sensor distortion can be modelled, all consumer grade digital cameras, including the ubiquitous “*smartphone*”, can acquire valuable geomorphic data (Micheletti *et al.*, 2014). Furthermore, the recent development of low-cost, sometimes free, internet-based processing systems enable the upload, processing and download of the derived 3-D data in just a few minutes, potentially during field data collection. This is in direct contrast

to traditional photogrammetric software, where the user is forced to define and to determine interior and exterior orientation parameters explicitly. Most SfM platforms are now fully automated. The advantage of SfM is that it provides a black-box tool where expert supervision is unnecessary. It may also be a disadvantage in that the user has much less involvement in data quality control and the origins of error in data may not be identifiable.

This paper presents guidelines and a workflow for the application of SfM photogrammetry with a hand-held camera, to help avoid generating such inaccurate datasets. Examples and considerations are taken from a study conducted by Micheletti *et al.* (2014) involving ground-based imagery. Although not discussed formally here, all principles also remain valid for images obtained using other approaches such as with Unmanned Aerial Vehicles (UAVs) or drones.

Photogrammetric heritage

The term Structure-from-Motion has evolved from the machine vision community, specifically for tracking points across sequences of images occupied from different positions (e.g. Spetsakis and Aloimonos, 1999; Boufama *et al.*, 1993; Szeliski and Kang, 1994). SfM owes its existence to innovations and mathematical models developed many generations ago, particularly in photogrammetry. The coplanarity condition, now used to establish the spatial relationship between images, was applied in the 1950 and 1960s for numerical aerial triangulation and mapping from aerial photography (Thompson, 1965). The bundle adjustment, which implements the collinearity condition to establish a mathematically rigorous relationship between image and object, was established later by Brown (1971, 1976), Kenefick *et al.* (1972) and Granshaw (1980). Only perfect metric cameras generate images which are distortion free. However, a “self-calibrating” bundle adjustment (Kenefick *et al.*, 1972; Faig and Moniwa, 1973) can model and estimate additional parameters suitable to represent a wide range of internal distortions associated with consumer grade digital cameras. Unfortunately, much of this important pioneering work necessary to establish both appropriate camera models

(e.g. Patias and Streilein, 1996; Shortis *et al.*, 1998) and appropriate geometry necessary for their accurate recovery (Fraser, 1984; Wester-Ebbinghaus, 1986), and the lessons that come from this work, is often overlooked. The freely available and fully automated software packages are flexible and do not assume that the same camera has been used to acquire all images. Each frame may therefore be calibrated individually and inappropriate geometry/image overlaps can generate inaccurate camera models and hence inaccurate datasets.

Method

Software

There are a range of SfM tools now available. PC software, smartphone and web-based apps usually provide similar services for 3-D model generation but differ in the range of post-processing options. Nevertheless, a distinction can be made between solutions that upload images to companies' servers to be processed and provide a download of the results afterwards (e.g. Autodesk 123D Catch, www.123dapp.com/catch or Microsoft Photosynth, www.photosynth.net) and tools that actually process the data on the local machine (e.g. Agisoft PhotoScan, www.agisoft.com, or VisualSFM developed by Wu, 2013, ccwum.me/vsfm). Most tools are available freely, but recently SfM algorithms have been implemented in more conventional and commercial close-range photogrammetry software requiring a subscription (e.g. PhotoModeler release 2014). Whilst, most offer the possibility to download or extract 3-D outputs, some online services still act primarily as web-based 3-D visualization platforms.

SfM services vary in their characteristics and options. Some software resamples images to speed up computations (e.g. Autodesk 123D Catch, currently reduces image resolution to 3 Mega Pixels). Thus, a high-resolution sensor is not usually required but this may limit the precision of generated data. It necessitates careful consideration of the distance between the sensor and the zone of interest, to maintain required resolution. The availability of data used to describe the camera geometry also varies, limiting objective assessment of internal geometry. In contrast, the output is commonly easy to use,

generally 3-D meshes with a basic control upon their density. Point clouds (the nodes of the meshes) can normally be exported in LAS or ASCII format files, allowing further analysis or use in other software.

Data acquisition

SfM involves a process that automatically finds and matches a limited number of common features between images which are then used to establish both interior and exterior orientation parameters. A subsequent procedure then extracts a high resolution and colour-coded point cloud to represent the object. For this reason, the acquisition of imagery with the right characteristics is critical. A range of cameras can be used but a digital SLR camera equipped with fixed focus lens will generate the most accurate data as widely varying zoom settings can cause difficulties (Shortis *et al.*, 2006, Sanz-Ablanedo *et al.*, 2012). Images do not need to be acquired from the same distance or have the same scale (see Figure 1). On the contrary, it is advisable to acquire multi-scale image sets which initially capture the whole site with a few frames before obtaining closer range images to capture the desired detail at the required precision. This is particularly important when capturing areas of detail which are physically obscured by other features (i.e. occlusions). The whole set of images is used for feature extraction, so it is fundamental to ensure that the scene is static and that exposures capture the detail required. Flash photography frequently creates inconsistent image textures which can confuse the feature-matching process (Micheletti *et al.*, 2014). The spatial relationship between images is more flexible than traditional photogrammetric image acquisition using stereo-pairs (Chandler, 1999, Remondino *et al.*, 2014). However, it is critical to acquire imagery from as many different spatial positions as possible. The wide range of image directions then creates a dataset with a strong geometry, important to recover both internal camera models and precise, and hence accurate, object coordinates. The exact number of photographs required is dictated on a case-by-case basis and is a function of occlusion, shape complexity and scale. A range between 10 and 100 should be a good starting point for most applications at close (cm to 10s of m) and intermediate (<

1km) scales. Micheletti *et al.* (2014) demonstrated that increasing the number of images produces denser meshes and improves model accuracy. More significantly, this investigation showed how larger datasets help to remove outliers when the number of images is already sufficient for a good representation of the surface of interest. Hence, very large datasets are not always necessary as even small image sets are able to provide outputs of very satisfying quality, provided image geometry remains strong throughout the area of interest. An important practical constraint is computer memory and the associated time users are willing to wait for results.

Transparent, reflective or homogeneous surfaces present difficulties because incorrect features can be linked during the automatic feature-matching process (Autodesk, 2014). Finally, and of importance to many geomorphic studies, the post-registration procedures combined with a clear idea of what these quantitative data are to be used for, must be planned in advance (see detailed section below).

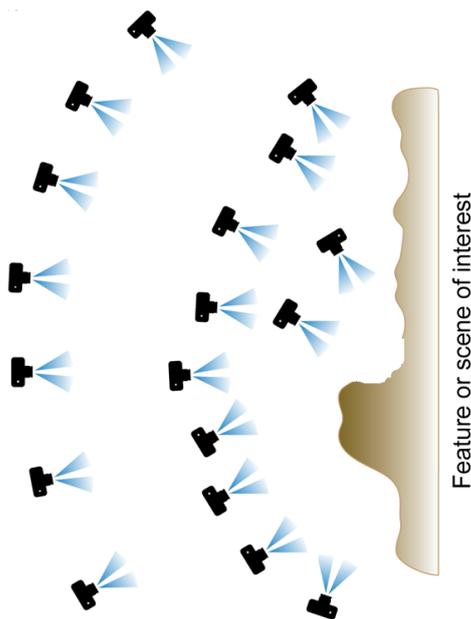


Figure 1: Example of imagery acquisition. Structure from Motion photogrammetry requires multiple photographs with large overlap collected from different positions and directions.

Guidelines and tips for imagery acquisition are often provided with tools (e.g. by Autodesk at www.123dapp.com). A summary

of key points has been provided by Micheletti *et al.*, (2014) and includes:

1. Plan camera survey and registration or scaling method in advance.
2. Capture the whole subject first, and then the detail, ensuring that occlusions are captured adequately (see item 3).
3. Ensure appropriate coverage. Basic principle: every point on the subject must appear on at least three images acquired from spatially different locations.
4. Static scene.
5. Consistent light.
6. Avoid overexposed and under-exposed images.
7. Avoid blurred images – normally arising from slow shutter speed and/or camera movement.
8. Avoid transparent, reflective or homogeneous surfaces.

As for sensors, SfM applications allow a wide range of surveying platforms options for camera deployment. Again, the best choice varies on a case-by-case basis, depending on object of interest and scale. Usually, hand-held devices and tripod-based terrestrial imagery are employed for small landforms. Larger scenes are nowadays mostly surveyed using small-scale UAVs (including multi-copters and fixed-wing drones, e.g. Ryan *et al.*, 2015). These platforms are becoming more popular amongst academics and industrial surveyors due to their increasing affordability. Their clear advantage is the possibility of placing the sensor in locations that would otherwise be difficult to capture with hand-held sensors. Nevertheless, the use of such platforms can create weak image geometry, poor camera models and hence low accuracy data (see below).

Post-processing and possible error sources

In contrast to traditional photogrammetry, SfM does not explicitly require use of ground control points (GCPs), clearly identifiable locations with known or assumed real-world coordinates. Consequently, the resulting

mesh is neither scaled, nor aligned to local gravity, and is therefore unrelated to an established external coordinate system. For many applications this is not a problem, particularly if just a 3-D visualisation or a simple and relative spatial record is all that is required. More usefully, the introduction of a known distance in the scene can define a real world scale.

Typically in geomorphology there is a need to quantify true morphological change by conducting repeat surveys (see Williams, 2012 for details about DEMs of Difference and change detection). If a consistent coordinate system is not defined through time, then detected change may result from a change of coordinate system as opposed to any actual geomorphological process. If the generated model needs to be computed in real-world coordinates or co-registered with either existing or future datasets, a transformation needs to be determined and applied. This is often done using rotation, translation and scaling parameters derived using common GCPs. This procedure has been adopted in many studies in geomorphology (e.g. James and Robson, 2012; Westoby *et al.*, 2012; Fonstad *et al.*, 2013; Micheletti *et al.*, 2014). Thus, coded or simple targets are often employed in SfM approaches (e.g. Figure 2), similar to conventional photogrammetry. Measurement of coded targets can be fully automated in some software but manual target measurement remains universal. Targets need to be clearly identifiable in the images, the number and distribution depending upon project characteristics. A minimum of five would be recommended, though more is preferable so that the quality of the transformation can be assessed independently. If the scene is too vast for the use of artificial targets, easily identifiable natural features can be used instead (Dowling *et al.*, 2009; Dandois and Ellis, 2010). However, these studies recognize the presence of uncertainty linked to the transformation. Westoby *et al.* (2012) indicate that error in the co-registration procedure can be linked to the manual identification of common points and consequently impact upon the accuracy of the derived transformation matrix. Despite this, and in the absence of a second dataset, geo-referencing models using GCPs can usually

provide a sufficiently precise registration for many geoscience applications.



Figure 2: Riverbank with targets printed on regular paper; evenly distributed for referencing purposes (Micheletti *et al.*, 2014).

If a second dataset is available, a possible operation to reduce co-registration errors is the application of an Iterative Closest Points (ICP) algorithm. The algorithm iteratively revises a transformation solution in order to minimize the spatial difference between two point clouds (Zhang, 1992), for areas where there has been no change between the dates of image acquisition. It has proven to be an efficient way to ensure that the coordinate system alignment between two point clouds is as close as possible (e.g. James and Robson, 2012; Micheletti *et al.*, 2014). If the aim is to monitor changes in an area or an object through time, it is suggested to isolate stable zones and to apply an ICP procedure (Micheletti *et al.*, 2014). This will improve the quality of change detection in areas where change has occurred.

As in traditional photogrammetric methods, every stage of a 3-D reconstruction using SfM photogrammetry can create significant errors that propagate through to the final product. The reliance upon a “black box” calibration routine to model camera geometry is particularly problematic. Weak image geometry will generate an imprecise, but more importantly, an inaccurate set of parameters to model camera geometry. A conventional block of vertical aerial imagery is geometrically weak and both a calibrated metric camera and abundant ground control points were traditionally required to maintain mapping accuracies as well as defining a coordinate system. Calibrating a camera “in-

situ” using a conventional block of vertical imagery acquired using a UAV is likely to generate inaccurate data. This typically manifests itself in the form of a systematic error surface or “dome” caused by an inaccurate lens model (Wackrow and Chandler, 2008; James and Robson, 2014) which can often be overlooked (e.g. Ouédraogo *et al.*, 2014). One simple recommendation is to strengthen image geometry by obtaining oblique imagery in addition to the vertical dataset acquired for object coverage. This requires particularly careful attention to be given to the design of UAV surveys.

Together, imagery acquisition and output registration remain delicate steps in otherwise highly automated SfM photogrammetry. For this reason, it is important to consider *a priori* the best strategy whilst considering the specific geoscience application and accuracies that are realistic.

Useful tools and alternatives

Commonly, freely available SfM packages have very limited post-processing functions. As a result, it is often necessary to rely on other software for registration or quantitative analyses. Since many packages offer the possibility to export mesh nodes in LAS or ASCII format files, it is not difficult to find appropriate software to read the data and, if needed, interpolate it to facilitate its use. If working on point clouds is desired, the point cloud data management software CloudCompare developed by EDF R&D (<http://progress.edf.com>) is a convenient solution in terms of both cost and performance (the software is freely available at www.danielgm.net/cc). CloudCompare provides basic manual registration tools, an application of the ICP algorithm and also a point cloud – point cloud comparison tool in the form of the chamfer matching algorithm (Barrow *et al.*, 1977). The chamfer matching algorithm returns values of dissimilarity between two datasets in the form of three-dimensional distances computed by associating each point in the compared dataset with its closest point in the reference

data. Furthermore, CloudCompare now supports the M3C2 cloud-to-cloud differencing algorithm (Brodu and Lague, 2012) as well, an alternative to the original cloud-to-cloud approach. Despite being more computationally demanding than elevation differencing of rasterized DEMs, cloud-to-cloud approaches remain a flexible alternative to comparing more complex 3-D datasets.

For point cloud registration, the transformation applied by Westoby *et al.* (2012) and Micheletti *et al.* (2014) uses a Matlab implementation of the Horn’s absolute orientation algorithm (Horn, 1987) called ABSOR (Jacobson, 2009).

Finally, Table 1 summarises the technical aspects of a range of hand-held sensors used to acquire high resolution topographic data.

Examples

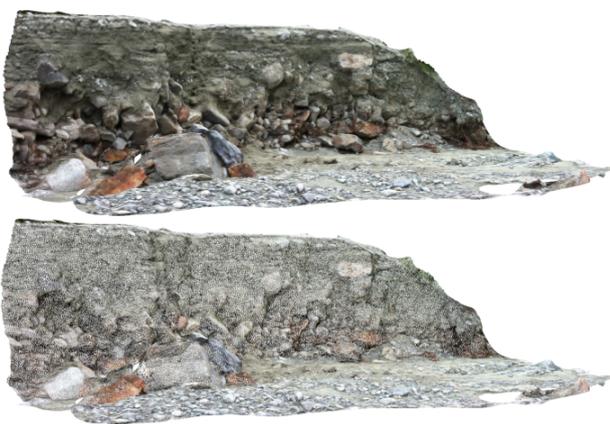
An investigation on the quality of SfM applications for close-range measurement and intermediate measurement scales is presented in Micheletti *et al.* (2014). The freely available, internet-based SfM service offered by Autodesk 123D Catch (www.123dapp.com/catch) was used in conjunction with imagery collected with both a smartphone (Apple iPhone 4) and a digital single-lens reflex camera (Nikon D7000). Extracted models were compared with PhotoModeler implementing traditional stereo based model extraction, whilst using terrestrial laser scanner (TLS) data as a benchmark.

Close-range measurement scale

For a close-range measurement study, the riverbank used (Figure 2) was 10 m long by 1.20 m high and characterized by heterogeneous texture, with grains of different sizes in a coarse-sand and gravel matrix. Using only 13 photographs it was possible to create automatically a three-dimensional point cloud using 123D Catch (Figure 3).

*Table 1: Technical aspects and some possible alternatives for the acquisition of high resolution topographic data in terrain surveys (Micheletti *et al.*, 2014). See also Young (2012).*

Technical aspect	Options	Main characteristics
Survey	Smartphone	Low cost, portable wireless internet access, low quality
	Consumer-grade digital sensor	Low cost, moderate quality
	High-quality digital SLR sensor	High quality, portable, moderate cost, no internet access
	Laser Scanning (TLS and ALS)	High precision, expensive, less portable
Image processing	Internet-based SfM	Free, near real time, fully automatic, lower quality
	Local software SfM	Mostly free and automatic, better quality expected
	Traditional « stereo » photogrammetry	High quality, subscription cost, expert knowledge
	SfM-MVS photogrammetry	As above, but also greater automation and reliability
Co-registration	Scaling	Fast and easy, comparison with other datasets not possible
	Tie points	Comparison with other datasets only, average precision
	Targets + GCP	Any coordinate system, high precision, not always possible
	ICP	Refinement of alignment, needs two co-registered datasets



*Figure 3: Image-covered mesh (above) and its wireframe (below) generated using smartphone imagery and 123D Catch (Micheletti *et al.*, 2014).*

Photogrammetric targets were used to transform the point cloud for comparison purposes using the methodology described

earlier. The comparison was performed in Cloud Compare using TLS data as a benchmark. Median error was 0.0044 and 0.0148 meters for the Nikon D7000 and for the smartphone imagery respectively. These values reduced to 0.0034 and 0.0079 meters after ICP application, proving how registration errors can play a critical role. Despite the downscaling applied to the images prior to data processing, the Nikon D7000 imagery provided optimum results. However, the advantage of a smartphone, in addition to cost, is the possibility of previewing output in the field using wireless communication systems. In any case, sub-centimetre precision could be achieved using SfM photogrammetry for close-range studies, if appropriate care is used during the co-registration procedure.

In the low light conditions prevalent in higher latitudes during the winter, tripod-based

terrestrial imagery can be valuable. A Nikon D7000 digital SLR camera was used to acquire 25 images from a range of distances, necessary to capture a 6 m section of another evolving riverbank in Northamptonshire, UK, on a cloudy day in November. A ground control network was established using a Trimble Total Station, suitable for longer term monitoring and registration purposes. Imagery has been processed using the internet-based, freely available 123D Catch support, but also with the commercial PhotoModeler-SfM/MVS software to perform a comparison (Figure 4). Although the number of points extracted differed (c. 1.9 million for PhotoModeler versus 255,045 for 123D Catch), results demonstrate the high accuracy of the free software. A cloud-to-cloud comparison between the two highlighted just minor differences, median error being 0.004 m for the non-vegetated area.

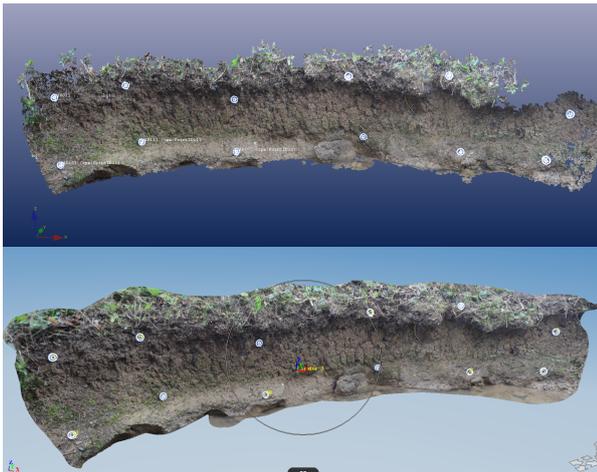


Figure 4: Data extracted using tripod-based terrestrial Nikon D7000 imagery with PhotoModeler SfM/MVS (above) and 123D Catch (below). Targets for registration purposes are distributed throughout the area of interest.

Intermediate measurement scale

An alluvial fan of approximately 87,000 m² was chosen to represent an intermediate case study. A number of complications characterise this experiment, particularly the size of the object and the receding oblique viewpoint from the valley-based imagery. Accordingly, features on the fan are not all captured at the ideal angles/levels of texture. Further, since the object was too large to use photogrammetric targets easily, well-defined

tie-points were used to manually identify and transform the reconstruction to the same coordinate system for comparison purposes. Despite this, results remain convincing; the smartphone-123D Catch models had median errors of 0.5998 and 0.4226 meters prior and after ICP application respectively.

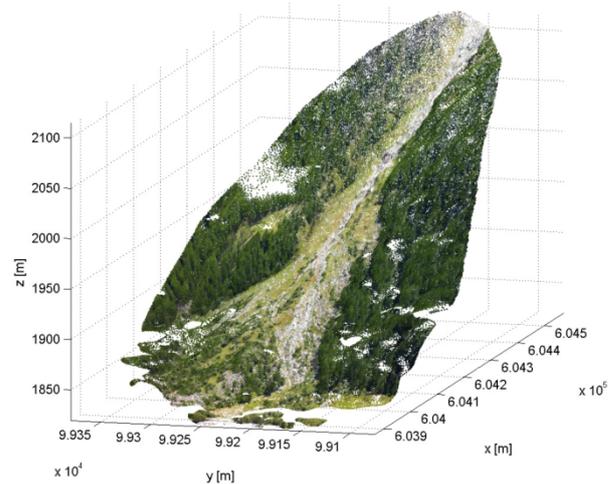


Figure 5: Alluvial fan point cloud generated using smartphone imagery and 123D Catch (Micheletti et al., 2014).

The linear degradation of precision with image scale is well-established in photogrammetry. Moreover, if riverbank and alluvial fan median error are scaled by the mean distance between feature of interest and sensor, accuracies are approximately 1:625; thus, this can be used to roughly anticipate the quality of a basic SfM study based upon freely available software.

If higher accuracies are demanded (1:1,000-1:5,000) then more conventional photogrammetric methods, which implement the SfM process, may remain preferable. Specifically, packages which allow the user to calibrate the camera and then apply appropriate sensor models to image sets will undoubtedly generate data of higher accuracies. This certainly requires a greater understanding of image geometry and general photogrammetric procedures than the freely available, but black box, SfM implementations. SfM photogrammetry can then rival the resolution and quality of surveys conducted using a terrestrial laser scanner (James and Robson, 2012) but with more rapid and convenient field procedures, less occlusions and at a fraction of the hardware costs.

Table 2: Geomorphological sub-disciplines and recent references-studies using SfM photogrammetry.

Geomorphological disciplines	Examples of application
Fluvial Environments	Fonstad <i>et al.</i> , 2013; Javernick <i>et al.</i> , 2014; Micheletti <i>et al.</i> , 2014; Woodget <i>et al.</i> , 2015
Glacial/Periglacial environments	Westoby <i>et al.</i> , 2012; Lucieer <i>et al.</i> , 2014; Ryan <i>et al.</i> , 2015
Hillslope landforms	James and Robson, 2012; Westoby <i>et al.</i> , 2012; Gomez-Gutierrez <i>et al.</i> , 2014; Micheletti <i>et al.</i> , 2014; Stumpf <i>et al.</i> , 2015
Coastal regions	James and Robson 2012; Westoby <i>et al.</i> , 2012
Aeolian landforms	Hugenholtz <i>et al.</i> , 2013
Agricultural watersheds	Ouédraogo <i>et al.</i> , 2014

Conclusion

Fully automated SfM processing of imagery collected by hand-held consumer grade digital cameras provide a valuable spatial object model for various geoscience applications (Table 2), especially at close-range and even with a smartphone. Historical developments in both computer vision and photogrammetry are now receiving particular attention by the major software conglomerates (Google, Microsoft and Autodesk) and it is likely that algorithms will improve radically and rapidly in the near future. Future perspectives are likely to include better sensors and production of very high resolution data, better imagery acquisition alternatives (especially related to the significant potential of UAVs) and enhanced algorithms and processing supports with highly automated routines. With these developments, SfM photogrammetry will become more and more accessible to non-expert users. Nevertheless, recognition of traditional photogrammetric principles remains critical for successful application. Whilst currently available SfM approaches have proven their strengths, even with small

image sets, especially for close range and intermediate scale applications (to 100s of m), the quality of derived data is clearly related to image quality, scale and geometry. If image geometry is weak in any area then inaccurate data can easily be generated, particularly if black box calibration routines are used to determine camera geometry. Geo-referencing can also be the cause of important errors and needs to be investigated to avoid misguided interpretations. As with any technique, experience gained through experimentation is necessary to ensure that expectations are both realistic and fulfilled.

Acknowledgements

Support in the field and access to the riverbank study field site (Figure 4) on the River Nene in Northamptonshire, UK was provided by Prof Steve Rice at Loughborough University.

References

- Alho P, Kukko A, Hyyppa H, Kaartinen H, Hyyppa J, Jaakkola A. 2009. Application of boat-based laser scanning for river survey. *Earth Surface Processes and Landforms* **34**: 1831-1838.
- Autodesk. 2014. 123D Catch. www.123dapp.com/catch.
- Barker R, Dixon L, Hooke J. 1997. Use of terrestrial photogrammetry for monitoring and measuring bank erosion. *Earth Surface Processes and Landforms* **22**: 1217-1227.
- Barrow HG, Tenenbaum JM, Bolles RC, Wolf HC. 1977. Parametric correspondence and chmfer matching: Two new techniques for image matching. *Proceedings of the 5th International Joint Conference on Artificial Intelligence: Cambridge, Massachusetts*, 659-663.
- Bennett GL, Molnar P, Eisenbeiss H, McArdell BW. 2012. Erosional power in the Swiss Alps: characterization of slope failure in the Illgraben. *Earth Surface Processes and Landforms* **37**: 1627-1640.
- Boufama B, Mohr R, Veillon F. 1993. Euclidean constraints on uncalibrated reconstruction. *Proceeding of the Fourth International Conference on Computer Vision, Berlin, Germany*, 466-470.

- Brasington J, Rumsby BT, McVey RA. 2000. Monitoring and modelling morphological change in braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms* **25**: 973-990.
- Brodu N, Lague D, 2012. 3D Terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. *Journal of Photogrammetry and Remote Sensing*, 68: 121-134
- Brown, DC, 1971. Close-Range Camera Calibration. *Photogrammetric Engineering*, 37(8): 855–866.
- Brown, DC. 1976. The bundle adjustment-progress and prospects. *International Archives of Photogrammetry*. 21 B3, Helsinki.
- Chandler, J., 1999. Effective application of automated digital photogrammetry. *Earth Surface Processes and Landforms*, **24**:51–63.
- Dandois JP, Ellis EC. 2010. Remote sensing of vegetation structure using computer vision. *Remote Sensing* **2**: 1157-1176.
- Dowling TI, Read AM, Gallant JC. 2009. Very high resolution DEM acquisition at low cost using a digital camera and free software. *18th World IMACS/MODSIM09 International Congress on Modelling and Simulation, Cairns, Australia* **25**: 973-990.
- Faig W, Moniwa H. 1973. Convergent photos for close-range. *Photogrammetric Engineering and Remote Sensing* **39**: 605-610.
- Fix RE, Burt TP. 1995. Global positioning system: an effective way to map a small area or catchment. *Earth Surface Processes and Landforms* **20**: 817-827.
- Fonstad MA, Dietrich JT, Courville BC, Carbonneau PE. 2013. Topographic structure from motion: a new development in photogrammetric measurements. *Earth Surface Processes and Landforms* **20**: 817-827.
- Fraser CS. 1984. Network design considerations for non-topographic photogrammetry. *Photogrammetric Engineering and Remote Sensing* **50**: 1115-1126.
- Furukawa Y, Ponce J. 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**: 1362–1376.
- Gomez-Gutierrez A, Schnabel S, Berenguer-Sempere F, Lavado-Contador F, Rubio-Delgado J. 2014. Using 3D photo-reconstruction methods to estimate gully headcut erosion. *Catena* **120**: 91-101.
- Granshaw SI. 1980. Bundle adjustment methods in engineering photogrammetry. *The Photogrammetric Record* **10**: 181-207.
- Heritage G, Hetherington, D. 2007. Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms* **32**: 66-74.
- Hodge R, Brasington J, Richards KS. 2009a. Analysing laser-scanned digital terrain models of gravel bed surfaces: linking morphology to sediment transport processes and hydraulics. *Sedimentology* **56**: 2024-2043.
- Hodge R, Brasington J, Richards KS. 2009a. In situ characterization of grain-scale fluvial morphology using terrestrial laser scanning. *Earth Surface Processes and Landforms* **34**: 954-968.
- Horn BKP. 1987. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America* **4**: 629-642.
- Hugenholtz CH, Whitehead K, Brown OW, Barchyn TE, Moorman BJ, LeClair A, Riddell K, Hamilton T. 2013. Geomorphological mapping with small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model. *Geomorphology* **194**: 16-24.
- Jacobson M. 2009. Absolute Orientation – Horn’s method ABSOR. <http://www.mathworks.com/matlabcentral/fileexchange/26186-absolute-orientation-horn-s-method>.
- James MR, Robson S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. *Journal of Geophysical Research* **117**: F03017.
- James MR, Robson S. 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms* **39**: 1413-1420.

- Javernick L, Brasington J, Caruso B. 2014. Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology* **213**: 166-182.
- Kenefick JF, Gyer MS, Harp BF. 1972. Analytical self-calibration. *Photogrammetric Engineering and Remote Sensing* **38**: 1117-1126.
- Lane SN, Richards KS, Chandler JH. 1994. Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms* **19**: 349-368.
- Lane SN. 2000. The measurement of river channel morphology using digital photogrammetry. *Photogrammetric Record* **16**: 937-961.
- Lucieer A, Turner D, King DH, Robinson SA. 2013. Using an Unmanned Aerial Vehicle (UAV) to capture micro-topography of Antarctic moss beds. *International Journal of Applied Earth Observation and Geoinformation* **27**: 53-62.
- Micheletti N, Chandler JH, Lane SN. 2014. Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, DOI: 10.1002/esp.3648.
- Ouédraogo MM, Degré A, Debouche C, Lisein J. 2014. The evaluation of unmanned aerial system-based photogrammetry and terrestrial laser scanning to generate DEMs of agricultural watersheds. *Geomorphology* **214** : 339-355.
- Patias, P. Streilein, A., 1996. Contribution of videogrammetry to the Architectural Restitution. Results Of The CIPA "O.Wagner Pavillion" test. *International Archives of Photogrammetry and Remote Sensing*, 31(b5): 457-462.
- Remondino, F, Spera, MG, Nocerino, E, Menna, F. and Nex, F. 2014. State of the art in high density image matching. *The Photogrammetric Record*, 29: 144-166.
- Rothermel M, Wenzel K, Fritsch D, Haala N. 2012. SURE: photogrammetry surface reconstruction from imagery. *LC3D Workshop*: Berlin, Germany, pp 9.
- Ryan JC, Hubbard AL, Box JE, Todd J, Christoffersen P, Carr JR, Holt TO, Snooke N. 2015. UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. *The Cryosphere* **9**:1-11.
- Sanz-Ablanedo E, Chandler JH, Wackrow R. 2012. Parametrising Internal Camera Geometry with Focusing Distance. *The Photogrammetric Record* **27**: 210-226.
- Schaefer M, Inkpen R. 2010. Towards a protocol for laser scanning of rock surfaces. *Earth Surface Processes and Landforms* **35**: 417-423.
- Shortis MR, Bellman CJ, Robson S, Johnston GJ, Johnson GW. 2006. Stability of zoom and fixed lenses used with digital SLR cameras. *International Archives Photogrammetry Remote Sensing and Spatial Information Sciences* **36**: 285-290.
- Shortis MR, Robson S, Beyer HA. 1998. Principal point behaviour and calibration parameter models for Kodak DCS cameras. *The Photogrammetric Record*, 16: 165-186.
- Snaveley N, Seitz SN, Szeliski R. 2008. Modeling the world from internet photo collections. *International Journal of Computer Vision* **80**: 189-210.
- Spetsakis ME, Aloimonos Y. 1991. A multi-frame approach to visual motion perception. *International Journal of Computer Vision* **6**: 245-255.
- Stumpf A, Malet JP, Allemand P, Pierrot-Deseilligny M, Skupinski G. 2015. Ground-based multi-view photogrammetry for the monitoring of landslide deformation and erosion. *Geomorphology* **231**: 130-145.
- Szeliski R, Kang SB. 1994. Recovering 3-D shape and motion from image streams using nonlinear least squares. *Journal of Visual Communication and Image Representation* **5**: 10-28.
- Thompson, EH. 1965. Review of methods independent model aerial triangulation. *The Photogrammetric Record*, 5: 72-79.
- Wackrow R, Chandler JH. 2008. A convergent image configuration for DEM extraction that minimises the systematic effects caused by an inaccurate lens model. *The Photogrammetric Record*, 23: 6-18.
- Westaway RM, Lane SN, Hicks DM. 2000. The development of an automated correction procedure for digital photogrammetry for the study of wide, shallow, gravel-bed rivers.

Earth Surface Processes and Landforms **25**: 209-226.

Wester-Ebbinghaus W. 1986. Analytical camera calibration. *International Archives of Photogrammetry and Remote Sensing* **26**: 77-84.

Westoby M, Brasington J, Glasser NF, Hambrey MJ, Reynolds MJ. 2012. Structure-from Motion photogrammetry: a low-cost, effective tool for geoscience applications. *Geomorphology* **179**: 300-314.

Williams RD. 2012. Section 2.3.3: DEMs of Difference. In: Clarke LE, Nield JM. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology: London, UK.

Woodget AS, Carbonneau PE, Visser F, Maddock IP. 2015. Quantifying submerged fluvial topography using hyperspatial resolution UAS imagery and structure from motion photogrammetry. *Earth Surface Processes and Landforms* **40**: 47-64.

Wu C. 2013. Towards Linear-time Incremental Structure from Motion, *International Conference on 3D Vision 3DV 2013*.

Young EJ. 2012. Section 2.1.3: dGPS. In: Clarke LE, Nield JM. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology: London, UK.

Zhang Z. 1992. Iterative point matching for registration of free-form curves. *International Journal of Computer Vision* **13**: 119-152.

Creating DEMs from Survey Data: Interpolation Methods and Determination of Accuracy

Andrew D F Bell¹

¹ School of Geography, Archaeology and Palaeoecology, Queens University Belfast, BT7 1NN, (abell19@qub.ac.uk)



ABSTRACT: Digital Elevation Models (DEMs) are becoming increasingly used in modern geomorphological studies. DEMs offer a number of benefits for the characterisation and monitoring of landforms. This paper aims to assess the impacts of interpolation approaches for generation of DEMs. The interpolation approach used for the raw survey data and any subsequent spatial analysis can be affected by errors in the DEMs, leading to potentially inaccurate conclusions being drawn. These errors can either be from the source data or as a consequence of the analysis procedure. Quantification and description of any potential error present for the interpolated datasets is carried out by Root Mean Square Error (RMSE). This is supplemented by analysis into the spatial structure and distribution of the datasets via summary statistics and Q-Q plots. Several of the most common interpolation approaches for the generation of survey data are reviewed here. A case study of applying multiple interpolation approaches to a Terrestrial LiDAR Scan (TLS) dataset is presented. This demonstrates the effect the interpolation approach has on the spatial structure and derivatives of DEMs. Whilst DEMs are used in many geomorphological studies there has to be a tailored, 'site-specific' interpolation approach based on study area, data source, terrain morphology and characteristics.

KEYWORDS: Digital Elevation Models, interpolation approaches, error statistics, terrain derivatives, spatial analysis

Introduction

The use of Digital Elevation Models (DEMs) in the analysis and characterisation of the landscape is beneficial in modern geomorphological studies. DEM spatial analysis not only provides a description of the landscape, but is the foundation of three-dimensional analysis from which other morphological descriptors can be derived (Zhou and Liu, 2004). These descriptors include slope, aspect, curvature (Pirotti and Tarolli, 2010) and roughness (McKean and Roering, 2004; Pollyea and Fariley, 2011). Interpolation of survey data for the creation of DEMs is becoming increasingly frequent in geomorphology (Prokop and Panholzer, 2009; Rayburg, 2009; Aguilar et al. 2010). DEMs are used in studies ranging from landslide analysis (McKean and Roering, 2004; Scheidl et al. 2008), rockfall analysis (Nguyen et al., 2011) fluvial geomorphology (Heritage and Hetherington, 2007) and

landscape characterisation (Glenn et al., 2006). The recent development of progressively higher resolution datasets such as aerial Light Detection and Ranging (LiDAR) systems and Terrestrial LiDAR Systems (TLS) results in the interpolation approaches to the raw data often being overlooked (Brunsdon, 2009). Increased data resolution results in demand for answering questions about data at finer resolutions.

Assessing the causes and propagation of error in DEMs is useful when dealing with large amounts of data (Fisher and Tate, 2006). The propagation of DEM error and the impact on terrain derivatives including slope and aspect is an issue when generating DEMs from survey data (Hunter and Goodchild, 2007). Errors in the base DEM or from the interpolation approach will have a detrimental effect on the terrain derivatives (Kienzle, 2004; Aguilar et al., 2005). Small errors in the generation of the DEM can lead

to inaccurate slope, aspect and curvature derivatives which in turn will lead to inaccurate predictions and conclusions (Januchowski et al., 2010). The interpolation approach used for spatial data analysis can influence the accuracy / quality of the surface produced (Lloyd and Atkinson, 2001; Heritage et al., 2009).

There is often little thought given to the implication that the choice of interpolation approach will have on the generation of the points to a regular DEM grid. This aim of this paper is to assess the impact of interpolation approaches for the generation of DEMs from survey data. Objectives include the creation of DEMs from survey data using multiple interpolation approaches. Summarising and identifying sources of error arising from the interpolation approach is also assessed. The method presented enables the interpolated DEMs to be assessed in terms of the statistical characteristics of error, spatial statistical structure and deviance of distribution as a means to easily understand spatial structure. Providing an insight into the best available interpolation approach when creating DEMs for analysis.

Interpolation methods available for creating DEMs

Accurate interpolation of survey data, is essential for accurate conclusions and validations to be accomplished. Numerous interpolation approaches can be applied to produce DEMs from survey data (Podobnikar, 2005). Interpolation approaches for the purposes of this article will be limited to four readily available interpolation approaches for the generation of DEMs from survey data.

Inverse Distance Weighting (IDW) is an exact interpolator with the predicted values at locations being the same as the observed values (Lloyd 2007). IDW works on a local neighbourhood approach on the assumption that the value at any unsampled point is a weighted average of the values of points within a certain cutoff distance. The weights are inversely proportional to the power of the distance (Burrough and McDonnell, 1998; Mitas and Mitasova, 2005). Advantages of IDW are the easy implementation of the technique within GIS programs and that

interpolations using IDW are not computer RAM intensive to produce.

Radial Basis Functions (RBF) are a group of exact interpolators that use a basic equation dependent on the distance between the interpolated point and the sampling points (Aguilar et al., 2005). RBF utilises splines, hypothetical surfaces that are fitted to some local subset of the data. The analogy often used for splines within RBF is the imitation of a rubber membrane passing through all the data points. The key advantage of the splines interpolated using a RBF from survey data is the amount of control, which can be achieved in the smoothing or tension of the final surface. A spline can be forced to fit through data points or can be smoothed (Lloyd 2007). The advantage of controlling the splines fit to the data is that it is possible to make predictions outside the range of data, so peaks of values that are not necessarily sampled may still be predicted. In addition, smoothing of the spline, as oppose to forcing it through all of the datapoints, resulting in a smoothed surface. This is beneficial where potential local small-scale variation in surfaces can affect the outcome of the generated surface, producing a noisy output. Disadvantages of the RBF are the stiffness and tension of the membrane applied for generation can create large gradients. Processing time is also significantly longer than other approaches due to the number of points being sampled. Additionally, derivatives of the process might cause difficulties in morphological analysis (Mitas and Mitasova, 2005).

Two RBF interpolators are used in this method, Spline with Tension (SPT) and Thin Plate Spline (TPS). SPT differs from TPS due to the tensioning variable that is introduced in the algorithm, which can provide a better fit to the data. Advantages of the SPT include the smoothing parameter whereas TPS exhibits a more rigid interpolation approach from the generation of surfaces. TPS does not offer as much control in terms of fitting of the predicted surface to the datapoints (Lloyd, 2007). TPS advantages include the retention of small-scale features, which is in contrast to weighted averages and trend surfaces (Burrough and McDonnell, 1998).

Ordinary Kriging (OK) is one of the most widely used forms of prediction from

geostatistical analysis, this allows the mean of the values, in this case elevation, to vary and is estimated for each prediction neighbourhood (Lloyd, 2007). The process of kriging can be simplified into; generation of a variogram, fitting a model to variogram, using a model in kriging and finalising the output of kriging. Kriging assumes that the spatial distribution of phenomena can be modeled using a random function. This random function for variation can be split into a deterministic component, representing change over the study area, and a stochastic, or random, component. The random function therefore reflects the uncertainty of spatial variables and parameters. Kriging develops the random function on the basis of the generation of a variogram, a measure of the spatial variability of a particular variable.

The variogram is used to assess the degree to which values differ according to how far apart they are in space (Lloyd, 2007). Lag of the variogram is used to describe distance by which observations are separated. Elements and measures of the variance are estimated by calculating the squared differences between paired observations separated by their lag. The variogram therefore characterizes the degree of difference in values as a function of the distance that they are separated by (Lloyd, 2007; pp. 142). A variogram can then be fitted with a mathematical model used as a tool to estimate how close a measured point is to an interpolated one.

The variogram model is used within kriging as a means of finding appropriate weights for available observations and using the model fitted to the variogram to obtain the predicted values. Kriging is computed using a weighted average assigning greater weights to closer locations, with these weights being computed from the variogram (Brunsdon, 2009). The generated predicted variables within the neighbourhood controlled by the variogram model are used for the interpolation and generation of surfaces (Goovaerts 2002; Lloyd and Atkinson 2002a; Lloyd and Atkinson 2002b; Lloyd and Atkinson 2006). Disadvantages of Kriging are it is a relatively complex process with numerous forms and numerous models that can be applied for the prediction of variables (Lloyd, 2007).

Method for interpolating DEMs and error assessment

Survey data can be recorded and interpreted from monitoring equipment that records a fixed point in space. The workflow for collection of any survey dataset will now be explained following this an example using a collected TLS dataset will be used to illustrate this method.

Generating DEMs from survey data

As described, many interpolation approaches exist for the generation of survey data. For the purpose of this article the four approaches outlined in the previous section will be used, these are popular methods used to generate DEMs from survey data and are readily available in most software packages. Post processing of the raw data is required before interpolation approaches can be applied. Operator based post-processing involves the removal of any erroneous or non-ground point from the dataset. Operator based post processing is not necessary if data are supplied in post-processed format. Post processing software is dependant on the survey data collection method and the manufacturer of survey equipment. This could be from total station, GPS, Remote sensing or LiDAR. This method allows interchangeable software procedures in preparing data for analysis. In general a processed ASCII .txt files, in this example from Leica Cyclone 5.4 (Leica, 2006), are exported and a shapefile is generated within the ArcGIS software (ESRI, 2010).

This method implements the Geostatistical Analyst Extension of ArcGIS 10. Following post processing of the raw data the interpolation approaches are applied and multiple DEMs generated. During interpolation a different number of nearest neighbours (16, 32, 64) are tested to assess the affect this has on the spatial structure and error statistics of the generated DEMs. Cross validation and error summary statistics in the form of the Root Mean Square Error (RMSE) are recorded during this stage. Summary statistical analysis of interpolated surface, along with computation time, are recorded in ArcGIS. Q-Q plots of interpolated surfaces are also generated within the Open Source R Programming Environment (R Project, 2012). Q-Q plots are used as a method for

assessing error of the spatial structure of interpolated surfaces. The analysis scheme is detailed in Figure 1 and will be explained in more details in the following sections.

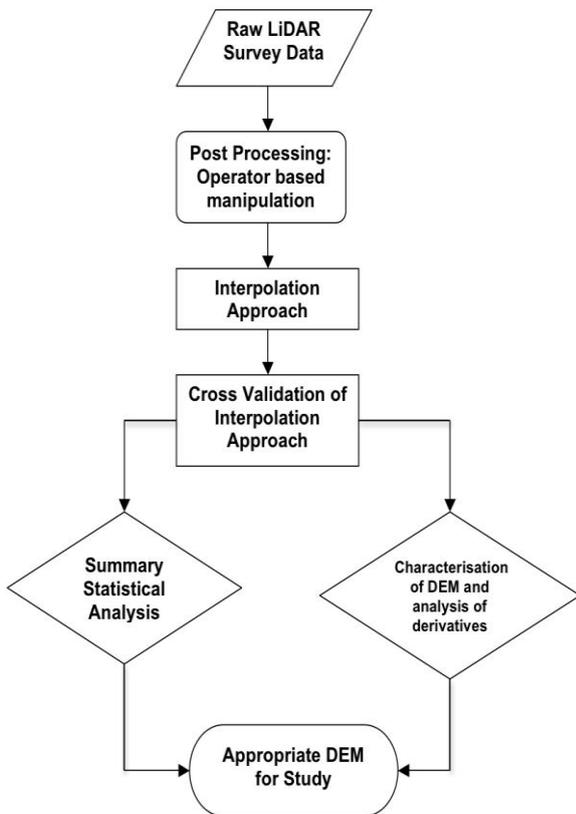


Figure 1: Analysis Scheme for assessment of interpolation approach

Error analysis of DEMs

One of the most common methods to measure DEM quality/accuracy is the Root Mean Square Error (RMSE) (Hunter and Goodchild, 1997; Fisher and Tate, 2006; Aquilar et al., 2005). The RMSE is given by the formula:

$$RMSE = \sqrt{\frac{\sum (z_{DEM} - z_{REF})^2}{n}} \quad (1)$$

Where Z_{DEM} = the measurement of elevation (or derivative) from DEM, Z_{REF} = higher accuracy measurement, from a sample n . RMSE is a form of cross validation of the dataset, which involves removal of one observation using the remaining observations to predict the value of the removed value. This is then returned to the dataset and the next observed value is removed, and then repeated for the entire dataset (Lloyd, 2010).

RMSE as validation measures the difference between observed versus predicted observations of the data providing a representation of the magnitude of error for a particular interpolation approach. The RMSE as a quality measure has a number of different formulas and can include the standard deviation of one surface against another (Keinzle, 2004; Fisher and Tate, 2006). Only formula (1) presented, is used in this example. The RMSE can be seen to give a quick and accurate assessment of the supposed DEM quality / accuracy from interpolation approaches, however, there are a number of drawbacks. RMSE is a global spatial measure, and therefore local spatial characteristics are not assessed. Statistical analysis and structure of the interpolated surfaces provide just as much information of the spatial structure and interpolation approach for the generation of DEMs from survey data (Wise, 2011).

Summary statistical analysis of generated surfaces

In addition to the RMSE, univariate statistics and Q-Q plots are used for an assessment of the interpolated surfaces. Univariate statistical analysis provides descriptive statistics for summarizing particular variables (Lloyd 2010). It includes statistical measures such as the minimum, maximum, standard deviation, skewness and kurtosis of the distribution of the interpolated surface. Univariate statistics have been used with reference to the assessment of interpolation methods by Heritage et al. (2007). One limitation of univariate statistics is that they obscure spatial variation within the study (Lloyd 2007). The ability for analysing distributions of error is based on the assumption that the distribution is normal and stationary. However, some studies suggest spatial autocorrelation statistics may be used for error from DEMs and survey data (Hunter and Goodchild, 1997; Gallay et al., 2010).

The Q-Q plot will be used to compare the structure of how the different interpolation approaches used to generate DEMs from survey data vary from each other in terms of application to the theoretical normal distribution. Therefore, providing an assessment of any sources of potential error when generating the DEMs. The diagnostic Q-Q plot should yield a straight line and any

variance from the straight line of the interpolated surfaces suggests a strong deviation from the normal distribution, therefore error and potential erroneous results (Hohle and Hohle, 2009).

A highly variable output of difference between observed and theoretical quantiles within the interpolation approach, coupled with high variability from the histogram suggests that the interpolated surface may not be the best estimate for the characterisation and generation of DEMs. Analysis of the DEMs and derivatives is carried out on selected DEMs. These are selected depending on the error statistics, generally choosing a range of DEMs generated from different interpolation approaches for assessment. Finally, the most appropriate DEM for the study is selected.

Case study: Statistical analysis of interpolation approaches to LiDAR data

The dataset being used is a TLS survey of Minnis North, a mud-flowslide landform along the A2 Coastal Road, Northern Ireland (Figure 2). Failures due to saturation of the exposed Lias Clays on the hillslope results in periodic failures impacting on the road, blocking off local communities (Smith and Warke, 2001). TLS monitoring is being carried out to assess potential movement and morphological changes on the site. DEMs for each of the interpolation approaches were generated and univariate statistics recorded (Table 1).



Figure 2: Leica HDS3000 Scan station, Terrestrial LiDAR scanning surveying slope morphology, Co. Antrim, Northern Ireland.

The RMSE for TPS are consistently the highest of all the interpolation approaches followed by IDW and OK, with the lowest RMSE recorded for SPT (Table 1). The highest RMSE error also correlates with an overestimation and error in the minimum and maximum heights when generating the DEM. All interpolated surfaces illustrate negative skewness quite close to the mean suggesting the high resolution of the Terrestrial LiDAR dataset enables accurate interpolation of surfaces. TPS performs least favourably in terms of skewness with the distribution being more negatively skewed than other techniques. This is supported by the kurtosis of the distribution with TPS 16 neighbours showing the most peaked distribution around the mean. The standard deviation of the TPS is highest for all interpolation approaches demonstrating greatest variation in values around the mean. A greater spread of values and peak around mean suggests potential sources of error and a reduction in accuracy of the interpolation approach.

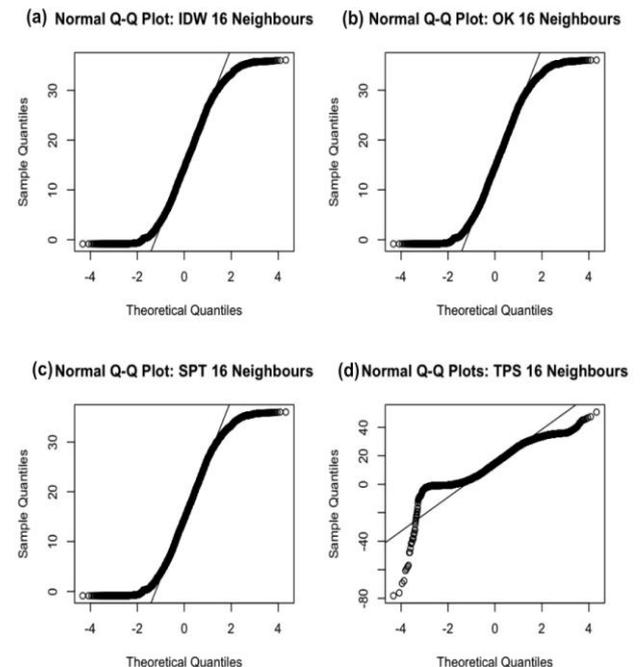


Figure 3: Probability Q-Q Plot of Interpolated Surfaces from LiDAR scan for (a) Inverse Distance Weighting 16 Neighbours (b) Ordinary Kriging 16 Neighbours (c) Spline with Tension with 16 Neighbours (d) Thin Plate Spline with 16 Neighbours

Table 1: Univariate statistical measure for multiple interpolation approaches for a single Terrestrial LiDAR Scan.

Interpolation Approaches; Inverse Distance Weighting (IDW), Ordinary Kriging (OK), Spline with Tension (SPT) and Thin Plate Splines (TPS). Summary statistics; Root Mean Square Error (RMSE), Mean Error (ME), Standard Deviation (St Dev), 1st Quartile (1st Q), 3rd Quartile (3rd Q), Interquartile Range (IQR), Skewness (Skew), Kurtosis (Kurt) and Processing Time (Time(s)).

		RMSE	ME	Mean	Min	Max	St Dev	1st Q	3rd Q	IQR	Skew	Kurt	Time (s)*
IDW	16	0.181	0.0008	16.611	-0.859	36.111	9.835	8.218	24.903	16.685	-0.040	1.882	84
	32	0.179	0.0009	16.614	-0.858	36.082	9.837	8.219	24.907	16.688	-0.040	1.881	95
	64	0.178	0.0011	16.618	-0.858	36.054	9.840	8.225	24.910	16.685	-0.040	1.880	136
OK	16	0.170	0.0002	16.670	-0.856	36.071	9.833	8.290	24.964	16.674	-0.043	1.882	187
	32	0.170	0.0002	16.672	-0.857	36.085	9.835	8.289	24.972	16.683	-0.043	1.881	382
	64	0.169	0.0003	16.611	-0.857	35.922	9.837	8.222	24.912	16.690	-0.041	1.879	1117
SPT	16	0.163	0.0005	16.611	-0.859	36.020	9.835	8.221	24.902	16.681	-0.040	1.882	357
	32	0.162	0.0004	16.612	-0.858	36.012	9.836	8.218	24.910	16.692	-0.041	1.881	789
	64	0.162	0.0001	16.613	-0.857	35.950	9.836	8.221	24.914	16.694	-0.041	1.880	2540
TPS	16	0.866	0.0002	16.670	-139.30	75.663	9.938	8.333	24.957	16.624	-0.173	3.259	257
	32	0.293	0.0002	16.665	-75.327	60.813	9.862	8.318	24.934	16.616	-0.066	2.081	457
	64	0.293	0.0002	16.664	-59.825	57.550	9.824	8.317	24.928	16.612	-0.052	1.969	1184

*Processing time on Macbook Pro, 2.4 GHz Intel Core 2 Duo, 8GB RAM running Bootcamp Partition, Windows 64 bit

The probability Q-Q plots illustrated in Figure 3 for a sample of the interpolated surfaces, illustrate the non-normality of the interpolated surfaces. The TPS shows distinct non-normality with the residuals and outliers causing the plot to become skewed. The other interpolation approaches show similar plots with the tails of the distributions reflecting the summary univariate statistics.

Effect of Interpolation Approach on DEM and Derivatives

The effect of the interpolation approach on the generation of the DEM for this case study is highlighted in Figure 4. DEMs with the highest RMSE and most varying univariate statistics were used to illustrate the changes. Therefore DEMs with 16 neighbours were used in the analysis (Table 1). The DEMs are shown along with a derivative of the DEM, in this case slope.

Figure 4 indicates the interpolated DEMs using IDW, OK and SPT show little variation in their structure and the terrain derivative of slope. This supports the initial univariate statistical analysis. Most notably is that the

TPS derived DEM illustrates large error and artefacts in both the DEM and subsequent terrain descriptor. This is shown by the predominance of lighter areas of higher peaked values. Clumping of these lighter areas are artefacts of the interpolation process. This is the result of an under and over sampling of the raw data creating erroneous interpolated artefacts in the surface. The subsequent second order derivative, slope, illustrates these artefacts, containing false highs in comparison to the other approaches. This should be noted as the artefacts could impact on a study of characterisation of landform morphology producing inaccurate results.

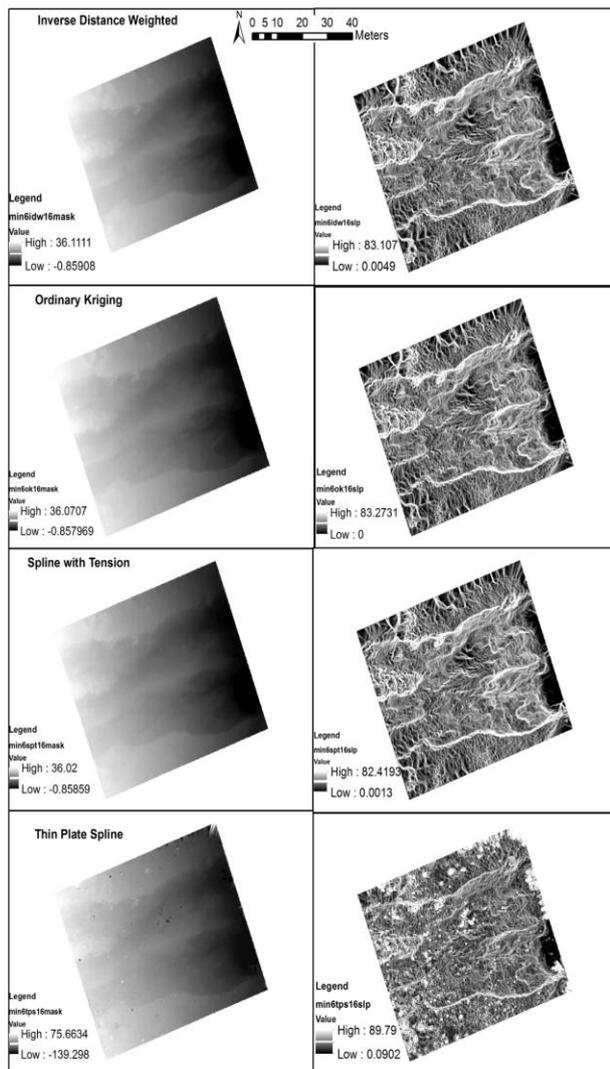


Figure 4: Multiple Interpolated DEMs with 16 neighbours. Left hand column is interpolated DEM and right column is the slope derived surfaces from the DEMs.

Advantages and limitations

Visual assessment of the interpolated DEMs highlights the potential error and the differences between the interpolation approaches. TPS, by the nature of the interpolation approach, makes predictions which are outside the range of the data values. Local small-scale variation in the dataset has produced a noisy output, a limitation of the TPS algorithm. The SPT counteracts this with the smoothing and tension parameters in the algorithm so that the predicted values are a closer fit to the actual raw survey data. This creates the lowest RMSE but not necessarily the best depiction of the raw survey data when analysing the output DEMs and statistics. Smoothing affects the interquartile range of

the data and in some cases the maximum of the data is beyond the maximum values of the data. This can be detrimental when generating DEMs of highly variable terrain, as it can result in certain areas being cut off. Disadvantages of SPT and TPS are that the computation time is considerably longer than the other interpolation approaches, and as the RBFs interpolate outside of the domain of data they therefore have the potential to create artefacts and false highs or lows in the data.

The geostatistical approach using OK for interpolation produces a smooth output which has the second lowest average RMSE for multiple neighbours. The use of the locally varying mean within kriging is beneficial in generating an accurate representation of the elevation surface. However, kriging assumes a normal distribution and estimates the modelled variogram, which may not fit with the data used.

One of the benefits of the IDW approach is that it only predicts in the ranges of the input data. This is illustrated by the summary statistics and DEMs (Table 1; Figure 3 & 4). The neighbourhood becomes smoother as the moving window is increased which is displayed by a reduction in the RMSE. Lloyd and Atkinson (2002) suggest that the use of IDW is acceptable when the sample spacing is small, as it generally is for LiDAR datasets. The cross validation and statistics illustrate IDW as comparable to more complex approaches that generate a slightly smaller RMSE (Lloyd, 2007). IDW saves on computing time compared to more complex interpolation procedures. Disadvantages are the potential for the spikes or a 'duck egg' effect when there is a clustering of data points. 'Duck eggs' are characteristic high or low spots when IDW is giving greater weights to clustering of data sampling locations. Generally this is not an issue when using LiDAR data as it is overcome by the high resolution of data, but must be considered for other survey techniques.

Conclusion

This paper presents a method for determining the accuracy of interpolation approaches for generating DEMs. Examples of common interpolation approaches for generating surfaces from survey data are assessed

through RMSE, spatial structure and Q-Q plots. Key findings indicate RMSE is a good indicator of error when assessing interpolation to DEMs. IDW is seen as a quick and easy exact interpolator and if used correctly, can be an effective means of generating DEMs from survey data. RBF are complex processor intensive functions and what is gained in some areas of the study can be lost in others with the potential for erroneous results to be higher. OK is similar to IDW in the production of the output surfaces and can be used if more control is required in the generation of the DEMs.

Results are dependant on survey data and site-specific conditions. This method provides the framework for the assessment of the interpolation approach discussing the advantages and disadvantages of each. Method presented can be applied to many environments and collected datasets. Researchers should investigate various interpolation approaches for generation of DEMs, choosing an appropriate tailored site-specific approach based on statistical information generated from the presented method.

References

- Aquilar FJ. Aguera F. Aquilar M.A. and Carvajal F. 2005. Effects of Terrain Morphology, Sampling density and Interpolation Methods on Grid DEM Accuracy. *Photogrammetric Engineering and Remote Sensing* **71**(7): 805-816
- Aquilar FJ. Mills JP. Delgado J. Aguilar MA. Negreiros JG. Perez JL. 2010. Modeling vertical error in LiDAR-derived digital elevation models, *ISPRS Journal of Photogrammetry and Remote Sensing* **65**: 103-110
- Brunsdon C. 2009. Geostatistical Analysis of LiDAR Data, Chapter 5 In: Heritage GL. Large ARG. *Laser Scanning for the Environmental Sciences*. Blackwell: Chichester
- Burrough PA. McDonnell RA. 1998. *Principles of Geographic Information Systems*. Oxford University Press: Oxford
- ESRI 2010. *ArcGIS Desktop: Version 10*. Environmental Systems Research Institute Redlands, CA.
- Fisher PF. Tate NJ. 2006. Causes and Consequences of error in digital elevation models. *Progress in Physical Geography*. **30**(4): 467-489
- Gallay M. Lloyd C. McKinley J. 2010. Using geographically weighted regression for analysing elevation error of high-resolution DEMs. *Accuracy 2010 Symposium, July20-23, Leicester, UK*.
- Goovaerts P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology* **228**: 113-129
- Glenn NF. Streuker DR. Chadwick DJ. Tackray GD. Dorsch SJ. 2006. Analysis of LiDAR-derived topographic information for the characterizing and differentiating landslide morphology and activity. *Geomorphology*. **73**:131-148
- Heritage G. Hertherington D. 2007. Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms*. **32**: 66-74
- Heritage GL. Milan DJ. Large ARG. Fuller IC. 2009. Influence of survey strategy and interpolation model on DEM quality. *Geomorphology*. **112**: 334-344
- Hohle J. Hohle M. 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing*. **64**: 398-406
- Hunter GJ. Goodchild MF. 1997. Modeling Uncertainty of Slope and Aspect Estimates Derived from Spatial Databases. *Geographical Analysis*. **29**(1): 35-49
- Kienzle S. 2004. The effect on First Order, Second Order and Compound Terrain Derivatives. *Transactions in GIS*. **8**(1):83-111
- Leica Cyclone 5.4. Leica Geosystems HDS Cyclone 5.4. *Leica Geosystems*. Switzerland. http://hds.leica-geosystems.com/en/Leica-Cyclone_6515.htm (Last Accessed: Feb 2012)

- Lloyd CD. Atkinson PM. 2002a. Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*. **23**(12): 2519-2524
- Lloyd CD. Atkinson PM. 2002b. Non-stationary approaches for mapping terrain and assessing prediction uncertainty. *Transaction in GIS*. **6**(1): 17-30
- Lloyd CD. Atkinson PM. 2006. Deriving ground surface digital elevation models from LiDAR data with geostatistics. *International Journal of Geographical Information Science*. **20**(5): 535-563
- Lloyd CD. 2007. *Local Models for Spatial Data Analysis*. Taylor Francis: Boca Raton
- Lloyd C. 2010. *Spatial Data Analysis: An Introduction for GIS users*. Oxford University Press: Oxford
- Mitas L. Mitasova H. 2005. Spatial Interpolation. Chapter 34 in: Longley PA. Goodchild MF. Maguire DJ. Rhind DW. (eds) *Geographical Information Systems: Principles, Techniques, Management and Applications* 2nd Edition. John Wiley & Sons: New Jersey
- Januchowski SR. Press RL. VanDerWal J. Edwards A. 2010. Characterizing Errors in digital elevation models and estimating the financial costs of accuracy. *International Journal of Geographical Information Science*. **24**(9): 1327-1347
- McKean J. Roering J. 2004. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry. *Geomorphology*. **57**: 331-351
- Nguyen HT. Fernandez-Steegeer TM. Waitr T. Rodrigues D. Azzam R. 2011. Use of terrestrial laser scanning for engineering geological applications on volcanic rock slopes – an example from Madeira island Portugal. *Natural Hazards and Earth Systems Sciences*. **11**: 807-817
- Pirotti F. Tarolli P. 2010. Suitability of LiDAR point density and derived landform curvature maps for channel network extraction. *Hydrological Processes*. **24**:1187-1197
- Podobnikar T. 2005. Production of integrated digital terrain model from multiple datasets of different quality. *International Journal of Geographical Information Science* **19**(1): 69-89
- Pollyea RM. Fairley JP. 2011. Estimating surface roughness of terrestrial laser scan data using orthogonal distance regression. *Geology*. **39**(7): 623-626
- Prokop A. Panholzer H. 2000. Assessing the capability of terrestrial laser scanning for monitoring slow moving landslides. *Natural Hazards and Earth System Sciences*. **9**: 1921-1928
- R – Project, 2012, <http://www.r-project.org/index.html>
Last accessed: 22nd March 2012
- Rayburg S. Thoms M. Neave M. 2009. A comparison of digital elevation models. *Geomorphology*. **106**: 261-270
- Scheidl C. Rickenmann D. Chiari M. 2008. The use of airborne data for the analysis of debris flow events in Switzerland. *Natural Hazards and Earth System Sciences*. **8**: 1113-1127
- Smith B. and Warke P. 2001, *Classic landforms of the Antrim Coast*, Geographical Association: Sheffield
- Wise S. 2011. Cross validation as a means of investigating DEM interpolation error. *Computers and Geosciences*. **37**: 978-991
- Zhou Q. Liu X. 2004. Analysis of errors of derived slope and aspect related to DEM data properties. *Computers and Geosciences*. **30**: 369-378

DEMs of Difference

Richard David Williams¹

¹ Institute of Geography and Earth Science, Aberystwyth University (rvw@aber.ac.uk)



ABSTRACT: A key aspect of geomorphological enquiry is concerned with quantitatively monitoring the development of the Earth's surface, in a diverse set of environments, and at a range of spatial scales and temporal frequencies. A variety of geomatics technologies facilitate the acquisition of multitemporal survey data that can be used to construct Digital Elevation Models (DEMs). The technique of producing a DEM of Difference (DoD) involves quantifying volumetric change between successive topographic surveys. Whilst the essence of the technique is relatively simple, distinguishing between real geomorphic change and survey noise requires appropriate approaches to error analysis to ensure that DoDs are reliable. This is especially important when DEMs have been constructed from fusions of data acquired using different survey or analysis techniques, causing vertical error to be spatially and/or temporally variable across component DEMs. This book chapter reviews example applications of DoDs from across the geomorphological discipline and then focuses upon examining morphological sediment budgeting in fluvial geomorphology. The chapter summarises approaches to error analysis, provides guidance on DEM acquisition, and reviews available software.

KEYWORDS: DEM of Difference (DoD), Deposition, Erosion, Error, Morphological Change, Sediment Budget

Introduction

Quantifying volumetric change is a primary objective for many investigations that consider landform development. Over the last decade, rapid progress in the development of geomatics technologies, and associated processing techniques, has enabled geomorphologists to develop monitoring campaigns that are capable of acquiring accurate Digital Elevation Models (DEMs) at temporal frequencies that are commensurate with rates of landform evolution (e.g. Favalli *et al.*, 2010; Fuller *et al.*, 2010; Fuller *et al.*, 2011; Williams *et al.*, 2011; Carrivick *et al.*, 2012) and at hitherto unprecedented spatial resolution (e.g. Milan *et al.*, 2007; Brasington *et al.*, 2012). This enables insight from morphological change to be coupled directly to process based observations. In addition, new image analysis techniques offer the potential to generate DEMs from archived aerial photos (Lane *et al.*, 2010). This provides the opportunity to extend timescales of enquiry using information from historical

collections and, based on knowledge of forcing events during the monitoring period, inferences can be made about processes that caused change. Whilst DEMs of Difference (DoDs, Wheaton *et al.*, 2010) provide insight into the interaction between process and form, they can also be used to assess the predictions of numerical morphodynamics models.

This chapter first provides an overview of mapping morphological change using the DoD technique. It then summarises pertinent examples of DoDs from a range of geomorphological fields and examines, in detail, the use of the morphological sediment budget method in fluvial geomorphology. Finally, the chapter discusses error analysis and provides guidance on generating DoDs. The focus of the chapter is on producing DoDs in situations that are characterised by elevation changes in one plane (i.e. 2.5 dimensions), which is typical of many Earth surface applications. In some cases change may be measured from a plane that is not

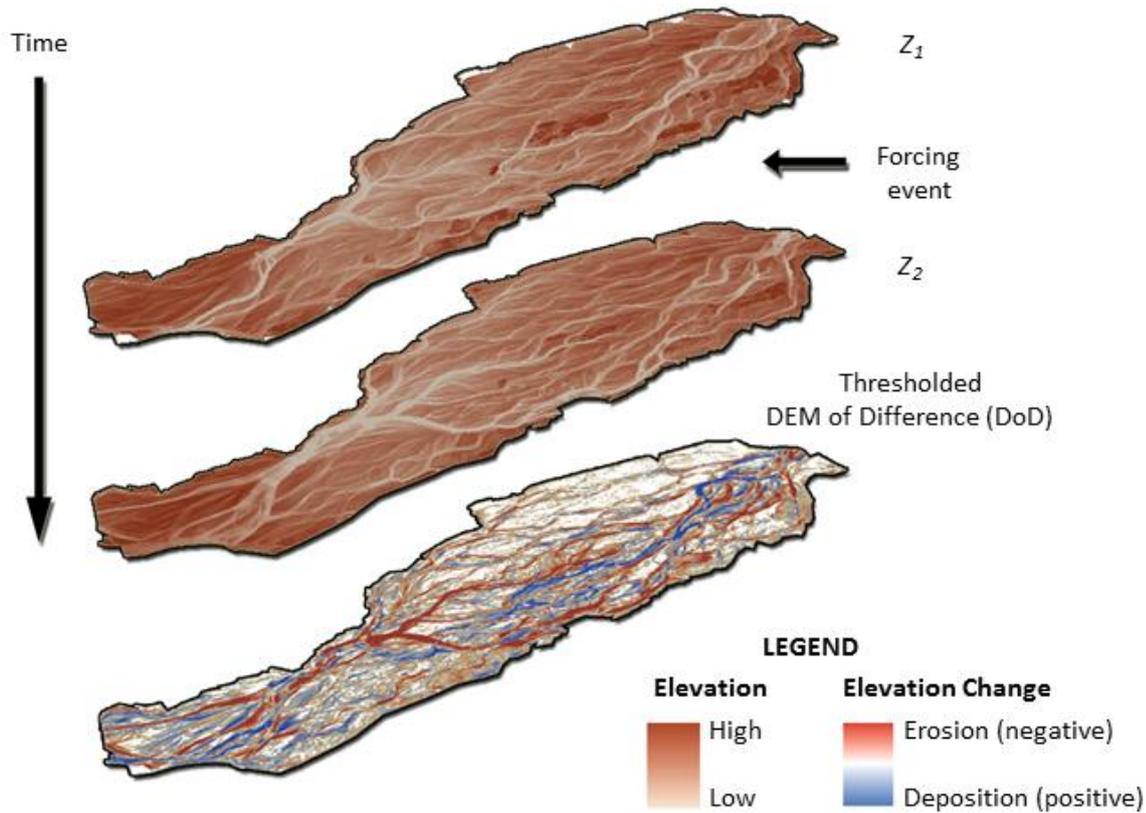


Figure 1: Thresholded DEM of Difference using data from the ReesScan project (Williams *et al.*, 2011). For illustration clarity, DEMs have been detrended by subtracting the streamwise bed slope from each DEM. Approximate flow direction is from right to left.

horizontal, for example when considering river bank retreat (Resop and Hession, 2010; O'Neal and Pizzuto, 2011) or glacial terminus dynamics (Quincey *et al.*, 2011). For 3d cases, where a 2.5d grid approach would bias vertical or overhanging components, Lague *et al.* (Submitted) provide guidance on 3d cloud comparison. The chapter concentrates on discussing survey data acquired in a field setting but the techniques discussed are equally applicable to deriving sequences of multitemporal DEMs from laboratory experiments (e.g. Brasington and Smart, 2003; van Dijk *et al.*, 2012).

Mapping Morphological Change

Geomorphic Change Detection (GCD, James *et al.*, 2012) can either be applied volumetrically, using DEMs (e.g. Rumsby *et al.*, 2008; Wheaton *et al.*, 2010), or in plan, where geomorphological features are delimited from remote sensing imagery or cartography (e.g. Gurnell, 1997; Surian, 1999; Hooke and Yorke, 2011). Here we are

concerned with volumetric GCD where two DEMs that share the same geodetic control (Bannister *et al.*, 1998) are subtracted from one another to reveal a mosaic of morphological change:

$$\delta E = Z_2 - Z_1 \quad (1)$$

where δE is a DEM showing change in elevation, Z_1 is a DEM that was surveyed earlier and Z_2 is a DEM that was surveyed later. Summing the total change across the DoD (δE) quantifies total volumetric change. Negative and positive values on a DoD map show erosion and deposition respectively (Figure 1). Application of equation 1 assumes that both DEMs are true representations of morphological form. Such an assumption is unlikely to be valid when considering field or laboratory observations and, based on Wheaton *et al.* (2010), the DEM elevation value, Z_{DEM} , is likely to contain a vertical error component, δz :

$$Z_{Actual} = Z_{DEM} \pm \delta z \quad (2)$$

where Z_{Actual} is the true elevation value. It is thus necessary to undertake error analysis to ensure that a DoD is reliable. Suitable error analysis techniques are discussed in the “Error Analysis” section below.

Applications

Geomorphology

A variety of geomatics technologies and processing techniques have been applied across the geomorphological discipline to quantify volumetric changes at a range of temporal frequencies and spatial extents (Table 1). These have included terrestrial, airborne and spaceborne photogrammetry, terrestrial laser scanning (TLS), airborne Light Detection and Ranging (LiDAR), Real Time Kinematic Global Positioning System (RTK-GPS) and total station survey. In coastal geomorphology sequences of multitemporal DEMs have been used to estimate beach level changes associated with the passage of a hurricane (Zhang *et al.*, 2005), cliff erosion rates during a year-long monitoring period (James and Robson, 2012) and estuarine bathymetric evolution at a 50-year frequency during a 150 year period (van der Wal *et al.*, 2002). DEMs of Difference have also been widely applied to monitor mass movements (Jaboyedoff *et al.*, 2012) where TLS has emerged as the benchmark technology for monitoring rockfalls (e.g. Rosser *et al.*, 2007; Oppikofer *et al.*, 2009). Landslides have been monitored using a range of geomatics technologies (e.g. Mora *et al.*, 2003; Chen *et al.*, 2006), and debris- and earth-flows have been monitored using airborne LiDAR (e.g. Bull *et al.*, 2010; DeLong *et al.*, 2012). Applications of successive DEMs in volcanology also exemplify the range of temporal frequencies and monitoring durations that can be considered. For example, investigations have quantified changes in lava flow fields every 15 minutes during several days (Favalli *et al.*, 2010), lava dome growth every day for over a year (Major *et al.*, 2009), bi-monthly monitoring of slope evolution during effusive eruption (Baldi *et al.*, 2008) and morphostructural change during several decades (Neri *et al.*, 2008). Sequences of multi-temporal DEMs have also been acquired and differenced in glaciology with applications in glacial, proglacial and

periglacial settings (e.g. Barrant *et al.*, 2009; Fischer *et al.*, 2011; Carrivick *et al.*, 2012).

Overall, the breadth of examples from across the geomorphological discipline illustrate that quantitative measures of morphological change provide a principal analysis technique for many investigations that consider change in Earth surface form. In fluvial geomorphology, considerable attention has been focused upon evaluating survey errors because geomorphic change is often only just detectable above the accuracy of the survey technique being applied. In contrast, investigations in other geomorphological fields tend to focus little attention on quantifying errors. In some cases this is justified since the magnitude of geomorphic change is of a far greater magnitude than survey errors. However, in other cases appropriate error analysis would provide a more rigorous estimate of morphological change. Indeed, of all the investigations listed in Table 1, only Favalli *et al.*'s (2010) investigation of an evolving lava field provides an example of rigorously assessing the reliability of multitemporal DEMs.

Fluvial Geomorphology

DoDs have been widely applied in fluvial geomorphology to (i) infer bedload sediment transport rates; (ii) interpret processes such as channel scour, fill, migration and avulsion; (iii) map the disturbance of ecological habitats; (iv) estimate bed level trends; (v) validate morphological models; and (vi) manage gravel extraction and replenishment schemes. Of these investigation types, the main difference between them is whether their objective is purely to map bed level change or to estimate the rate of bedload sediment transport. Table 2 lists pertinent studies that have applied DoDs in fluvial geomorphology and summarises the survey technologies applied.

Since bedload is the primary determinant of channel morphology (Leopold, 1992; Church, 2006), the morphometric method provides an indirect alternative to the notoriously difficult task of directly sampling and measuring bedload transport rates (Gomez, 1991), so long as the scale of application is sufficiently large (Hicks and Gomez, 2005). Ashmore and Church (1998) provide a salient review of the method, which is based upon a continuity relation for the bedload transport rate:

Table 1: Examples of investigations that have applied DEM differencing in a range of geomorphological settings. See Table 2 for a list of fluvial examples.

Geomorphological field	Application	Geomatics technology	Monitoring duration	Survey frequency	Approximate spatial extent	Reference
Coasts	Estuary change	Bathymetric charts (lead lines and echo sounding)	150 years	Half-century	217 km ²	van der Wal <i>et al.</i> (2002)
	Estuary change induced by earthquakes	Airborne LiDAR	5 months	Start and end of monitoring	5 km ²	Measures <i>et al.</i> (2011)
	Beach changes after a hurricane	Airborne LiDAR	Event	Pre- and post-event	40 km long coastline	Zhang <i>et al.</i> (2005)
	Cliff and gully erosion	Airborne LiDAR	6 years	Start and end of monitoring	77 km long coastline	Young and Ashford (2006)
	Cliff erosion	TLS	16 months	Monthly	0.1 km ²	Rosser <i>et al.</i> (2005)
	Cliff erosion	TLS	1 year	Start and end of monitoring	0.005 km ²	Hobbs <i>et al.</i> (2010)
	Cliff erosion	Oblique terrestrial imagery: SfM and MultiView Stereo	1 year	7 surveys during 1 year	0.05 km long coastline	James <i>et al.</i> (2012)
Fluvial reworking of sediment stores	Talus cone erosion	TLS	3 months	Start and end of monitoring	0.009 km ²	Morche <i>et al.</i> (2008)
	Cut / fill of gully and alluvial fan	Kinematic GPS	32 months	3 - 5 months	0.5 km ²	Fuller and Marden (2010)
Glaciology	Glacier surface elevation change	Aerial photogrammetry	1 year	Start and end of monitoring	6.3 km ²	Hubbard <i>et al.</i> (2000)
	Glacier surface elevation change	Aerial photogrammetry and cartographic data	18 years	Start and end of monitoring	5.5 km ²	Rippin <i>et al.</i> (2003)
	Rockglacier movement	TLS	8 years	1 month - 3 years	0.04 km ²	Avian <i>et al.</i> (2009)
	Glacier surface elevation change	Aerial photogrammetry and airborne LiDAR	2 years	Start and end of monitoring	6 km ²	Barrand <i>et al.</i> (2009)
	Debris covered glacier margins	Airborne LiDAR	4 years	Start and end of monitoring	0.5 km ²	Abermann <i>et al.</i> (2010)
	Permafrost affected bedrock and glacier ice	Aerial photogrammetry and airborne LiDAR	51 years	2 - 22 years	6.5 km ²	Fischer <i>et al.</i> (2011)
	Forefield sediment redistribution	Airborne LiDAR	2 years	Start and end of monitoring	2 km ²	Irvine-Fynn <i>et al.</i> (2011)
	Proglacial and braidplain change	Airborne LiDAR and TLS	2 years	1 day - 1 year	1.5 km ²	Carrivick <i>et al.</i> (2012)
	Glacier surface elevation change	TLS	5 days	Daily	0.05 km ²	Nield <i>et al.</i> (2012)

Table 1 continued

Geomorphological field	Application	Geomatics technology	Monitoring duration	Survey frequency	Approximate spatial extent	Reference
Mass movements	Mudflow	Cartographic data	45 years	1 - 16 years	1.2 km ²	van Westen and Lulie Getahun (2003)
	Landslide	Kinematic GPS	18 months	7 - 11 months	0.04 km ²	Mora <i>et al.</i> (2003)
	Earthquake triggered landslide and river erosion of deposit	Aerial photogrammetry and airborne LiDAR	14 years	1 - 11 years	6 km ²	Chen <i>et al.</i> (2006)
	Rockfall and slope failure (coast)	TLS	32 months	Monthly	0.1 km ²	Rosser <i>et al.</i> (2007)
	Rockfall and slope failure (deglaciated terrain)	TLS	1 year	1 day - 6 months	0.06 km ²	Oppikofer <i>et al.</i> (2008)
	Landslide (deep seated, Tertiary sediments)	Aerial photogrammetry and airborne LiDAR	50 years	6 - 21 years	0.8 km ²	Dewitte <i>et al.</i> (2008)
	Rockslide (fjord)	TLS	2 years	Annual	0.6 km ²	Oppikofer <i>et al.</i> (2009)
	Landslide (slope undercut by river)	TLS	18 months	2 - 6 months	0.01 km ²	Prokop and Panholzer (2009)
	Rockfall from landslide scar	TLS	10 months	2 - 8 months	0.004 km ²	Abellán <i>et al.</i> (2010)
	Debris flow and flood	Airborne LiDAR	Event	Pre- and post-event	0.4 km ²	Bull <i>et al.</i> (2010)
Earthflow (soil and weathered bedrock)	Airborne LiDAR	4 years	Start and end of monitoring	0.06 km ²	DeLong <i>et al.</i> (2012)	
Seismology	Deformation due to surface rupture	Airborne LiDAR	Event	Pre- and post-event	50 km long multi-fault	
Volcanology	Landslide	Aerial photogrammetry	18 years	3 - 10 years	7 km ²	Fabris and Pesci (2005)
	Slope evolution during an eruption	Aerial photogrammetry	4 years	8 days – 14 months	1 km ²	Baldi <i>et al.</i> (2008)
	Summit morphological change due to eruptive processes and deformation	Airborne LiDAR and aerial photogrammetry	21 years	2 – 12 years	7 km ²	Neri <i>et al.</i> (2008)
	Lava dome growth	Oblique terrestrial imagery	17 months	24-48 hours	1 km ²	Major <i>et al.</i> (2009)
	Lava flow dynamics	Airborne LiDAR	2 days	15 minute intervals	28 km ²	Favalli <i>et al.</i> (2010)
	Growth and deformation of a scoria cones	Airborne LiDAR	4 years	Annual	2 km ²	Fornaciai <i>et al.</i> (2010)
	Lahar	Airborne LiDAR	Event	Pre- and post-event	62 km long river	Procter <i>et al.</i> (2010)
	Crater wall collapse	TLS	4 years	17 – 32 months	9 km ²	Pesci <i>et al.</i> (2011)

Table 2: Examples of geomatics technologies applied to undertake morphological budgeting in fluvial geomorphology.

Geomatics technology	Examples
Aerial photographs, bathymetric charts and bank height from longitudinal profile and maps	Popov (1962)
Total station surveyed cross-section	Griffiths (1979; Goff and Ashmore (1994); Martin and Church (1995) Milne and Sear (1997) Brewer and Passmore (2002); Fuller <i>et al.</i> (2002)
Combination of aerial photogrammetry, sonar and total station surveyed cross-section	McLean and Church (1999)
Combination of ground-based photogrammetry and total station distributed point survey	Lane <i>et al.</i> (1994)
Total station distributed point survey	Eaton and Lapointe (2001); Fuller <i>et al.</i> (2003); Rice <i>et al.</i> (2009); Milan <i>et al.</i> (2011)
RTK-GPS distributed point survey	Brasington <i>et al.</i> (2000); Brasington <i>et al.</i> (2003); Fuller <i>et al.</i> (2011); Fuller <i>et al.</i> (2012)
Airborne photogrammetry and image analysis	Brasington <i>et al.</i> (2003); Lane <i>et al.</i> (2003); Lane <i>et al.</i> (2010)
Airborne LiDAR	Lane <i>et al.</i> (2003); Hofle <i>et al.</i> (2009)
TLS	Milan <i>et al.</i> (2007); Resop and Hession (2010); Wheaton <i>et al.</i> (2010); O'Neal and Pizzuto (2011)
TLS (mobile platform)	Alho <i>et al.</i> (2011); Williams <i>et al.</i> (2011)
Non-metric photogrammetry from a pole	Bird <i>et al.</i> (2010)
Aerial photos and measurements of average bed depth change	Carson and Griffiths (1989); Ham and Church (2000)

$$V_o = V_i - (1 - \rho) \frac{\delta S}{\delta t} \quad (3)$$

where V_o and V_i are volumes of sediment flux out of and into the reach respectively, ρ is porosity, δS is change in storage and δt is change in time (Figure 2). Early work on the morphometric method was undertaken by Popov (1962), expanded by Neill (1971; 1987) and the technique has subsequently been developed using a range of geomatics technologies to infer sediment transport rates (e.g. Ferguson and Ashworth, 1992; Goff and Ashmore, 1994; Lane *et al.*, 1995; Eaton and Lapointe, 2001; Martin and Ham, 2005). The major empirical challenges in applying the method are associated with: (i) closing the sediment budget by estimating sediment flux through the reach of interest; and (ii) quantifying error in DoDs. In addition, compensating cycles of cut and fill that occur at a temporal frequency greater than the monitoring frequency have the potential to introduce bias.

To calculate sediment flux through a reach using only morphological information it is necessary to estimate sediment travel distance, as described in Ashmore and Church (1998):

$$Q_b = V_e \frac{L_t}{L_r} t \quad (4)$$

where Q_b is the bulk sediment transport rate, V_e is the volume of erosion, L_t and L_r are the transport and reach distances respectively and t is the time between morphological surveys. Step length can be estimated by measuring distances between the paired centroids of erosion and deposit volumes along the main channel. However, volume matches between pairs can be difficult to establish, and searches may be futile, particularly in complex braided rivers (Ashmore and Church, 1998; Eaton and

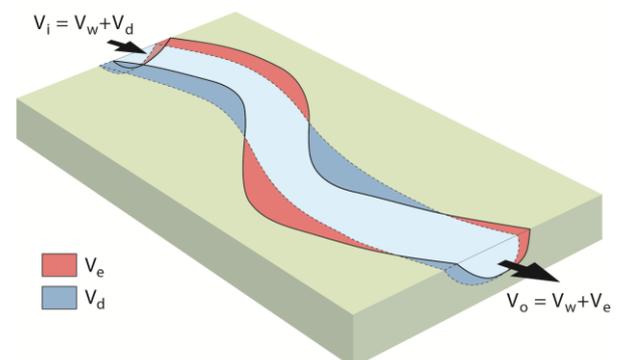


Figure 2: Application of morphometric method in fluvial geomorphology using equation 3. V_i and V_o are the volumetric sediment input and output. V_w is the volumetric washload material. V_e and V_d are the volumetric sediment erosion and deposition. Based on Church (2006).

Lapointe, 2001). Alternatives approaches to estimating sediment flux through the reach of interest include bedload sampling (Lane *et al.*, 1995), and using tracer pebbles to estimate travel distances and sediment mobility patterns (Schwendel *et al.*, 2010).

The advent of remote sensing techniques enables spatially continuous surveys of topography and has largely replaced the prism based method of interpolating cross-sections in a streamwise direction to estimate volumetric change (Griffiths, 1979; Ferguson and Ashworth, 1992; Martin and Church, 1995) or techniques to estimate the aerial extent of bed material depth changes from aerial photos (Ham and Church, 2000). In some situations, however, survey by regular cross-sections remains the primary means of monitoring channel morphology for long-term (annual to decadal) sediment budgeting. Indeed, many unitary authorities commission cross-section surveys to monitor channel topography due to the cost-effective nature of this survey technique. Discussions on the balance between cross-section spacing and accuracy in morphological budgeting can be found in Lane *et al.* (2003), Hicks (2012) and Lindsay and Ashmore (2002).

Error Analysis

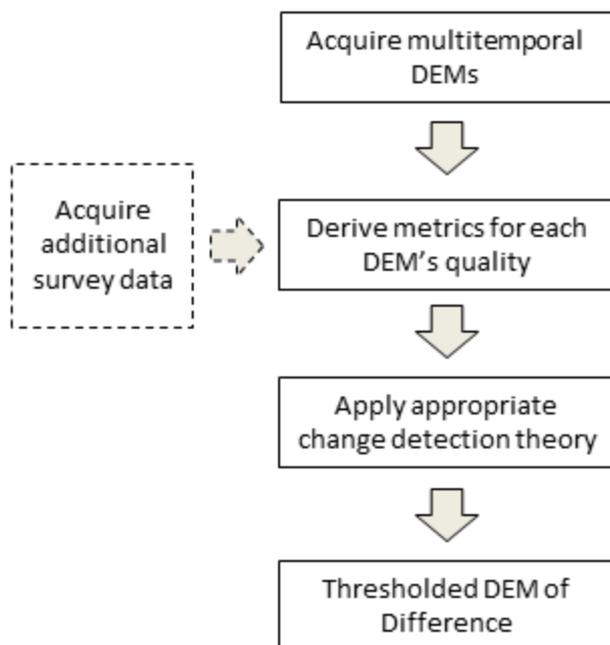


Figure 3: Workflow for DEM differencing with error assessment.

DEMs are unlikely to be exact representations of the Earth's surface due to a variety of uncertainties including those associated with sampling, topographic complexity, geodetic control, survey point precision, processing methods, interpolation and resolution. The production of DoDs can propagate and amplify these uncertainties and it is therefore essential to identify and minimise errors (James *et al.*, 2012). It is useful to consider the ability to detect geomorphic change as a signal, S , to noise, N , ratio:

$$\frac{S}{N} = \frac{V_{GC}}{V_E} \quad (5)$$

Where V_{GC} is variability due to geomorphic change and V_E is variability caused by error (James *et al.*, 2012). This concept highlights the notion that change detection is likely to be more reliable when measured change is of a greater magnitude than associated survey errors. In many cases, however, the magnitude of geomorphic change is similar to the magnitude of uncertainties and appropriate error analyses must therefore be applied to produce reliable DoDs. Moreover, errors are likely to be spatially variable and the signal to noise ratio is likely to vary across an area of interest. The development of appropriate morphological budgeting error analyses has received substantial research attention in fluvial geomorphology (e.g. Brasington *et al.*, 2000; Brasington *et al.*, 2003; Fuller *et al.*, 2003; Lane *et al.*, 2003; Heritage *et al.*, 2009; Wheaton *et al.*, 2010; Milan *et al.*, 2011). In particular, the need to develop reliable DoD techniques has been necessary because channel change is often relatively subtle and of a similar magnitude to DEM uncertainties, especially in the case of deposited gravel sheets (Brasington *et al.*, 2003).

Figure 3 shows a workflow that illustrates the main stages in producing a DoD that is thresholded based on an error assessment. Various approaches can be applied to measure DEM quality (Wheaton *et al.*, 2010) including manufacturer reported instrument precision, repeat observations of control points, bootstrapping experiments, repeat surveys of unchanged areas, fuzzy terrain models and geostatistical techniques. A number of change detection workflows have been proposed. These can be broadly divided into techniques that: (i) apply a

minimum level of detection; (ii) map probabilistic thresholding using a user defined Confidence Interval; (iii) consider the spatial variability of uncertainty from multiple parameters; and (iv) assess the spatial coherence of erosion and deposition. Each of these techniques is summarised below, and the associated advantages and disadvantages of each approach are discussed.

Minimum Level of Detection

The combined error in a DoD, δU_{DOD} , that results from the addition or subtraction of two DEMs, Z_1 and Z_2 , can be estimated from the root sum square of errors:

$$\delta U_{DOD} = \sqrt{\delta z_1^2 + \delta z_2^2} \quad (6)$$

where δz_1 and δz_2 are the errors associated with Z_1 and Z_2 respectively. For example, if δz_1 and δz_2 were both 0.05 m then δU_{DOD} would be 0.07 m. To apply the Minimum Level of Detection (LOD_{Min}), the value of δU_{DOD} is applied as a constant threshold across the DoD. This approach is conservative because only geomorphic change that is greater than the LOD_{Min} is deemed to be reliable. Indeed, testing by Brasington et al. (2003: Figure 8) and Wheaton et al. (2010: Figure 3) indicates that volumetric and areal estimates of morphological change are very sensitive to the LOD_{Min} threshold and information on real geomorphic change is likely to be lost below the threshold. Nevertheless, conservative estimates of geomorphic change have been obtained by defining the LOD_{Min} threshold using grain size information (Schwendel et al., 2010; Fuller et al., 2011; Fuller et al., 2012), and by assessing instrument and registration error (O'Neal and Pizzuto, 2011). A variation on Equation 6 is provided by Fuller et al. (2003: Equation 6) who add an additional term to consider the covariance between the component DEMs.

Probabilistic thresholding using a user defined Confidence Interval

A more rigorous approach to producing a DEM of Difference is to subject the total error to probabilistic thresholding at a user defined confidence interval (Brasington et al., 2000;

Brasington et al., 2003; Lane et al., 2003). By assuming that estimates of δz are approximated by the standard deviation error, σ , and have a normal distribution, equation 6 can be modified to:

$$U_{crit} = t\sqrt{SDE_1^2 + SDE_2^2} \quad (7)$$

where U_{crit} is the critical threshold error, SDE_1 and SDE_2 are the standard deviations of error for Z_1 and Z_2 respectively, and t is the critical t-value for a two-tailed Student's t-distribution for a chosen Confidence Interval:

$$t = \frac{|z_1 - z_2|}{\delta u} \quad (8)$$

where $|z_2 - z_1|$ is the absolute value of the DEM of Difference. For tests at the 1σ or 68% Confidence Interval, $t \geq 1$ and for tests at the 2σ or 95% Confidence Interval, $t \geq 1.96$.

Probabilistic thresholding provides a technique to remove systematic bias through the filtering of elevation changes based on the confidence that detected change is real. Moreover, the user can decide a suitable Confidence Interval for the analysis. However, σ_1 and σ_2 are not necessary uniform across their respective elevation surfaces and the estimation of their values usually requires the use of quality control points that are compared to a surface (Brasington et al., 2000). Applications of probabilistic thresholding have commonly spatially segmented σ_1 and σ_2 based on the influence of riverbed conditions (i.e. dry-dry, wet-wet, wet-dry, dry-wet) on the DEMs that are being differenced (Brasington et al., 2003; Lane et al., 2003; Milan et al., 2007; Bird et al., 2010). Williams et al. (2011) extended this approach to riverbed condition segmentation by assigning spatially variable σ values across dry areas using values of σ estimated from detrended TLS survey data (Figure 4 **Error! Reference source not found.**). Overall, whilst probabilistic approaches to DoD production are likely to produce more reliable estimates of morphological change than approaches based on a minimum LOD, geomorphic changes that are small in elevation but significant in areal extent, such as floodplain sedimentation, may still be misclassified as noise rather than actual morphological change.

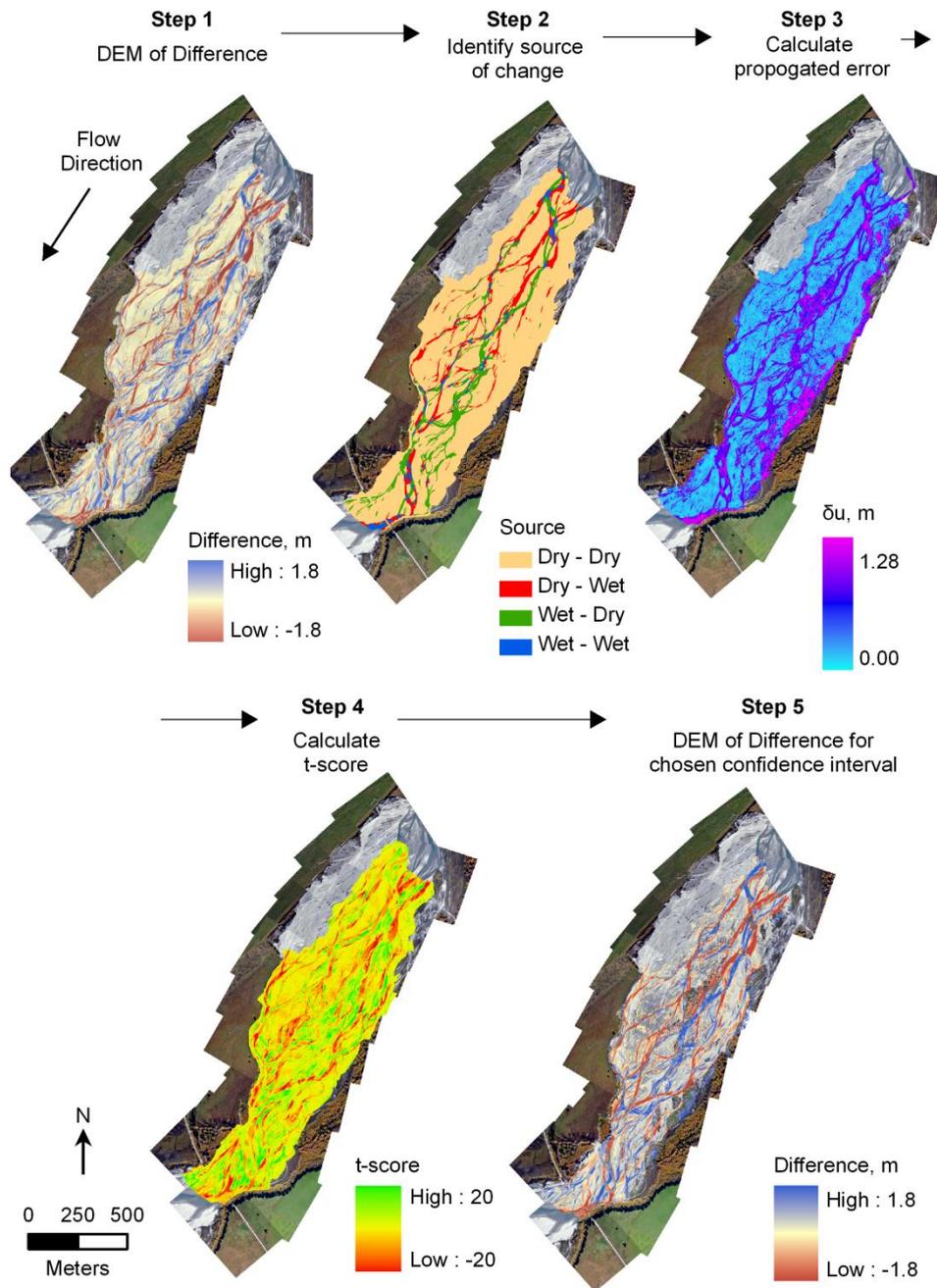


Figure 4: Workflow for producing a DEM of Difference by applying probabilistic thresholding using a user defined Confidence Interval. Reproduced, with permission, from Williams et al. (2011: Figure 8).

Mapping spatial variability of uncertainty from multiple parameters

The incorporation of spatial variability in uncertainty across component DEMs has recently been considered by both Wheaton et al. (2010) and Milan et al. (2011). Whilst Milan et al. (2011) incorporate form

roughness into their error budget, and produce DoDs that are less biased than spatially uniform approaches, Wheaton et al. (2010) use a fuzzy inference system to estimate error from multiple factors that contribute to DEM uncertainty. In brief, Wheaton et al.'s (2010) technique has three stages (Figure 5). First, the factors that

contribute to uncertainty in a DEM grid cell are identified and the magnitude of their error

is rated on a linguistic scale (e.g. low,

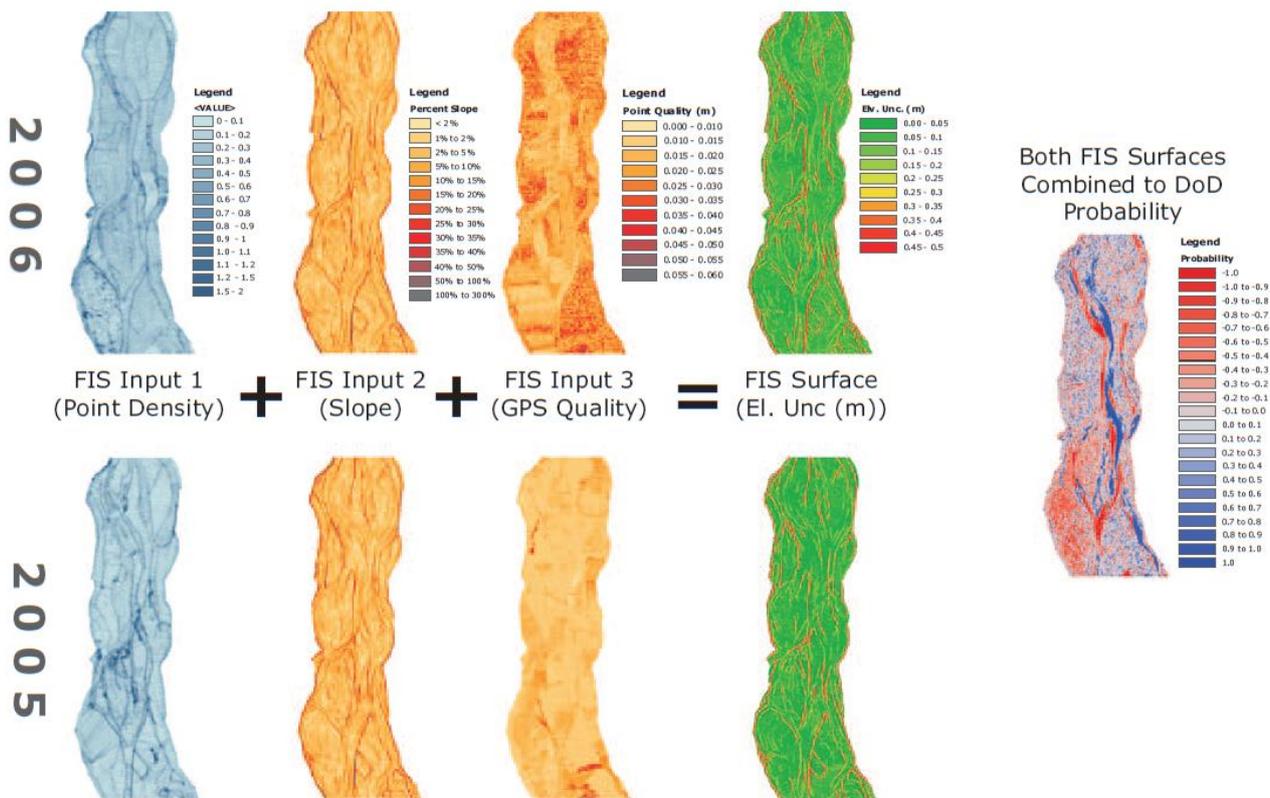


Figure 5. Example of mapping spatial variability of DEM of Difference (DoD) uncertainty using multiple parameters. For each DEM (2005 and 2006) the three inputs (point density, slope and GPS quality) are combined in a three rule fuzzy inference system to produce a defuzzified prediction of elevation uncertainty (where a probability > 0 is deposition and < 0 is erosion). Reproduced, with permission, from Wheaton et al. (2010: Figure 7).

medium, high). Second, a set of rules is defined that considers all combinations of the contributing factors and results in error output, also on a linguistic scale (e.g. low, average, high, extreme). Finally, each DEM grid cell is assigned a weighted uncertainty value. Wheaton et al. (2010) applied this technique to five annual surveys of a 700 m long reach of the River Feshie, Scotland, to develop a fuzzy inference system that incorporated error associated with survey point density, slope and GPS point quality. A MATLAB toolbox is available as a supplement to their paper that enables user configuration of the fuzzy inference system. The fuzzy logic approach enables DoDs to incorporate spatial variability of errors even though it is difficult to define error in precise terms due to the relative dearth of independent survey data on the spatial distribution of error magnitude.

Spatial coherence of erosion and deposition

Cut and fill tend to occur in spatially coherent patterns. Therefore, areas of contiguous and coherent erosion or deposition should have a higher probability of being characterised as such than those that are associated with unstructured patterns. Wheaton et al. (2010) developed a procedure that first defines coherent units of erosion or deposition using a moving window, and then produces a probability that change in a grid cell is true. The authors then calculated a conditional probability that coupled this spatial coherence approach with their technique to map the spatial variability of uncertainty using a fuzzy inference system. The spatial coherence approach could, however, be applied together with other approaches, such as probabilistic thresholding using a user defined Confidence Interval (Section 4.2). Nevertheless, Wheaton et al. (2010) used

data from the River Feshie to show that incorporating spatial coherence produces DEMs of Difference that retain greater volumes of erosion and deposition compared to applying a technique that only considers the probability of change on a cell by cell basis. Accounting for spatial coherence in DoDs that are subjected to error analysis therefore appears to ensure that the rich detail of real geomorphic change is preserved.

Guidance

DEM acquisition and quality

The most important factor for determining the reliability of a DoD is the accuracy of the individual DEMs and their coregistration. The largest errors in DoDs arise in areas with high form and surface roughness, and sparse survey point densities. Therefore the most expedient way to improve the reliability of erosion and deposition estimates is to acquire DEMs that are characterised by levels of precision and accuracy that are commensurate with the magnitude of errors that are acceptable. An investigation's objectives will, of course, also determine the survey frequency and spatial extent that is necessary and these too will impact detection accuracy. The application of novel remote sensing methodologies, such as Structure from Motion (James and Robson, 2012; Westoby *et al.*, in press) and TLS (Heritage and Hetherington, 2007; Williams *et al.*, 2011) are enabling increasingly rich and dense point cloud datasets to be generated at a relatively low cost. However, filtering and classifying dense datasets to generate a DEM is not necessarily straightforward and requires the use of suitable processing techniques (e.g. Brasington *et al.*, 2012; Brodu and Lague, 2012; Rychkov *et al.*, 2012). Moreover, awareness of appropriate DEM generation techniques is integral to producing high quality DEMs (Heritage *et al.*, 2009; Schwendel *et al.*, 2012). Since other chapters in this volume describe key principles and practices associated with various geomatics technologies that are utilised to produce DEMs, further discussion on minimising survey errors is not warranted here. However, it is pertinent to note that any investigation that intends to acquire survey data for generating DEMs should include consideration of how DEM quality will be

measured. This is particularly important when different survey techniques are fused together. For example, in fluvial geomorphology the wetted channel problem (Hicks, 2012) usually requires different geomatics technologies to map exposed and inundated areas of the riverbed (e.g. Brasington and Smart, 2003; Westaway *et al.*, 2003; Williams *et al.*, 2011; Legleiter, 2012).

Software

DoDs can be produced using a variety of Geographic Information System (GIS) and programming software (e.g. Golden Software Surfer, ESRI ArcGIS, Mathworks MATLAB). A very useful utility for ArcGIS is the Geomorphic Change Detection Toolbox (<http://gcd.joewheaton.org/home>). This Toolbox includes procedures to prepare data, undertake change detection using various uncertainty methods, perform batch runs and segment DoDs. Results are output in a variety of forms including GIS grid files, charts and text files. Help on using the Toolbox is available through a series of online video tutorials and help forums. Whilst users should be aware of the principles associated with geomorphic change detection theory before using automated toolboxes, such software reduces the time required to produce DoDs and thus enables attention to be focused upon interpretation of estimated morphological changes.

Error analysis

The development of appropriate error analyses for DoDs has primarily been undertaken within the field of fluvial geomorphology. This interest has been driven by desire to reliably estimate relatively small magnitudes of geomorphic change relative to uncertainty when applying morphological sediment budgeting in gravel-bed river environments. The techniques developed within fluvial geomorphology are, however, transferable to other geomorphological fields. Indeed, applying appropriate error analyses should be *de rigueur* for reliably estimating morphological change across all fields of geomorphological enquiry.

Limitations

A range of geomatics technologies enable the acquisition of precise topographic data during repeat surveys. However, the accuracy of a volumetric estimate of morphological change is limited by the temporal frequency of the successive surveys. For example, in fluvial geomorphology, increasing the temporal interval between two surveys is likely to increase the probability that a DoD will incorporate compensating cycles of scour and fill. This is particularly likely if more than one competent flow event has occurred during the intervening period. The DoD technique thus provides a lower-bound estimate of volumetric change. Ultimately, the utility of a volumetric estimate of change will depend upon the history of forcing events and the characteristics of the environmental setting that is being examined.

Acknowledgements

This chapter was written whilst RDW was in receipt of an Aberystwyth University Postgraduate Research Studentship. Dr Stephen Tooth is thanked for providing comments on an early draft. Comments from two anonymous reviewers also improved the chapter.

References

Abellán A, Calvet J, Vilaplana JM, Blanchard J. 2010. Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. *Geomorphology*, **119**: 162-171. DOI: 10.1016/j.geomorph.2010.03.016

Abermann J, Fischer A, Lambrecht A, Geist T. 2010. On the potential of very high-resolution repeat DEMs in glacial and periglacial environments. *Cryosphere*, **4**: 53-65. DOI: doi:10.5194/tc-4-53-2010

Alho P, Vaaja M, Kukko A, Kasvi E, Kurkela M, Hyyppä J, Hyyppä H, Kaartinen H. 2011. Mobile laser scanning in fluvial geomorphology: mapping and change detection of point bars. *Zeitschrift Fur Geomorphologie*, **55**: 31-50. DOI: 10.1127/0372-8854/2011/0055s2-0044

Ashmore PE, Church MA. 1998. Sediment transport and river morphology: a paradigm

for study. In *Gravel Bed Rivers in the Environment*, Klingeman PC, Beschta RL, Komar BD, Bradley JB (eds). Water Resources Publications: Oregon; 115-139.

Avian M, Kellerer-Pirklbauer A, Bauer A. 2009. LiDAR for monitoring mass movements in permafrost environments at the cirque Hinteres Langtal, Austria, between 2000 and 2008. *Natural Hazards and Earth Systems Science*, **9**: 1087-1094. DOI: 10.5194/nhess-9-1087-2009

Baldi P, Coltelli M, Fabris M, Marsella M, Tommasi P. 2008. High precision photogrammetry for monitoring the evolution of the NW flank of Stromboli volcano during and after the 2002-2003 eruption. *Bulletin of Volcanology*, **70**: 703-715. DOI: 10.1007/s00445-007-0162-1

Bannister A, Raymond S, Baker R. 1998. *Surveying*, Prentice Hall: New Jersey.

Barrand NE, Murray T, James TD, Barr SL, Mills JP. 2009. Optimizing photogrammetric DEMs for glacier volume change assessment using laser-scanning derived ground-control points. *Journal of Glaciology*, **55**: 106-116. DOI: 10.3189/002214309788609001

Bird S, Hogan D, Schwab J. 2010. Photogrammetric monitoring of small streams under a riparian forest canopy. *Earth Surface Processes and Landforms*, **35**: 952-970. DOI: 10.1002/esp.2001

Brasington J, Rumsby BT, Mcvey RA. 2000. Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms*, **25**: 973-990. DOI: 10.1002/1096-9837(200008)25:9<973::AID-ESP111>3.0.CO;2-Y

Brasington J, Langham J, Rumsby B. 2003. Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport. *Geomorphology*, **53**: 299-316. DOI: 10.1016/S0169-555X(02)00320-3

Brasington J, Smart RMA. 2003. Close range digital photogrammetric analysis of experimental drainage basin evolution. *Earth Surface Processes and Landforms*, **28**: 231-247. DOI: 10.1002/esp.480

Brasington J, Vericat D, Rychkov I. 2012. Modelling river bed morphology, roughness and surface sedimentology using high

- resolution terrestrial laser scanning. *Water Resources Research*, **48**: W11519. DOI: 10.1029/2012WR012223
- Brewer PA, Passmore DG. 2002. Sediment budgeting techniques in gravel-bed rivers. In *Sediment Flux to Basins: Causes, Controls and Consequences*, Jones SJ, Frostick LE (eds). Geology Society: London; 97-113. DOI: 10.1144/GSL.SP.2002.191.01.07
- Brodu N, Lague D. 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*, **68**: 121-134. DOI: 10.1016/j.isprsjprs.2012.01.006
- Bull JM, Miller H, Gravley DM, Costello D, Hikuroa DCH, Dix JK. 2010. Assessing debris flows using LIDAR differencing: 18 May 2005 Matata event, New Zealand. *Geomorphology*, **124**: 75-84. DOI: 10.1016/j.geomorph.2010.08.011
- Carrivick JL, Geilhausen M, Warburton J, Dickson NE, Carver SJ, Evans AJ, Brown LE. 2012. Contemporary geomorphological activity throughout the proglacial area of an alpine catchment. *Geomorphology*. DOI: 10.1016/j.geomorph.2012.03.029
- Carson MA, Griffiths GA. 1989. Gravel transport in the braided Waimakariri River - mechanisms, measurements and predictions. *Journal of Hydrology*, **109**: 201-220. DOI: 10.1016/0022-1694(89)90016-4,
- Chen R-F, Chang K-J, Angelier J, Chan Y-C, Deffontaines B, Lee C-T, Lin M-L. 2006. Topographical changes revealed by high-resolution airborne LiDAR data: The 1999 Tsaoling landslide induced by the Chi-Chi earthquake. *Engineering Geology*, **88**: 160-172. DOI: 10.1016/j.enggeo.2006.09.008
- Church M. 2006. Bed material transport and the morphology of alluvial river channels. *Annual Review of Earth and Planetary Sciences*, **34**: 325-354. DOI: 10.1146/annurev.earth.33.092203.122721
- Delong SB, Prentice CS, Hilley GE, Ebert Y. 2012. Multitemporal ALSM change detection, sediment delivery, and process mapping at an active earthflow. *Earth Surface Processes and Landforms*, **37**: 262-272. DOI: 10.1002/esp.2234
- Dewitte O, Jasselette JC, Cornet Y, Van Den Eeckhaut M, Collignon A, Poesen J, Demoulin A. 2008. Tracking landslide displacements by multi-temporal DTMs: A combined aerial stereophotogrammetric and LIDAR approach in western Belgium. *Engineering Geology*, **99**: 11-22. DOI: 10.1016/j.enggeo.2008.02.006
- Eaton BC, Lapointe MF. 2001. Effects of large floods on sediment transport and reach morphology in the cobble-bed Sainte Marguerite River. *Geomorphology*, **40**: 291-309. DOI: 10.1016/s0169-555x(01)00056-3
- Fabris M, Pesci A. 2005. Automated DEM extraction in digital aerial photogrammetry: precision and validation for mass movement monitoring. *Annals of Geophysics*, **48**: 973-988. DOI: 10.4401/ag-3247
- Favalli M, Fornaciai A, Mazzarini F, Harris A, Neri M, Behncke B, Pareschi MT, Tarquini S, Boschi E. 2010. Evolution of an active lava flow field using a multitemporal LIDAR acquisition. *Journal of Geophysical Research-Solid Earth*, **115**: B11203. DOI: 10.1029/2010jb007463
- Ferguson R, Ashworth PJ. 1992. Spatial patterns of bedload transport and channel change in braided and near-braided rivers. In *Dynamics of Gravel-bed Rivers*, Billi P, Hey RD, Thorne CR, Tacconi P (eds). John Wiley & Sons Ltd; 477-492.
- Fischer L, Eisenbeiss H, Käab A, Huggel C, Haeberli W. 2011. Monitoring topographic changes in a periglacial high-mountain face using high-resolution DTMs, Monte Rosa East Face, Italian Alps. *Permafrost and Periglacial Processes*, **22**: 140-152. DOI: 10.1002/ppp.717
- Fornaciai A, Behncke B, Favalli M, Neri M, Tarquini S, Boschi E. 2010. Detecting short-term evolution of Etnean scoria cones: a LIDAR-based approach. *Bulletin of Volcanology*, **72**: 1209-1222. DOI: 10.1007/s00445-010-0394-3
- Fuller IC, Passmore DG, Heritage GL, Large ARG, Brewer PA. 2002. Annual sediment budgets in an unstable gravel-bed river: the River Coquet, northern England. In *Sediment Flux to Basins: Causes, Controls and Consequences*, Jones SJ, Frostick LE (eds). Geology Society: London; 115-131.
- Fuller IC, Large ARG, Charlton ME, Heritage GL, Milan DJ. 2003. Reach-scale sediment

- transfers: an evaluation of two morphological budgeting approaches. *Earth Surface Processes and Landforms*, **28**: 889-903. DOI: 10.1002/esp.1011
- Fuller IC, Marden M. 2010. Rapid channel response to variability in sediment supply: Cutting and filling of the Tarndale Fan, Waipaoa catchment, New Zealand. *Marine Geology*, **270**: 45-54. DOI: 10.1016/j.margeo.2009.10.004
- Fuller IC, Basher L, Marden M, Massey C. 2011. Using Morphological Adjustments to Appraise Sediment Flux. *Journal of Hydrology (New Zealand)*, **50**: 59-79. DOI: 10.1144/GSL.SP.2002.191.01.08
- Fuller IC, Richardson JM, Basher L, Dykes RC, Vale SS. 2012. Responses to river management? Geomorphic change over decadal and annual timescales in two gravel-bed rivers in New Zealand. In *River Channels: Types, Dynamics and Changes*, Molina D (ed). Nova Science: New York; 137-163.
- Goff JR, Ashmore P. 1994. Gravel transport and morphological change in braided Sunwapta River, Alberta, Canada. *Earth Surface Processes and Landforms*, **19**: 195-212. DOI: 10.1002/esp.3290190302
- Gomez B. 1991. Bedload transport. *Earth-Science Reviews*, **31**: 89-132. DOI: 10.1016/0012-8252(91)90017-a
- Griffiths GA. 1979. Recent sedimentation history of the Waimakariri River, New Zealand. *Journal of Hydrology (New Zealand)*, **18**: 6-28.
- Gurnell AM. 1997. Channel change on the River Dee meanders, 1946–1992, from the analysis of air photographs. *Regulated Rivers: Research & Management*, **13**: 13-26.
- Ham DG, Church M. 2000. Bed-material transport estimated from channel morphodynamics: Chilliwack River, British Columbia. *Earth Surface Processes and Landforms*, **25**: 1123-1142. DOI: 10.1002/1096-9837(200009)25:10<1123::AID-ESP122>3.0.CO;2-9
- Heritage GL, Hetherington DJ. 2007. Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms*, **32**: 66-74. DOI: 10.1002/esp.1375
- Heritage GL, Milan DJ, Large ARG, Fuller IC. 2009. Influence of survey strategy and interpolation model on DEM quality. *Geomorphology*, **112**: 334-344. DOI: 10.1016/j.geomorph.2009.06.024
- Hicks DM, Gomez B. 2005. Sediment Transport. In *Tools in Fluvial Geomorphology*, Kondolf GM, Piegay H (eds). Wiley: Chichester, England; 425-461. DOI: 10.1002/0470868333.ch15
- Hicks DM. 2012. Remotely Sensed Topographic Change in Gravel Riverbeds with Flowing Channels. In *Gravel-Bed Rivers*. Wiley; 303-314. DOI: 10.1002/9781119952497.ch23
- Hobbs PRN, Gibson A, Jones L, Pennington C, Jenkins G, Pearson S, Freeborough K. 2010. Monitoring coastal change using terrestrial LiDAR. *Geological Society, London, Special Publications*, **345**: 117-127. DOI: 10.1144/sp345.12
- Hofle B, Vetter M, Pfeifer N, Mandlbürger G, Stotter J. 2009. Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, **34**: 1635-1649. DOI: 10.1002/esp.1853
- Hooke JM, Yorke L. 2011. Channel bar dynamics on multi-decadal timescales in an active meandering river. *Earth Surface Processes and Landforms*, **36**: 1910-1928. DOI: 10.1002/esp.2214
- Hubbard A, Willis I, Sharp M, Mair D, Nienow P, Hubbard B, Blatter H. 2000. Glacier mass-balance determination by remote sensing and high-resolution modelling. *Journal of Glaciology*, **46**: 491-498. DOI: 10.3189/172756500781833016
- Irvine-Fynn TDL, Barrand NE, Porter PR, Hodson AJ, Murray T. 2011. Recent High-Arctic glacial sediment redistribution: A process perspective using airborne lidar. *Geomorphology*, **125**: 27-39. DOI: 10.1016/j.geomorph.2010.08.012
- Jaboyedoff M, Oppikofer T, Abellán A, Derron M-H, Loya A, Metzger R, Pedrazzini A. 2012. Use of LIDAR in landslide investigations: a review. *Natural Hazards*, **61**: 5-28. DOI: 10.1007/s11069-010-9634-2
- James LA, Hodgson ME, Ghoshal S, Latiolais MM. 2012. Geomorphic change detection using historic maps and DEM differencing:

- The temporal dimension of geospatial analysis. *Geomorphology*, **137**: 181-198. DOI: 10.1016/j.geomorph.2010.10.039
- James MR, Robson S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *J. Geophys. Res.*, **117**: F03017. DOI: 10.1029/2011jf002289
- Lague D, Brodu N, Leroux J. Submitted. A new method for high precision 3D deformation measurement of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*.
- Lane SN, Chandler JH, Richards KS. 1994. Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms*, **19**: 349-368. DOI: 10.1002/esp.3290190406
- Lane SN, Richards K, Chandler J. 1995. Morphological estimation of the time-integrated bed load transport rate. *Water Resources Research*, **31**: 761-772. DOI: 10.1029/94WR01726
- Lane SN, Westaway RM, Hicks DM. 2003. Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surface Processes and Landforms*, **28**: 249-271. DOI: 10.1002/esp.483
- Lane SN, Widdison PE, Thomas RE, Ashworth PJ, Best JL, Lunt IA, Smith GHS, Simpson CJ. 2010. Quantification of braided river channel change using archival digital image analysis. *Earth Surface Processes and Landforms*, **35**: 971-985. DOI: 10.1002/esp.2015
- Legleiter CJ. 2012. Remote measurement of river morphology via fusion of LiDAR topography and spectrally based bathymetry. *Earth Surface Processes and Landforms*, **37**: 499-518. DOI: 10.1002/esp.2262
- Leopold LB. 1992. Sediment size that determines channel geometry. In *Dynamics of Gravel-Bed Rivers*, Billi P, Hey RD, Thorne CR, Tacconi P (eds). Wiley: Chichester; 297-311.
- Lindsay JB, Ashmore PE. 2002. The effects of survey frequency on estimates of scour and fill in a braided river model. *Earth Surface Processes and Landforms*, **27**: 27-43. DOI: 10.1002/esp.282
- Major JJ, Dzurisin D, Schilling SP, Poland MP. 2009. Monitoring lava-dome growth during the 2004-2008 Mount St. Helens, Washington, eruption using oblique terrestrial photography. *Earth and Planetary Science Letters*, **286**: 243-254. DOI: 10.1016/j.epsl.2009.06.034
- Martin Y, Church M. 1995. Bed-material transport estimated from channel surveys: Vedder River, British Columbia. *Earth Surface Processes and Landforms*, **20**: 347-361. DOI: 10.1002/esp.3290200405
- Martin Y, Ham D. 2005. Testing bedload transport formulae using morphologic transport estimates and field data: lower Fraser River, British Columbia. *Earth Surface Processes and Landforms*, **30**: 1265-1282. DOI: 10.1002/esp.1200
- Mclean DG, Church M. 1999. Sediment transport along lower Fraser River: 2. Estimates based on the long-term gravel budget. *Water Resources Research*, **35**: 2549-2559. DOI: 10.1029/1999wr900102
- Measures R, Hicks DM, Shankar U, Bind J, Arnold J, Zeldis J. 2011. *Mapping earthquake induced topographical change and liquefaction in the Avon-Heathcote Estuary*, Environment Canterbury Regional Council: Christchurch, New Zealand, 32.
- Milan DJ, Heritage GL, Hetherington D. 2007. Application of a 3D laser scanner in the assessment of erosion and deposition volumes and channel change in a proglacial river. *Earth Surface Processes and Landforms*, **32**: 1657-1674. DOI: 10.1002/esp.1592
- Milan DJ, Heritage GL, Large ARG, Fuller IC. 2011. Filtering spatial error from DEMs: Implications for morphological change estimation. *Geomorphology*, **125**: 160-171. DOI: 10.1016/j.geomorph.2010.09.012
- Milne JA, Sear DA. 1997. Modelling river channel topography using GIS. *International Journal of Geographical Information Science*, **11**: 499-519. DOI: 10.1080/136588197242275
- Mora P, Baldi P, Casula G, Fabris M, Ghirotti M, Mazzini E, Pesci A. 2003. Global Positioning Systems and digital photogrammetry for the monitoring of mass

- movements: application to the Ca' di Malta landslide (northern Apennines, Italy). *Engineering Geology*, **68**: 103-121. DOI: 10.1016/s0013-7952(02)00200-4
- Morche D, Schmidt K-H, Sahling I, Herkommer M, Kutschera J. 2008. Volume changes of Alpine sediment stores in a state of post-event disequilibrium and the implications for downstream hydrology and bed load transport. *Norwegian Journal of Geography*, **62**: 89-101. DOI: 10.1080/00291950802095079
- Mukoyama S. 2011. Estimation of ground deformation caused by the earthquake (M7.2) in Japan, 2008, from the geomorphic image analysis of high resolution LiDAR DEMs. *Journal of Mountain Science*, **8**: 239-245. DOI: 10.1007/s11629-011-2106-7
- Neill CR. 1971. River bed transport related to meander migration rates. *Journal of Waterways, Harbors and Coastal Engineering, American Society of Civil Engineers*, **97**: 783-786.
- Neill CR. 1987. Sediment balance considerations linking long-term transport and channel processes. In *Sediment transport in gravel bed rivers*, Thorne CR, Bathurst JC, Hey RD (eds). John Wiley & Sons: Chichester; 225-240.
- Neri M, Mazzarini F, Tarquini S, Bisson M, Isola I, Behncke B, Pareschi MT. 2008. The changing face of Mount Etna's summit area documented with Lidar technology. *Geophysical Research Letters*, **35**: L09305. DOI: 10.1029/2008gl033740
- Nield JM, Chiverrell RC, Darby SE, Leyland J, Vircavs LH, Jacobs B. 2012. Complex spatial feedbacks of tephra redistribution, ice melt and surface roughness modulate ablation on tephra covered glaciers. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp.3352
- O'neal MA, Pizzuto JE. 2011. The rates and spatial patterns of annual riverbank erosion revealed through terrestrial laser-scanner surveys of the South River, Virginia. *Earth Surface Processes and Landforms*, **36**: 695-701. DOI: 10.1002/esp.2098
- Oppikofer T, Jaboyedoff M, Keusen H-R. 2008. Collapse at the eastern Eiger flank in the Swiss Alps. *Nature Geoscience*, **1**: 531-535. DOI: 10.1038/ngeo258
- Oppikofer T, Jaboyedoff M, Blikra L, Derron MH, Metzger R. 2009. Characterization and monitoring of the Åknes rockslide using terrestrial laser scanning. *Natural Hazards and Earth Systems Science*, **9**: 1003-1019. DOI: 10.5194/nhess-9-1003-2009
- Oskin ME, Arrowsmith JR, Corona AH, Elliott AJ, Fletcher JM, Fielding EJ, Gold PO, Garcia JJG, Hudnut KW, Liu-Zeng J, Teran OJ. 2012. Near-Field Deformation from the El Mayor-Cucapah Earthquake Revealed by Differential LIDAR. *Science*, **335**: 702-705. DOI: 10.1126/science.1213778
- Pesci A, Teza G, Casula G, Loddo F, De Martino P, Dolce M, Obrizzo F, Pingue F. 2011. Multitemporal laser scanner-based observation of the Mt. Vesuvius crater: Characterization of overall geometry and recognition of landslide events. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66**: 327-336. DOI: 10.1016/j.isprsjprs.2010.12.002
- Popov IV. 1962. Application of morphological analysis to the evaluation of the general channel deformations of the River Ob. *Soviet Hydrology*, **3**: 267-324.
- Procter J, Cronin SJ, Fuller IC, Lube G, Manville V. 2010. Quantifying the geomorphic impacts of a lake-breakout lahar, Mount Ruapehu, New Zealand. *Geology*, **38**: 67-70. DOI: 10.1130/g30129.1
- Prokop A, Panholzer H. 2009. Assessing the capability of terrestrial laser scanning for monitoring slow moving landslides. *Nat. Hazards Earth Syst. Sci.*, **9**: 1921-1928. DOI: 10.5194/nhess-9-1921-2009
- Quincey DJ, Bunting P, Williams RD, Herrmann J. 2011. Ice dynamics at the terminus of the Fox Glacier, measured by terrestrial laser scanning. *American Geophysical Union, Fall Meeting 2011*. San Francisco.
- Resop J, Hession W. 2010. Terrestrial Laser Scanning for Monitoring Streambank Retreat: Comparison with Traditional Surveying Techniques. *Journal of Hydraulic Engineering*, **136**: 794-798. DOI: doi:10.1061/(ASCE)HY.1943-7900.0000233
- Rice SP, Church M, Wooldridge CL, Hickin EJ. 2009. Morphology and evolution of bars in a wandering gravel-bed river; lower Fraser river, British Columbia, Canada.

- Sedimentology*, **56**: 709-736. DOI: 10.1111/j.1365-3091.2008.00994.x
- Rippin D, Willis I, Arnold N, Hodson A, Moore J, Kohler J, Björnsson H. 2003. Changes in geometry and subglacial drainage of Midre Lovénbreen, Svalbard, determined from digital elevation models. *Earth Surface Processes and Landforms*, **28**: 273-298. DOI: 10.1002/esp.485
- Rosser N, Lim M, Petley D, Dunning S, Allison R. 2007. Patterns of precursory rockfall prior to slope failure. *J. Geophys. Res.*, **112**: F04014. DOI: 10.1029/2006jf000642
- Rosser NJ, Petley DN, Lim M, Dunning SA, Allison RJ. 2005. Terrestrial laser scanning for monitoring the process of hard rock coastal cliff erosion. *Quarterly Journal of Engineering Geology and Hydrogeology*, **38**: 363-375. DOI: 10.1144/1470-9236/05-008
- Rumsby BT, Brasington J, Langham JA, McLelland SJ, Middleton R, Rollinson G. 2008. Monitoring and modelling particle and reach-scale morphological change in gravel-bed rivers: Applications and challenges. *Geomorphology*, **93**: 40-54. DOI: 10.1016/j.geomorph.2006.12.017
- Rychkov I, Brasington J, Vericat D. 2012. Computational and methodological aspects of terrestrial surface analysis based on point clouds. *Computers & Geosciences*, **42**: 64-70. DOI: 10.1016/j.cageo.2012.02.011
- Schwendel AC, Fuller IC, Death RG. 2010. Morphological dynamics of upland headwater streams in the southern North Island of New Zealand. *New Zealand Geographer*, **66**: 14-32. DOI: 10.1111/j.1745-7939.2010.01170.x
- Schwendel AC, Fuller IC, Death RG. 2012. Assessing DEM interpolation methods for effective representation of upland stream morphology for rapid appraisal of bed stability. *River Research and Applications*, **28**: 567-584. DOI: 10.1002/rra.1475
- Surian N. 1999. Channel changes due to river regulation: the case of the Piave River, Italy. *Earth Surface Processes and Landforms*, **24**: 1135-1151. DOI: 10.1002/(SICI)1096-9837(199911)24:12<1135::AID-ESP40>3.0.CO;2-F
- Van Der Wal D, Pye K, Neal A. 2002. Long-term morphological change in the Ribble Estuary, northwest England. *Marine Geology*, **189**: 249-266. DOI: 10.1016/s0025-3227(02)00476-0
- Van Dijk WM, Van De Lageweg WI, Kleinhans MG. 2012. Experimental meandering river with chute cutoffs. *J. Geophys. Res.*, **117**: F03023. DOI: 10.1029/2011jf002314
- Van Westen CJ, Lulie Getahun F. 2003. Analyzing the evolution of the Tessina landslide using aerial photographs and digital elevation models. *Geomorphology*, **54**: 77-89. DOI: 10.1016/s0169-555x(03)00057-6
- Westaway RM, Lane SN, Hicks DM. 2003. Remote survey of large-scale braided, gravel-bed rivers using digital photogrammetry and image analysis. *International Journal of Remote Sensing*, **24**: 795-815. DOI: 10.1080/01431160110113070
- Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. in press. 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for geoscience applications. *Geomorphology*. DOI: 10.1016/j.geomorph.2012.08.021
- Wheaton JM, Brasington J, Darby SE, Sear DA. 2010. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. *Earth Surface Processes and Landforms*, **35**: 136-156. DOI: 10.1002/esp.1886
- Williams RD, Brasington J, Vericat D, Hicks DM, Labrosse F, Neal M. 2011. Monitoring braided river change using terrestrial laser scanning and optical bathymetric mapping. In *Geomorphological mapping: methods and applications*, Smith M, Paron P, Griffiths J (eds). Elsevier: Amsterdam; 507-532. DOI: 10.1016/B978-0-444-53446-0.00020-3
- Young AP, Ashford SA. 2006. Application of Airborne LIDAR for Seacliff Volumetric Change and Beach-Sediment Budget Contributions. *Journal of Coastal Research*, **22**: 307-318.
- Zhang K, Whitman D, Leatherman S, Robertson W. 2005. Quantification of Beach Changes Caused by Hurricane Floyd along Florida's Atlantic Coast Using Airborne Laser Surveys. *Journal of Coastal Research*, **21**: 44-134. DOI: 10.2112/02057.1

The uDig Spatial Toolbox for hydro-geomorphic analysis

Wuletawu Abera¹, Andrea Antonello², Silvia Franceschi², Giuseppe Formetta³ and Riccardo Rigon¹

¹ Department of Civil Environmental & Mechanical Engineering, University of Trento, Italy (riccardo.rigon@ing.unitn.it)

² Hydrologis s.r.l., Via Siemens 19, Bolzano, Italy (andrea.antonello@hydrologis.com)

³ University of Calabria, Cosenza, Italy (giuseppeform@libero.it)



ABSTRACT: Geographical Information Systems (GIS) are now widely used in hydrology and geomorphology to automate basin, hillslope and stream network analyses. Several commercial GIS software packages are available which provide terrain analysis functionalities for hydrogeomorphology, however, these are often prohibitively expensive. JGrasstools in uDig GIS, is, instead free and Open Source. Recently, uDig was integrated with significant resources for environmental analysis. The Spatial Toolbox for uDig is a specialized toolset for topographical analysis, geomorphometry and hydrology. A large number of tools are included in the toolbox for terrain analysis, river network delineation, and basin topology characterization. They are designed to meet the research needs of academics and scientists, but it remains simple enough in operation to be used for student instruction and professional use. JGrasstools and uDig are developed in Java, which ensures their portability in all operating systems running a Java Virtual Machine. This chapter demonstrates the capabilities of the uDig Spatial Toolbox, which range from the extraction of landform parameters to more advanced DEM manipulations and hydro-geomorphological modelling.

KEYWORDS: Hydrology, geomorphology, GIS, Open Source, catchment analysis, network extraction

Introduction

Since efforts in the late 1980s (Bras *et al.*, 1986, Band 1986, 1993; Grayson *et al.*, 1991) much progress has been made in extending terrain modelling and implementing the mathematical findings of geomorphometry (e.g. Evans *et al.*, 2003) into usable tools (e.g. Wilson and Gallant, 2000, Pike, 2002, Rigon *et al.*, 2006). Furthermore, the availability of Digital Elevation Models (DEMs) has promoted the automatic derivation of river basin features by researchers and practitioners in hydrology and geomorphology.

Given that the tools available were sometimes prohibitively expensive, some

researchers provided their tools as a free product (e.g. Lindsay, 2005; Wood, 2002), but, with few exceptions (e.g. Garbrecht and Martz, 1997; Mitasova and Neteler, 2004, and the suite Sextante (<http://www.sextantegis.com/docs.html>), they provided just the executable of their code, and did not disclose the source code. Since then, with the objective to offer open source alternatives for terrain analysis, various software has been developed. A brief overview of the main software will now be given, this is supplemented by a comparative table of algorithms implemented by each provided in Appendix A (Table A1). LandSerf is an open source tool designed to provide high quality geomorphological visualization and analysis (Wood, 2009), which includes

specific tools for fractal analyses of landscape surfaces. Whitebox Geospatial Analysis Tools, formerly known as TAS, was developed with the objective of providing free and improved visualizations and spatial analyses in GIS (Lindsay, 2005). TauDEM (Terrain Analysis Using Digital Elevation Models) derives from decades of theoretical and applicative work in hydrologic DEM analysis and watershed delineation by Tarboton (e.g. Tarboton, 1997). GRASS GIS is for many purposes similar to the uDig tools presented in this chapter (e.g. Jasiewicz and Metz, 2011). GeoNet derives from recent research by Passalacqua and coworkers on filtering landscape geometries with wavelet tools (e.g. Lashermes and Foufloula-Georgiou, 2007), and on channel initiation (Passalacqua *et al.* 2010a, 2010b).

The uDig (User-friendly Desktop Internet GIS) Spatial toolbox merges the visualization and spatial analysis capabilities commonly found in raster GIS packages with an extensive list of sub-programs specifically designed for research in hydrology and geomorphology. In comparison to many of the other toolkits mentioned, the Spatial Toolbox is a real GIS toolkit (like the one in GRASS, see Table A1) with the advantages of being able to access geographical databases, transform and treat several common geographical data formats, handle and conjointly use vectorial and raster data, and generate the most common data formats in output. Additionally, while most users will find the sole availability of executable code satisfactory, only the full availability of source code internals provides researchers with complete control over the final results of their analyses. For this reasons, the uDig Spatial Toolbox was designed to provide a user-friendly, open source, well-documented, new-generation, GIS for specific applications in hydrology and geomorphology, but also effective for more generic environmental applications.

For historical reasons, the tools in the uDig Spatial Toolbox are also called JGrasstools. They are organized into four toolboxes: Raster processing (RP); Vector processing (VP); HortonMachine (HM); and Others. Most of the geomorphometry analysis tools are in the HortonMachine toolbox. This chapter will concentrate on those functionalities which are

useful to geomorphological analyses contained in the HortonMachine toolbox, but will also touch on those command options that can produce vectorial features of geomorphological entities, without going into a detailed description. Table 1 lists the acronyms used in this book chapter. The Raster processing toolbox has basic tools for raster corrections and operations (that work through the Map Calculator in ArcGIS), whereas the HortonMachine has functionality that ranges from standard analysis of DEMs (such as slope, aspect, curvature) to more specific hydro-geomorphological modeling solutions, which are explained under each category. Raster processing and HortonMachine have some tools which output vectorial features, the Vector Processing toolbox provides many tools for vectorial hydro-geomorphological analysis. Table 2 presents the general functionalities of the four toolboxes.

Table 1: The list of acronyms used in this chapter.

Acronyms	
RP	Raster processing
VP	Vector processing
HM	HortonMachine
OMS	Object Modeling system
Grass	Geographical Resources & Support System
JGrass	Java Geographical Resources & Support System
uDig	User-friendly Desktop Internet GIS
TauDEM	Terrain Analysis Using Digital Elevation Models
OGC	Open GIS Consortium
PNS	Pfafstetter numbering schemes
TAS	Terrain Analysis System
JAMI	Just Another Meteorological Interpolation
NaN	Not a number

Table 2: The list of uDig spatial toolboxes and summary of their general functionalities.

Toolbox	Functionalities
1. Raster processing	Raster data correction and calculations
2. Vector processing	Wide range of Vector data analyses such as vectorizer, buffer zone, line and polygon topological analyses etc.
3. HortonMachine	From simple digital raster terrain analyses to more advanced hydro-geomorphological analyses
4. Others	Design of water supply and sewer systems for urban environments, and other tools

In the next sections, some selected raster processing and HortonMachine tools are described in detail.

In showcasing the virtues of the uDig Spatial Toolbox, the Posina River Basin has been selected as a case study to illustrate the application of some tools. The Posina River Basin is located in the north-western part of the Pre-Alps of Vicenza, between the Astico

Valley and Monte Pasubio. The surface area of the basin at Stancari is 116 km². Geomorphologically, the basin shape is roughly circular and enclosed by a series of mountains with elevations reaching 2000 m and above (Borga *et al.*, 2000). The location and associated DEM of the Posina River Basin are shown in Figure 1.

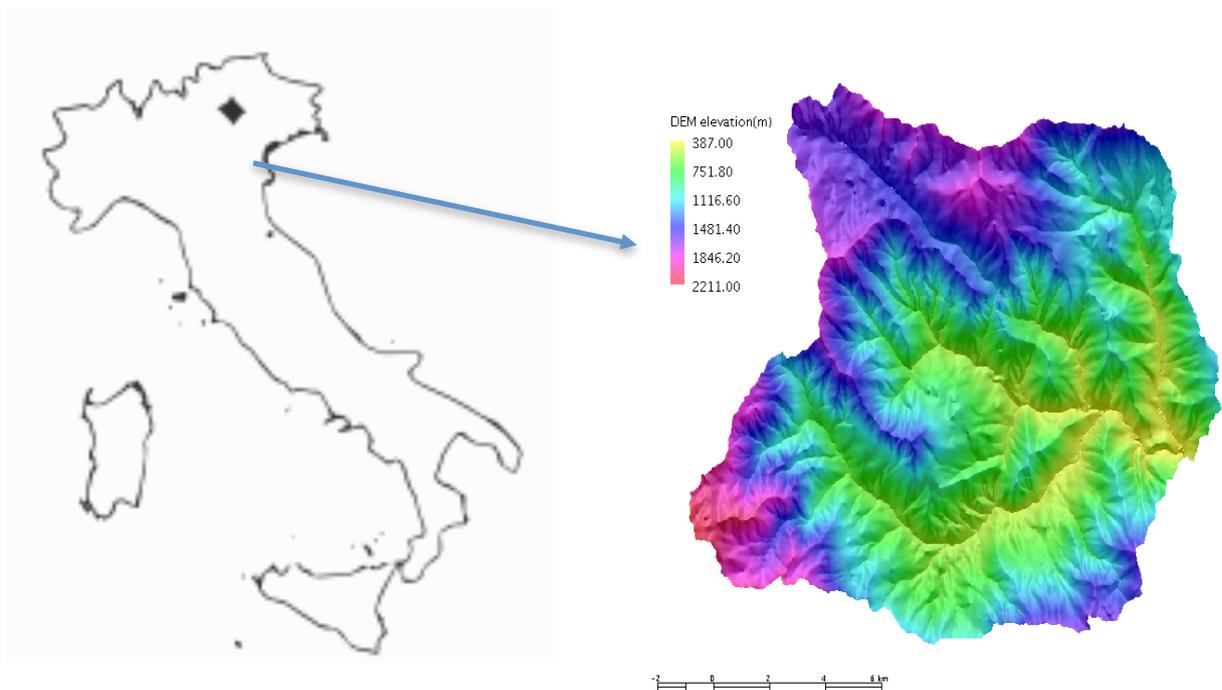


Figure 1: The location and the DEM of Posina river basin, the case study.

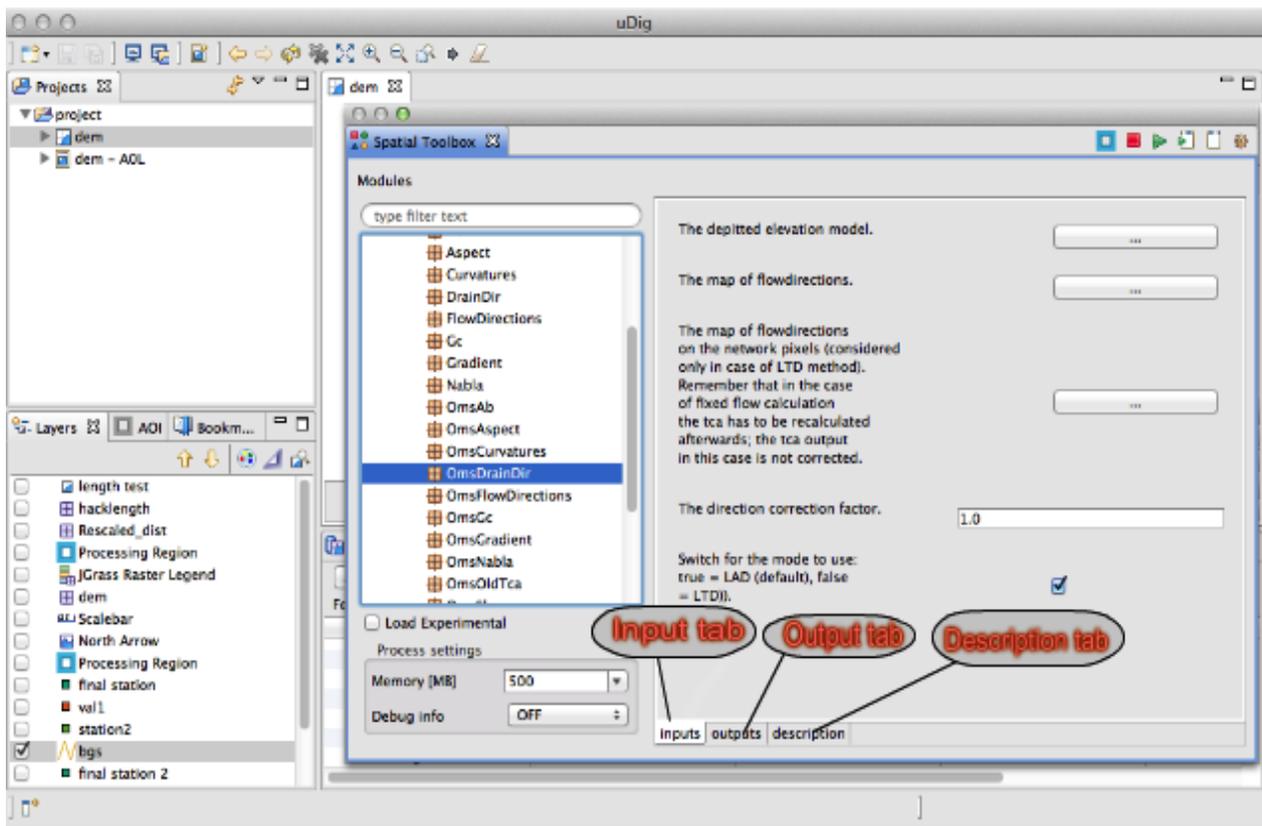


Figure 2: uDig GIS interface with spatial toolbox overlay and the three tabs: inputs, outputs and descriptions tabs at the bottom of the toolbox.

Raster Processing

The **Raster Processing toolbox** contains a set of tools for the preparation of topographic data for hydro-geomorphological applications. Figure 2 shows the typical appearance of the uDig Spatial Toolbox. When the Spatial Toolbox is open, for any tool, there are three tabs that serve to input and the output variables. The third tab contains the help associated with the selected the tool, with a short explanation of what the tool does. Raster processing toolbox contains tools for the quick manipulation of raster maps (the relevant tools are listed in Table 3). For all the tools listed in Table 3, which are useful for modifying and improving DEMs and provide input data for geomorphological and hydrological analyses, information can be found in the Help of the toolbox package.

There are about 30 tools under the RP toolbox that can be implemented. However, there is not the scope to explain all of them here, therefore a little more detail will be provided about a selection of the more interesting tools within the RP toolbox.

Among the list of tools presented in Table 3, a very useful one for advanced raster map manipulation is the raster calculator

(**MapCalc**), which allows complex calculations involving numerical and logical functions. With Mapcalc one can perform the most common mathematical operation on a map, modify map's values, combine maps, multiply, divide one map by another map (i.e. the values contained in one map for the corresponding values of another map), select part of a map, and so on.

Ranglookup, **Rastercorrector**, and **Rasterconverter** provide summary statistics of raster maps. The **Rangelookup** tool is particularly important as it identifies the raster data included between user-defined range values.

RasterSummary tool is useful as it provides basic summary statistics (such as the minimum, maximum, mean, standard deviation, histogram, and the NaN value) of the raster map.

Table 3: Some of the tools available in the Raster Processing toolbox of uDig.

Raster Processing Tools	Functionalities
BobTheBuilder	Builds human artifacts (such as dams) on a raster map
CutOut	Raster masking and cutout with some threshold
KernelDensity	Estimates the kernel density
MapCalc	Performs map algebra on raster map
Mosaic & ImageMosaicCreator	Patches rasters and creates mosaics of shapefiles for images
PointRasterizer & LineRasterizer	Rasterizes vectorial point and line features respectively
CannyEdgedetector	Performs edge detection operations
Profile	Creates profiles over raster maps
RangeLookUp	Reclassifies and assigns values of maps for a given ranges of raster values
RasterConverter	Converts rasters from one format to another
RasterCorrector	Corrects some raster values
RasterDiff	Calculates the difference between two rasters
RasterReprojector	Re-projects maps
RasterResolutionResampler	Resamples the raster map coverage
RasterSummary	Calculates the summary statistics of a raster map
RasterVectorIntersector	Analyzes raster maps within a polygon vector (intersection)

A group of tools for rasterizing vector data includes **BobTheBuilder**, **PointVectorizer** and **LineVectorizer**. **BobTheBuilder** rasterizes human artifacts, such as dams and buildings, which could be useful to include in the raster maps. **PointVectorizer** and **LineVectorizer** rasterize point and line features, such as measurement stations and river channels respectively.

Finally, **SurfaceInterpolator** is useful for interpolating landscape data (such as elevation and temperature), from point measurement to the whole study area. Two surface interpolation algorithms are incorporated in this tool: the Thin Plate Spline (TPS) Interpolator and the Inverse Weight Distance (IWD) Interpolator (e.g. Goovaerts 2000). These methods can be applied to create Digital Terrain Models (DTM) from a set of GPS points or digitized maps, as well as models of other continuous environmental variables, for instance, surface temperature.

HortonMachine Functionality in Geomorphometry

The **HortonMachine toolbox** is organized into seven broad categories of commands: DEM manipulation; Geomorphology indices; Network related analysis; Hydro-geomorphology model tools; Basin related tools; Hillslope related attribute tools; and spatial statistics tools. Each of these will now be outlined in turn, and Tables 4 and 5 present the selection of tools useful for hydro-geomorphological applications.

DEM manipulation toolbox

The **DEM manipulation** tools contain subprograms used for preparing DEMs for analysis. These subprograms include routines to remove flats, spikes, and depressions from DEMs (**pitfiller**), to extract streams (**ExtractNetwork**), to extract sub-

Table 4: Some of the tools for DEM manipulations, geomorphology, hydro-geomorphology and statistics available in the HortonMachine toolbox (for hydro-geomorphological terminology, please refer at <http://www.physicalgeography.net/glossary.html>).

HortonMachine Tools	Functionalities
1. DEM Manipulations	e.g. Moore <i>et al.</i> , 1991; Palacios-Velez <i>et al.</i> , 1986; Rigon <i>et al.</i> , 2006
ExtractBasin	Extracts a basin by using the flow direction map
Markoutlets	Marks the outlets of a basin on the drainage direction map
Pitfiller	Fills the depression points of the DEM
SplitSubbasins	Labels the sub-basins of a basin using stream ordering
Wateroutlet	Extracts the watershed for a defined outlet
2. Geomorphology tools	e.g. Orlandini <i>et al.</i> , 2003; Tarboton, 1997; Mitasova and Neteler, 2004; Moore <i>et al.</i> , 1991; Garbrecht and Martz, 1997.
Aspect, slope, Gradient, curvature	Calculate aspect, slope, gradient and curvature type of the map respectively
FlowDirections, DrainDir, LeastCostFlowDirections	Calculate the D8 method drainage direction, drainage directions minimizing the deviation from the real flow, and least cost method drainage directions respectively
Tca, Gc	Calculate contributing areas and topographic classes, respectively
3. Hydro-geomorphology	
Hillshade	Calculates the shadows of the DEM
Skyview	Calculates the skyview factor of the DEM
Insolation	Estimates the amount of shortwave radiation on a surface for a given of time
4. Statistics	
Cb	Calculates the histogram and the statistical moments of a set of data from a map with respect to another map
SumDownStream	Calculates the sum values of a map from upstream to downstream following the flowdirections
Jami	An interpolation method
Variogram,	Calculates the experimental semivariogram
Kriging	Implements the ordinary kriging interpolation algorithm

basins (**ExtractBasin**, **SplitSubbasins**), and to find the basin outlets (**Wateroutlet**). Depression filling is perhaps the most widely implemented algorithm for depression removal and is found in all the terrain analysis tools (e.g. TauDEM - Tarboton, 1997; Rivix - Peckham, 2009; TAS - Lindsay, 2005; GRASS - Jasiewicz and Metz, 2011). JGrasstools uses the algorithm presented by Bras *et al.* (1986).

Geomorphology toolbox

The **Geomorphology** toolbox contains tools for calculation of slopes, curvatures, drainage directions and contributing areas, among many others. One of the simplest geomorphological attributes maps is the aspect map, a map that shows which side a slope is directed, this can be calculated using the **Aspect** tool from the DEM (Figure 3).

Terrain attributes are based on local neighbourhoods and reflect a simple application of the differential geometry of curves on surfaces (Peckham and Jordan, 2007). Algorithms involving upslope or downslope calculations (i.e. those within the basic topographic attributes and network related measures tools) rely on the steepest descent (or 'D8') flow-routing algorithm (O'Callaghan and Mark, 1984) because of the need for unique, non-diverging flowpaths. Two other algorithms involving analysis of neighbourhoods are implemented in uDig because using the "pure" D8 method for the drainage direction estimation causes deviation from the "real" flow direction identified by the gradients. The first algorithm implemented according to Orlandini *et al.* (2003) is the D8-LAD (least angular deviation), which minimizes the total angular deviation. The second algorithm, D8-LTD (least transversal deviation), minimizes the total deviation length of the flow going downstream. uDig (through the uDig Spatial toolbox) is currently the only GIS that contains these algorithms.

A third algorithm available is the multiple flow directions algorithm, first implemented by Fairfield and Leymarie (1991). This is used mainly for comparison, since this effect is barely found in nature (e.g. Orlandini *et al.*

2012), and D8-LAD and D8-LTD recover very precisely the real drainage directions. Gradient calculation (**Gradient, Slope**) is another standard tool present in all modern toolboxes.

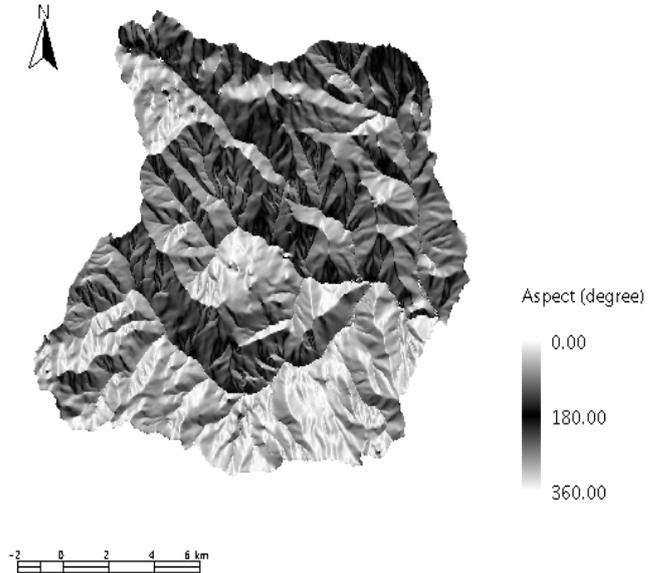


Figure 3: Aspect map of the Posina river basin, with enhanced visualization using the style editor tools in uDig GIS.

Table 5: Some of the tools available in the **Network sub** toolbox and their general functionalities.

HortonMachine Network tools	Functionalities
ExtractNetwork	Extracts the raster network from DEM (Orlandini <i>et al.</i> 2012; Montgomery and Dietrich, 1988, 1989)
HackLength	Calculates the distance of each pixel to the divides going upstream along the flow directions (Rigon <i>et al.</i> , 1996)
NetDiff	Calculates the difference between the value of a quantity in two network points with different numbering
Netnumbering	Assigns identification (id) numbers to the network links
NetworkAttributesBuilder	Extracts the network as a shapefile and adds networks attributes to it some (Rodriguez-Iturbe and Rinaldo., 1997; Rigon <i>et al.</i> , 1996)
DistanceToOutlet	Calculates the planar projection of the distance of each pixel to the outlet (D'Odorico and Rigon, 2003)
NetShape2Flow	Transforms the network shapefile to a flow raster map

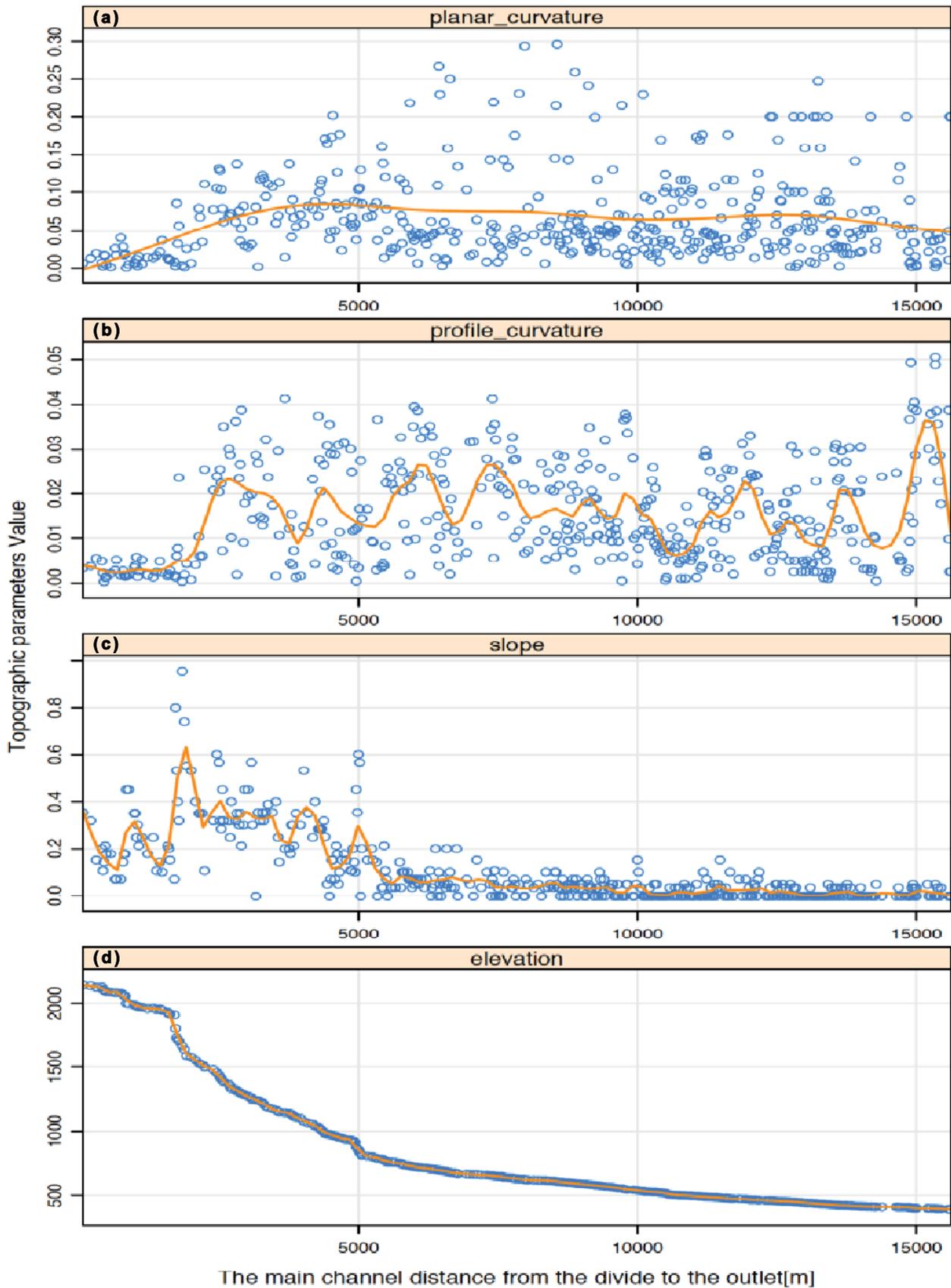


Figure 4: The topographic parameters: (a) planar curvature, (b) profile curvature, (c) slope and, (d) elevation, in the main stream of case study basin of Posina.

The classification of topographic sites into three different classes of curvature is another important tool (**Tc**, **Gc**). Longitudinal (or profile), normal and planar curvatures for each pixel are helpful to estimate the deviation of the gradient vector. Profile curvature measures the topographic curvature (i.e. the gradient deviation) along a flow line following the steepest descent path, and planform curvature measures the curvature of contour lines on topographic maps. Detailed description of different landform curvatures is found, for instance, in Tarolli *et al.* (2012). In combination with some other general tools, provided by the uDig Spatial Toolbox, these tools can provide the information shown in Figure 4 (see also complementary material at <http://abouthydrology.blogspot.it/2014/05/the-udig-spatial-toolbox-paper.html>). Figure 4 shows the elevation, slope and curvatures along the main stream of the Posina. Two knickpoints are particularly evident in the elevation plot, which reflect changes of gradient and curvature. The curvatures, in turn, are both positive starting after the first knickpoint, as expected, and identifying the presence of convergent-convex sites typical of a valley, and of channel geomorphology.

Network toolbox

The main tasks available in the **Network toolbox** in uDig are related to watershed extraction, various basin morphometric analyses, stream network extraction and analysis. These are presented in Table 5. The stream network extraction tool uses three alternative approaches: total contributing area threshold; slope-area threshold; and, curvature based. The first, and most common, method of extracting a channel is by setting some threshold on the total contributing area (**Tca**), representing the total area of upslope cells. Cells with a total contributing area greater than a given threshold area are considered to be flow channel, since Tca is considered a surrogate of discharge (O'Callaghan and Mark 1984).

In addition to the contributing-area threshold method, a slope-area threshold method based on work by Montgomery and Dietrich (1992) and a curvature based stream delineation method (Tarboton and Armes, 2001) have been implemented in the stream network extraction tool. Furthermore, stream

network analysis includes utilities to order channel streams (using Hack and Horton-Strahler ordering schemes, e.g. Rodriguez-Iturbe and Rinaldo, 1997). From this ordering, it is possible to derive statistics associated to the network, to estimate Shreve's magnitude, and to measure link-average slopes and lengths and from them estimate, for instance, Horton laws (e.g. number and length of channels per Horton order, bifurcation ratio, and length ratio; see Rodriguez-Iturbe and Rinaldo, 1997).

The **HackStream** tool provides the channel ordering based on Hack's stream enumeration (Rigon *et al.*, 1996). In Hack's ordering, the main channel of the network is assigned the order 1, the channels that flow into it are assigned the order of 2, and the branches that flow into channels of order 2 are assigned the order of 3, and so on.

The most common and popular method of channel classification is according to the Horton-Strahler ordering scheme (Horton, 1945; Strahler, 1957), which is implemented in the **NetworkAttributesBuilder** tool: the network is divided into links that connect either two tributary junctions (internal links) or a tributary junction and a channel source point (external links: Rigon *et al.*, 1996). This ordering system assigns order 1 to the source; and when two or more streams of the same order, n , meet they form a stream of order, $n+1$. When two streams of different orders, n and m with $n > m$, meet the order of the channel they form remains with the order of the greater of the two, n . The **NetworkAttributesBuilder** produces not only the raster enumeration, but also the vectorial features of the stream ordering.

Another method for labeling channel links and associated hillslope is the so-called Pfafstetter coding method (e.g. Verdin and Verdin, 1999). It provides the topographical connectivity between channels and hillslopes. The technical description of the Pfafstetter numbering schemes (PNS) as implemented in the **Pfaf** tool is given by Formetta *et al.* 2013a. The generalization of this coding system, implemented in the uDig Spatial Toolbox, can also take account of the presence of dams and irrigation channels. The **Pfaf** tool produces a shapefile (i.e. a vectorial feature) that contains, besides the enumeration itself (as shown in Figure 5), the

associated properties, such as the starting and ending point of a link, the elevation drop and other properties.

Hillslope toolbox

The **Hillslope** toolbox contents are presented in Table 6. They include tools for the classification of hillslope points into categories derived from information about curvatures, tools for evaluating distances of hillslope points to streams, and tools for calculating statistics of any quantity in a hillslope (Ghesla and Rigon, 2006). In addition to the estimation of pixels curvature, based on the longitudinal (profile) and transversal curvatures mentioned before, the topographic class (Tc) tool subdivides the sites of a basin in different topographic classes. The program has two outputs: the more detailed nine topographic classes (Parsons, 1988) and an aggregated topographic classification with three fundamental classes. Figure 6 is a visual comparison of an example of detailed nine and aggregated three topographic class maps of Posina river basin.

Planar curvature represents the degree of divergence or convergence perpendicular to the flow direction, and profile curvature shows convexity or concavity along the flow direction. By combining these two main curvatures, the topographic class (Tc) tool identifies 9 classes, which are three planar type sites (parallel-planar, divergent-planar convergent-planar sites), three convex type

sites (parallel-convex, divergent-convex and convergent-convex sites), and three concave type sites (divergent-concave, parallel-concave and convergent-concave sites). These attributes can be summarized into three fundamentals classes (concave, convex and planar sites). The graphical depiction of the curvature classification of hillslopes is shown in Figure 7.

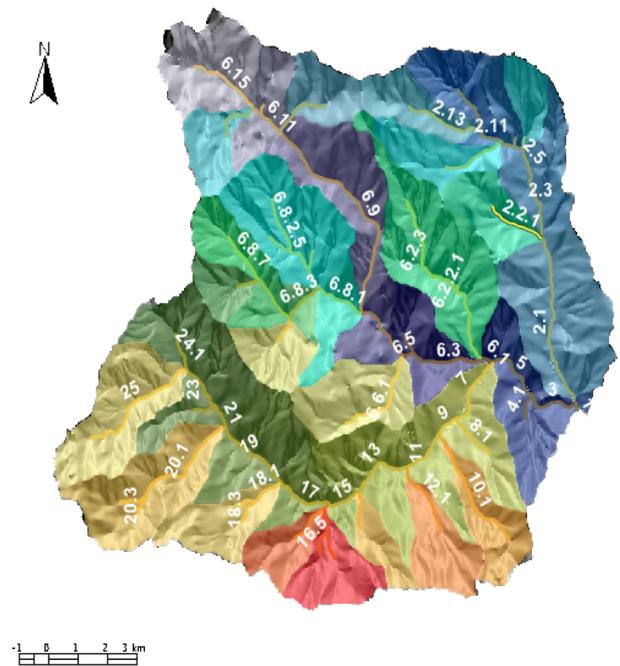


Figure 5: The pfafstetter enumeration scheme for the Posina river basin, as implemented in uDig GIS spatial toolbox for channel networks and hillslopes.

Table 6: The list of tools in Hillslope and Basin sub toolbox and their general functionalities.

HortonMachine tools	Functionalities
Hillslope toolbox	e.g. Parsons, 1988; Rodriguez-Iturbe and Rinaldo., 1997
H2CA	Estimates some attributes of hillslopes associated to a common channel network.
H2cd	Calculates hillslope distance from river network
Tc	Subdivides hillslopes into topographic classes
Basin toolbox	e.g. Rigon <i>et al.</i> 2011; D’Odorico and Rigon, 2003
BasinShape	Creates sub-basin shape file following the netnumbering tool
RescaledDistance	Calculates the rescaled distance of each pixel from the outlet
TopIndex	Calculates the topographic index of each sites

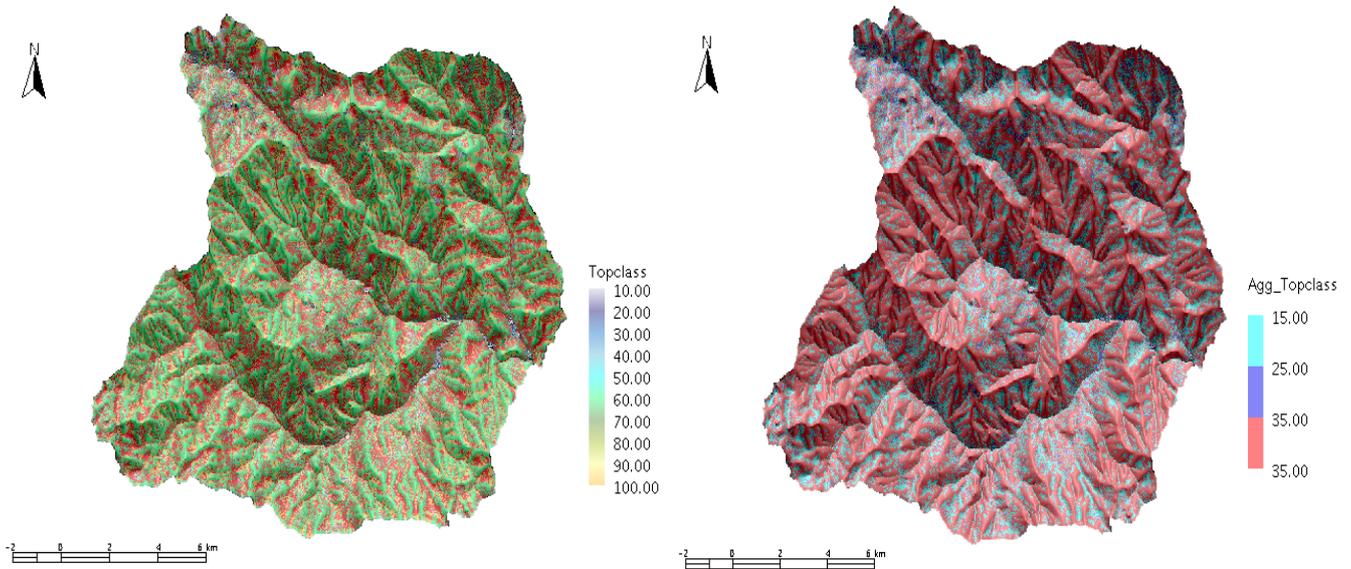


Figure 6: Hillslope topographic classes map of the Posina river basin. The first map (left) shows the nine hillslope classes based on topographic curvature (planar-planar sites (10), convex-planar sites (20), concave-planar sites (30), planar-convex sites (40), convex-convex sites (50), concave-convex sites (60), planar-concave sites (70), convex-concave sites (80), and concave-concave sites (90)). The second map (right) shows the three principal topographic classes (concave sites (15), planar sites (25), and convex sites (35)) of the basin.

		Contour line		
		Divergent	Parallel	Convergent
Flow line	Convex			
	Planar			
	Concave			

Figure 7: The subdivision of the hillslope sites according to their curvature (after Parsons, 1988).

In general terms, divergent-convex landforms are associated with the dominance of hillslope processes, while convergent-concave landforms are associated with

valley-dominated erosion (e.g. Tarolli and Dalla Fontana, 2009). Mapping these divergent and convergent sites is essential for the geomorphological and hydrological analyses of a basin, the local divergence and convergence roughly identifying convex zones as hillslope zones, the concave zones as valleys, which are subject to different processes (as also enlighten by Figure 4).

Other tools from this toolbox were used, for instance, to produce the calculations in D’Odorico and Rigon (2003) to evaluate the distance of any point in a hillslope to a channel. The hillslope to channel distance (**H2cd**) calculates the distance of each point on the hillslope to the channel network following the steepest descent (see Figure 8).

H2CA calculates the distance a drop of water released (or rained) in any point in a hillslope takes to arrive into a channel. **H2CA** plus **H2cd** is the total length from any point in a basin to the basin outlet. It is useful to separate these tools so as to associate to each of them a different residence time, as was done by Rinaldo *et al.* (1995).

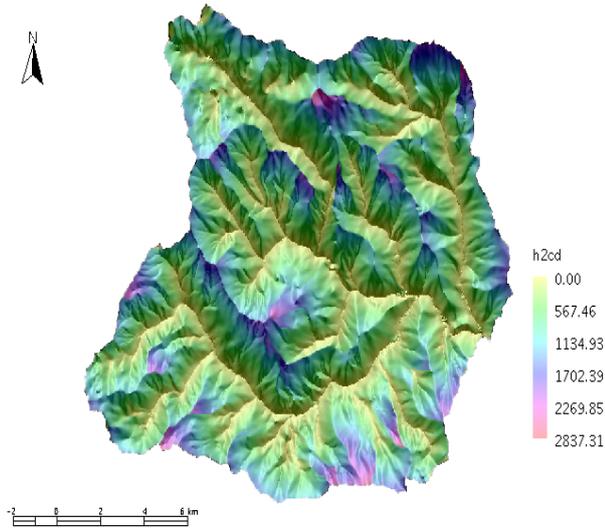


Figure 8: The map of the distance of each hillslope pixel to the channel (h2cd) in the Posina basin.

Figure 9, shows the distribution of distances from any point in a hillslope to channel versus the distance of the hillslope to the outlet. The figure was obtained after a little manipulation of the data (produced by the tool) made with R (<http://www.r-project.org>, please see the complimentary material). It clearly shows that the mean hillslope lengths of the Posina catchment are increasing downstream. The command **Drainage Density** (which equals the total network length per contributing area) can be used to obtain the homonymic quantity. Historically the two quantities, H2CA and drainage density were thought to be inversely proportional (e.g. Rodriguez-Iturbe and Rinaldo, 1997), and the second was often used to infer the first because easier to estimate from maps.

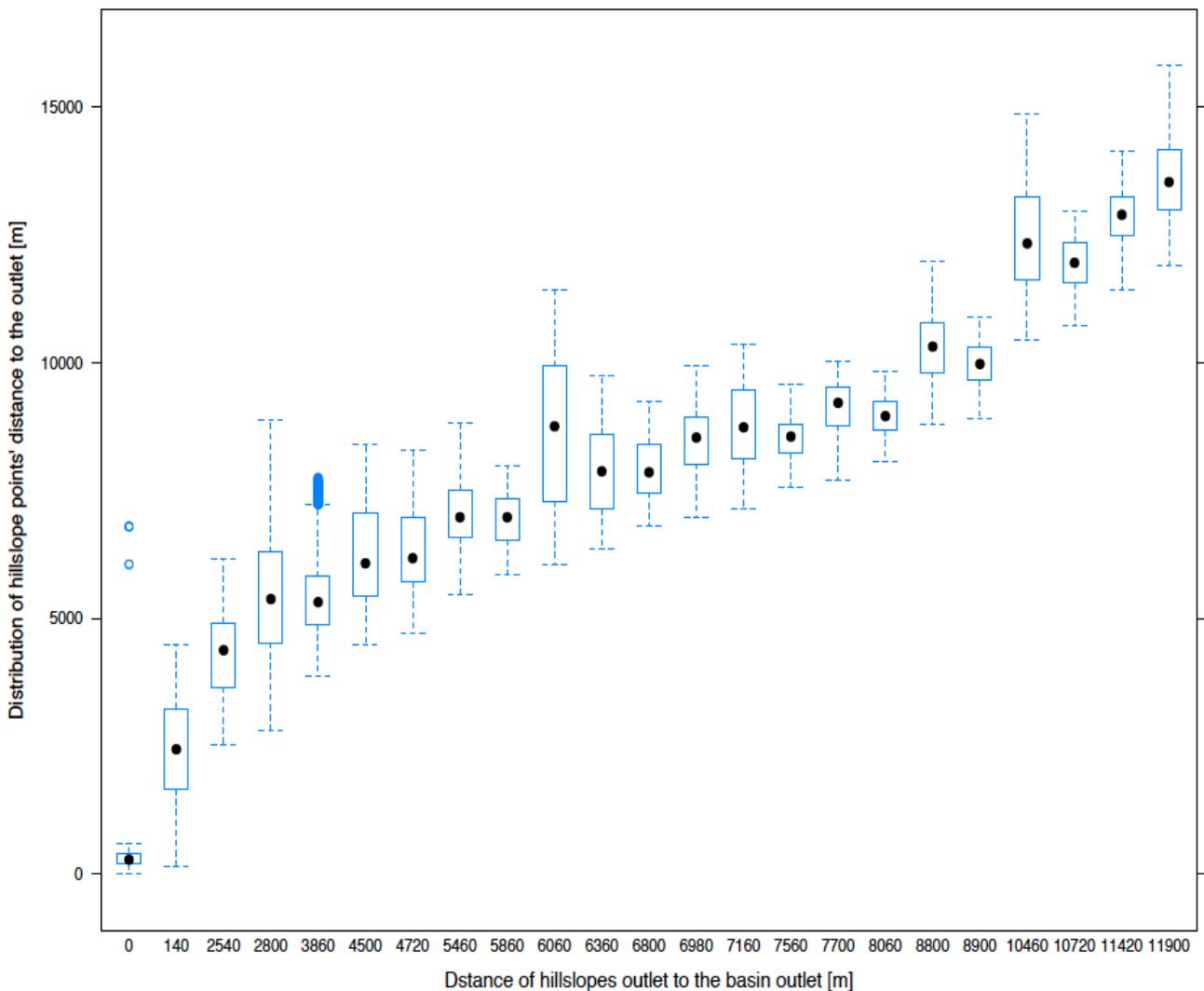


Figure 9: The distribution of hillslope pixel distance to the outlet versus the mean hillslope distance.

Basin toolbox

The **Basin toolbox** in the HortonMachine toolbox contains models to estimate basin-wide characteristics (shown in Table 6). They include, among others, methods to evaluate the width function (e.g. Rigon *et al.* 2011; D'Odorico and Rigon, 2003) and the rescaled width function (**RescaledDistance**; Rinaldo *et al.*, 1995), and topographic index (**TopIndex**) which is commonly used to quantify topographic control on hydrological processes which accumulate soil moisture (Beven and Kirkby, 1979). This has been criticised as a model for deriving maps of soil (see Barling *et al.*, 1994; Lanni *et al.*, 2012), however, the topographic index still remains a useful visualisation of the process of saturation, which can serve as a first approximation to understand which points saturate first (e.g. Crave and Gascuel-Oudou, 1997; Hjerdt *et al.* 2004). The rescaled distance is the distance of each pixel from the

outlet measured along the drainage directions, weighted by the ratio of the water velocity in channels and on the hillslope. If the ratio of velocities is taken equal to one, the normal planar projection of the distances to outlet for any point in a basin is obtained.

The topographic index classifies the basin based on its ability to generate surface flow, according to Beven and Kirkby (1979). As is known, sites with a higher topographic index tend to become saturated before sites with a lower topographic index. The map showing the topographic index and the rescaled distances for each pixel of the study basin is shown in Figure 10. **BasinShape** is a tool which creates feature collections of sub-basins extracted by the netnumbering tool. It is useful for extracting important information from each sub-basin, such as area, perimeter, max elevation, minimum elevation, mean elevation etc.

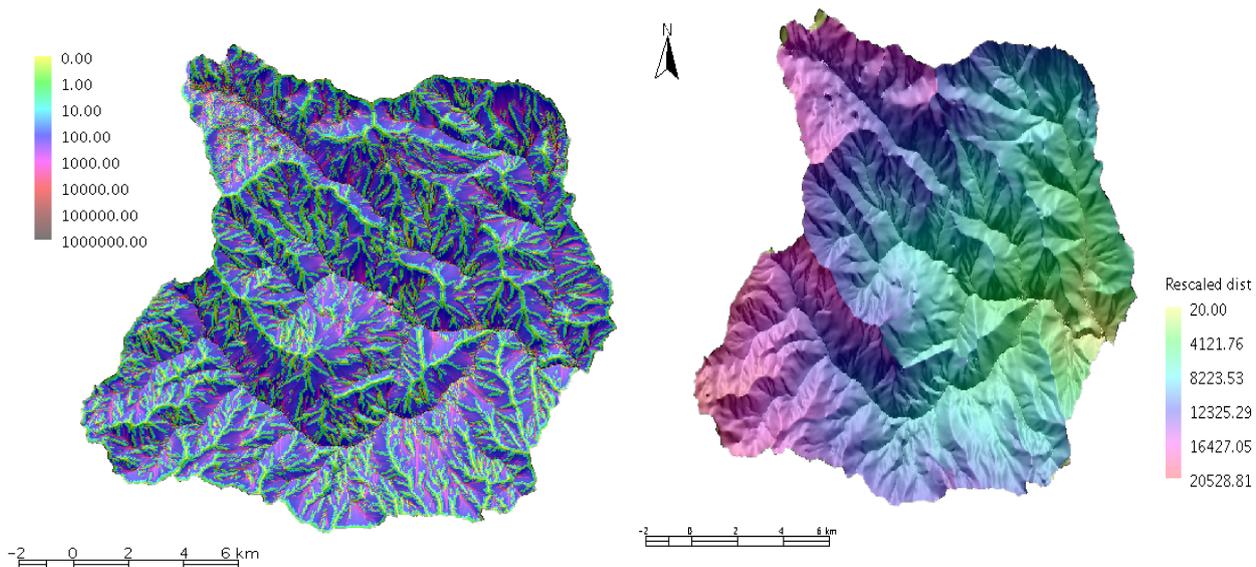


Figure 10: The map of the topographic index (left) and the rescaled distance (right) for the Posina river basin.

Statistics toolbox

In addition to the terrain analysis functionalities described above, the spatial toolbox also incorporates statistical tools (presented in Table 4). Among these, there are tools for both deterministic and geostatistic interpolation algorithms. These include Just Another Meteo Interpolator (JAMI) and kriging interpolation tools. JAMI is

a robust approach of interpolating different meteorological data presented in Formetta (2013). The geostatistical technique implemented in the statistical toolbox is kriging. At the moment, the ordinary kriging algorithm (Goovaerts 1997, 2000) is the one implemented in the toolbox. If input data are provided as time series, the **kriging** runs over all the time steps, estimating a different semivariogram model, and the parameters

used for kriging interpolation, for each time step. The kriging tool provides both point (non-regular grid) and regular raster grid outputs. Figure 11 and 12 are examples of grid and

point interpolations obtained using exponential semivariogram model fitting. Figure 11 is a scatter plot comparing measured and interpolated hourly meteorological data (temperature).

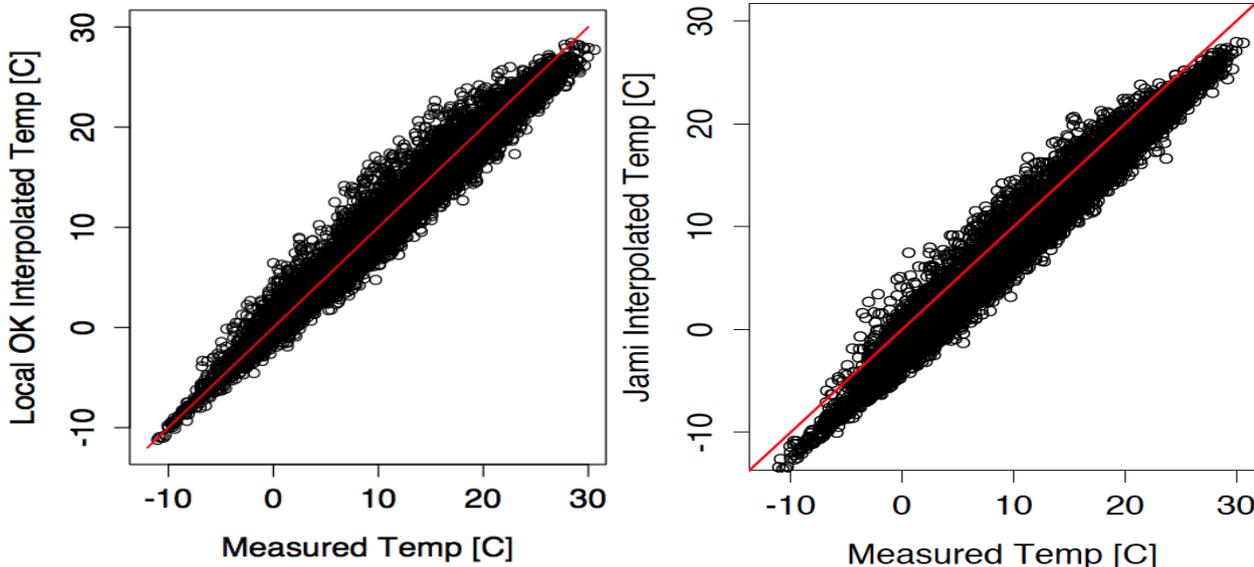


Figure 11: Scatter plot of measured and ordinary kriging interpolated (left: with $R^2 = 0.78$) and JAMI interpolated (right: with $R^2 = 0.74$) hourly temperature for one year (1995), in one of the measurement station in Posina river basin.

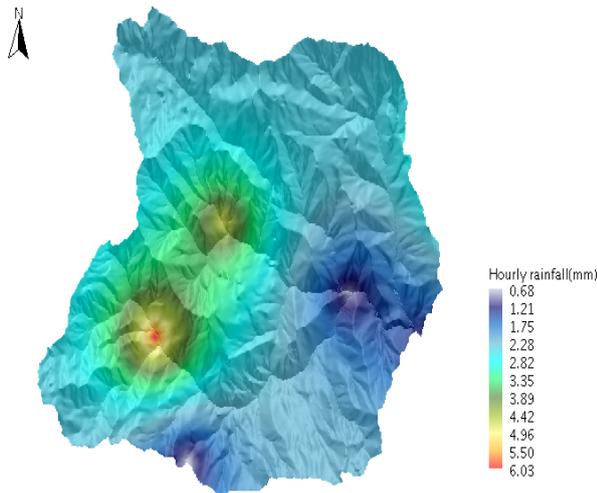


Figure 12: The grid-based rainfall interpolation using Ordinary kriging for the Posina basin at time step 1994-01-01 01:00 hour.

Further characteristics of the Spatial Toolbox

The geomorphological tools described in the previous sections are made even more effective by the general characteristics that

the Spatial Toolbox inherits from uDig, which are summarize briefly below.

Visualization

The graphical user interface (GUI) of the Spatial Toolbox allows multiple images to be displayed simultaneously with transparency effects, facilitating the visual inspection of multiple terrain attributes. Displayed images can also be combined with shaded-relief images to enhance visualization of terrain. In these 'composite-relief models', variations in colour correspond to the displayed attribute and tonal variations correspond to hill shading. Vector data may be overlaid onto raster images to enhance data visualization and interpretation. Spatial Toolbox is distributed with a standard set of colour palettes, which have been set as the default. Nonetheless, users can eventually create custom palettes for specific purposes using the **Palette manager**

Importing and exporting data

DEMs are the main input data to Spatial toolbox, but the program can utilize many other types of spatial data, including satellite

imagery. Raster import / export functions include read and write ArcView raster formats, GRASS images, Surfer grids, Autodesk .dwg and device independent bitmaps. The program can also read all the supported raster data included in the GDAL (<http://www.gdal.org>) library and the Shuttle Radar Topography Mission (SRTM) data (<http://www.ppp.org>). JGrasstools in Spatial toolbox reads and writes shapefiles, GRASS ASCII and native (supported since the GRASS 5.0) vector file formats, and delimited XYZ vector point files. Graphical output (i.e. displayed images with vector overlays) can be saved as MS-Windows bitmap (.wbmb), jpg, jpeg, and Portable Network Graphics files (.png), which can be read by most graphical packages and several word-processing programs.

Program development and availability

The spatial toolbox was originally developed as JGrass, and is now available as a collection of different tools (jGrasstools). The capabilities of the program have extended considerably in the past few years. The Spatial toolbox operates under Microsoft Window, Linux and Mac OsX platforms. Hardware requirements vary depending on the size of DEM or image being processed, but the standard requirements are 512MB RAM.

Spatial toolbox latest source code is also available for download at: <https://code.google.com/p/jgrasstools/>.

Program development is ongoing (the stable version is uDig 1.4.0 and jgrasstools0.7.7). The Hydrologis team has made a wiki page (<http://code.google.com/p/jgrasstools/>) in which the evolution of the software is tracked and documented. A mailing list for users and developers is available at jgrasstools@googlegroups.com and the authors welcome comments and feedback. The Spatial toolbox is maintained by the authors of the paper.

Complementary material that explains how the Figures in the paper were generated can be found at: <http://abouthydrology.blogspot.it/2014/05/the-udig-spatial-toolbox-paper.html>

Concluding remarks

The uDig Spatial toolbox is a powerful, research-grade environmental modeling environment. The main tools have been described in this chapter, however many more tools are available, particularly an advanced Hydrological model called JGrass-NewAGE modelling system (Formetta *et al.* 2011, 2013b, 2014), complete sub-models for estimating rainfall-runoff, radiation, evapotranspiration, snow water equivalent, landslide models like SHALSTAB (Montgomery and Dietrich, 1994), and CISLAM (Lanni *et al.*, 2012) peak flow modelling (Rigon *et al.*, 2011).

Tools in JGrasstools are ideal for both research and student instruction due to ease of use and free availability. Therefore, we believe, the uDig Spatial toolbox is suited to be used in research and education in physical geography, hydrology, geomorphology, climatology, environmental science and watershed modelling.

Acknowledgements

Funding for developing spatial toolbox has been provided by the University of Trento, Department of Civil and Environmental Engineering and CUDAM (University Centre for the Hydrogeological Defence of the Mountain Environment). The HydroAlp grants from Provincia Autonoma di Bolzano is the main sponsor of the project. At present, the modules are developed at CUDAM and HydroloGIS.

References

- Band LE. 1993. Extraction of channel networks and topographic parameters from digital elevation data. In: Beven K, Kirkby MJ (Eds.) Channel Network Hydrology/ Wiley: New York, pp. 13-42.
- Band LE. 1986. Topographic partition of watersheds with digital elevation models. *Water Resources Research*. **22**: 15-24. DOI: 10.1029/WR022i001p00015.
- Barling RD, Moore ID, Grayson RB. 1994. A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content. *Water Resources Research*. **30**:1029-1044. DOI: 10.1029/93WR03346.

- Beven K, Kirkby M. 1979. A physically based variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, **24**: 43-69.
- Borga M, Anganostou EN, Frank E. 2000. On the use of real-time radar rainfall estimates for flood prediction in mountainous basins. *Journal of Geophysical Research* **105**: 2269-2280. DOI: 10.1029/1999JD900270.
- Bras R, Tarboton D, Rodriguez-Iturbe I. 1986. The analysis of river basins and channel networks using digital terrain data. Cambridge Mass., Ralph M. Parsons Lab, M.I.T., *Tech. Rep. No. 326*.
- Crave A, Gascuel-Oudou C. 1997. The influence of topography on time and space distribution of soil surface water content. *Hydrological Processes* **11 (2)**: 203-210. DOI:10.1002/(SICI)1099-1085(199702)11:2%3C203::AID-HYP432%3E3.0.CO;2-K.
- D'Odorico P, Rigon R. 2003. Hillslope and channels contribution to the hydrologic response. *Water Resources Research* **39**:1-9. DOI: 10.1029/2002WR001708.
- Evans IS, Dikau R, Tokunaga E, Ohmori H, Hirano M.(eds). 2003. *Concept and Modelling in Geomorphology: International Perspective*, Terrapub, Tokyo
- Fairfield J, Leymarie P. 1991. Drainage networks from grid digital elevation models. *Water Resources Research* **27**: 709-717. DOI: 10.1029/90WR02658.
- Formetta G. 2013. Hydrological Modelling with Components: the OMS3 NewAge-JGrass System. PhD Thesis: Trento University, Italy.
- Formetta G, Antonello A, Franceschi S, David O, Rigon R. 2014. Hydrological modelling with components: A GIS-based open-source framework. *Environmental Modelling & Software* **55**: 190-200. DOI: 10.1016/j.envsoft.2014.01.019
- Formetta G, Antonello A, Franceschi S, David O, Rigon R. 2013a. The basin delineation and the built of a digital watershed model within the JGrass-NewAGE system. *Boletín Geológico y Minero: Special Issue "Advanced GIS terrain analysis for geophysical applications"*.
- Formetta G, Rigon R, Ch´avez JL, David O. 2013b. Modeling shortwave solar radiation using the JGrass-NewAge system. *Geoscientific Model Development* **6**: 915-928.
- Formetta G, Mantilla R, Franceschi S, Antonello A, Rigon R. 2011. The JGrass-NewAge system for forecasting and managing the hydrological budgets at the basin scale: models of flow generation and propagation/routing. *Geoscientific Model Development*, **4**: 943-955. DOI: 10.5194/gmd-4-943-2011.
- Garbrecht J, Martz LW. 1997. The assignment of drainage direction over flat surfaces in raster digital elevation models. *Journal of Hydrology* **193**: 204-213. DOI: 10.1016/S0022-1694(96)03138-1.
- Ghesla E, Rigon R. 2006. *A Tutorial for Preparing GEOTop Input Files with JGrass*: http://www.ing.unitn.it/dica/tools/download/Quaderni/tutorial_input_geotop_ENG.pdf.
- Goovaerts P. 2000.. Geostatistical approaches for interpolating elevation into spatial interpolation of rainfall. *Journal of Hydrology* **228**: 113-129. DOI: 10.1016/S0022-1694(00)00144-X.
- Goovaerts P. 1997. *Geostatistics for Natural Resource Evaluations*. Oxford University Press, USA.
- Grayson R, Moore I, Landson A. 1991. Digital terrain modelling: a review of hydrological, geological and biological applications. *Hydrological Processes* **5**: 3-30.
- Hjerdt, KN, McDonnell JJ, Seibert J, Rodhe A. 2004. A new topographic index to quantify downslope controls on local drainage. *Water Resources Research* **40(5)**: W05602. DOI: 10.1029/2004WR003130.
- Horton RE. 1945. Erosional development of streams and their drainage basins: hydro-physical approach to quantitative morphology. *Geological Society of America Bulletin* **56**: 275-370.
- Jasiewicz J, Metz M. 2011. A new GRASS GIS toolkit for Hortonian analysis of drainage networks. *Computers and Geosciences* **37(8)**: 1162-1173. DOI: 10.1016/j.cageo.2011.03.003.
- Lanni C, Borga M, Rigon R, Tarolli P. 2012. Modelling shallow landslide susceptibility by means of a subsurface flow path connectivity index and estimates of soil depth spatial distribution, *Hydrological Earth System*

- Science* **16**: 3959-3971. DOI: 10.5194/hess-16-3959-2012.
- Lashermes B, Foufloula-Georgiou E. 2007. Area and width functions of river networks: New results on multifractal properties. *Water Resources Research* **43**: W09405, DOI: 10.1029/2006WR005329.
- Lindsay JB. 2005. The Terrain Analysis System: tool for hydro-geomorphic applications. *Hydrological Processes* **19**:1123-1130. DOI: 10.1002/hyp.5818.
- Mitasova H, Neteler M. 2004. GRASS as Open Source - Free Software GIS: accomplishments and perspectives. *Transactions in GIS* **8**: 145-154. DOI: 10.1111/j.1467-9671.2004.00172.x.
- Montgomery DR, Dietrich WE. 1994. A physically based model for the topographic control on shallow landsliding. *Water Resources Research* **30**: 1153-1171. DOI: 10.1029/93WR02979.
- Montgomery DR, Dietrich WE. 1992. Channel initiation and the problem of landscape scale. *Science* **255**: 826-830. DOI: 10.1126/science.255.5046.826.
- Montgomery DR, Dietrich WE. 1989. Source areas, drainage density, and channel initiation. *Water Resources Research* **25**: 1907-1918. DOI: 10.1029/WR025i008p01907.
- Montgomery DR, Dietrich WE. 1988. Where do channels begin? *Nature* **336**: 232-234. DOI: 10.1038/336232a0.
- Moore ID, Grayson RB, Ladson AR. 1991. Digital terrain modelling : A review of hydrological, geomorphological and biological applications. *Hydrological Process* **5**: 3-30. DOI: 10.1002/hyp.3360050103.
- O'Callaghan J, Mark DM. 1984. The extraction of drainage networks from digital elevation data. *Computer Vision Graph* **28**: 323-344. DOI: 10.1016/S0734-189X(84)80011-0.
- Orlandini S, Moretti G, Corticelli MA, Santangelo PA, Capra A, Rivola R, Albertson JD. 2012. Evaluation of flow direction methods against field observations of overland flow dispersion, *Water Resources Research* **48**: W10523. doi:10.1029/2012WR012067. DOI: 10.1029/2012WR012067.
- Orlandini S, Moretti G, Franchini M, Aldighieri M, Testa B. 2003. Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevations models. *Water Resources Research* **39**(6): 1144. DOI: 10.1029/2002WR001639.
- Palacios-Velez O, Cuevas-Renaud B. 1986. Automated river-course, ridge and basin delineation from digital elevation data. *Journal of Hydrology*. **86**: 299-314. DOI: 10.1016/0022-1694(86)90169-1.
- Passalacqua P, Tarolli P, Foufloula-Georgiou E. 2010a. Testing space-scale methodologies for automatic geomorphic feature extraction from LiDAR in a complex mountainous landscape. *Water Resources Research* **46**: W11535, DOI:10.1029/2009WR008812.
- Passalacqua P, Do Trung T, Foufloula-Georgiou E, Sapiro G, Dietrich WE. 2010b. A geometric framework for channel network extraction from LiDAR: Nonlinear diffusion and geodesic paths. *Journal of Geophysical Research* **115**: F01002. DOI: 10.1029/2009JF001254.
- Parsons AJ. 1988. *Hillslope Form*. Routledge: London. DOI: 10.4324/9780203330913.
- Peckham RJ, Jordan G. 2007 Digital Terrain Modelling. *Lecture Notes In Geoinformation and Cartography*. Springer: London, pp. 1–326.
- Peckham SD. 2009. Geomorphometry in RiverTools. In: Hengl T, Reuter HI. (Eds.) *Geomorphometry: Concepts, Software and Applications*. Elsevier: Amsterdam, pp. 311-324.
- Pike RJ. 2002. *A Biography of Terrian Modeling (Geomorphometry): The Quantitative Representation of Topography*, Supplement 4.0. Open-File Report 02-465.
- Rigon R, D'Odorico P, Bertoldi G. 2011. The geomorphic structure of the runoff peak, *Hydrology and Earth System Science* **15**: 1853-1863, DOI:10.5194/hess-15-1853-2011.
- Rigon R, Ghesla E, Tiso C, Cozzini A. 2006. *The Horton Machine: a System for DEM analysis*. ISBN **10**: 88-8443-147-6. E-book: <http://www.ing.unitn.it/dica/tools/download/Quaderni/Horton%20manual.pdf>.
- Rigon R, Rodriguez-Iturb I, Maritan A, Giacometti A, Tarboton D, Rinaldo A. 1996. On Hack's law. *Water Resources Research* **32**: 3367-3374. DOI: 10.1029/96WR02397.
- Rinaldo A, Vogel G, Rigon R, Rodriguez-Iturb I. 1995. Can one gauge the shape of a

basin?, *Water Resources Research* **31(4)**: 1119-1127. DOI: 10.1029/94WR03290.

Rodriguez-Iturbe I, Rinaldo A. 1997. *Fractal River Networks. Chance and Self-Organization*. Cambridge University Press, New York.

Strahler AN. 1957. Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* **38(6)**: 913-920. DOI: 10.1029/TR038i006p00913.

Tarboton DG. 1997. A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models. *Water Resources Research* **33(2)**: 309-319.

Tarboton DG, Ames DP. 2001. Advances in the mapping of flow networks from digital elevation data. In *World Water and Environmental Resources Congress*, Orlando, Florida, May 20-24, ASCE: <http://www.neng.usu.edu/cee/faculty/dtarb/asce2001.pdf>.

Tarolli P, Dalla Fontana G. 2009. Hillslope-to-

valley transition morphology: New opportunities from high resolution DTMs. *Geomorphology* **113(1- 2)**: 47-56. DOI: 10.1016/j.geomorph.2009.02.006.

Tarolli P, Sofia G, Dalla Fontana G. 2012. Geomorphic features extraction from high-resolution topography: landslide crowns and bank erosion. *Natural Hazards* **61(1)**: 65-83. DOI: 10.1007/s11069-010-9695-2.

Verdin KL, Verdin JP. 1999. A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology* **218**:1-12. DOI: 10.1016/S0022-1694(99)00011-6.

Wilson JP, Gallant JC. 2000. *Terrain Analysis*. John Wiley & Sons: Chichester.

Wood J. 2002. *Java Programming for Spatial Sciences*. Taylor and Francis: London. DOI: 10.1201/9780203166178.

Wood J. 2009. The LandSerf Manual. Version 1.0, 3rd December 2009: <http://www soi.city.ac.uk/~jwo/landserf/landserf230/doc/landserfManual.pdf>.

Appendix A

Table A1: Comparison of some selected open, free GIS software for hydro-geomorphological analysis.

	TauDEM	LandSerf	GeoNet	Whitebox Geospatial Analysis tools	GRASS	uDig Spatial Toolbox
Surface parameters and topographic feature extraction	Slope, aspect, curvature, topographic	Slope, aspect, curvature	Slope, curvature, landform features	Slope, aspect, curvature	Slope, aspect, curvature, landforms	Slope, aspect, curvature, sky view
Flow Direction	D ∞ , D8	D8	D ∞ , D8	D ∞ , D8, FD8	D8, MFD	D8, D8-LAD, D8-LTD
Channal delineation	tca, area-slope, area at concave	tca	Yes	tca	tca, slope-area method	tca, slope-area, concave sites
Enumerate networks	-	-	Horton, Strahler	Horton-Strahler stream order	Horton, Strahler, Shreve, pfaftstater	Horton, Strahler, Hack, netnumber, pfaftstater
Delineation of basin and subbasin	Yes	Yes	Yes	Yes	Yes	Yes
Data Interpolation algorithms	Thiessen polygon	Yes	x	Nearest neighbour	Yes	JAMI, Kriging
Flow distances	Yes	Yes	Yes	Yes	Yes	Yes
Raster calculator and Raster summary	Yes	Yes	Yes	Yes	Yes	Yes
Topographic index	Yes	-	-	Yes	Yes	Yes
Vector analysis capability	Yes	Yes	Yes	Yes	Buffer, polygonizer, clip, merge, reshape	Buffer, polygonizer, clip, merge, reshape
Map layer visualisation	Yes	Yes	Yes	Hillshade layer	3D vector map, Voxel	Aspect, 3D
Raster and vector transformation	Yes	Yes	Yes	Yes	Yes	Yes
Surface interpolator	IDW, TIN, splines	Cressman interpolation	Yes	IDW, nearest neighbour	IDW, nearest neighbour	IDW, Delaunay triangulation, Thin Plate Spline

2.4.2. Terrain analysis and landform recognition

Calogero Schillaci¹, Andreas Braun¹ & Jan Kropáček¹

¹Department of Geosciences, Tuebingen University, Germany (calogero.schillaci@student.uni-tuebingen.de)



ABSTRACT: In the last decades, Geographic Information Systems (GIS) allowed the detailed analysis of land surface, whereas the development in Remote Sensing (RS) offered increasingly detailed Digital Elevation Models (DEMs) and multispectral imagery. This combination fostered a rapidly evolving research in the field of geomorphology. In this contribution, a workflow for a Digital Terrain Analysis (DTA) is introduced, based on a free open source software (FOSS) called SAGA, and an overview on landform recognition and extraction from DEMs provided. Furthermore, a variety of DEM data sources are outlined, on which the multiscale DTA can be performed. An overview of DEMs, pre-processing and visualization techniques are elaborated and a case study in the Ethiopian highlands is used to compare three different resolution DEMs (90, 30, 2 m) in terms of their potential for: visual analysis, landform characterization and hydrological application. A typical workflow for the generation of primary and secondary terrain derivatives is therefore illustrated, as well as landform identification, as well as a description of the challenges and problems that can occur in this context.

KEYWORDS: Digital Terrain Analysis; Landforms; SAGA; SRTM; Digital Elevation Models

Introduction

Terrain analysis

The aim of Terrain Analysis (TA) is the explanation of the arrangement of the Earth's surface as well as their classification based on the surface pattern similarities (Strahler, 1957; Wilson and Gallant, 2000). Landforms can be defined as specific geomorphic features on the Earth, ranging from large-scale features (such as plains and mountain ranges) to minor features (such as individual gullies, faults and valleys), both man-made or from natural genesis which have a defined range of physical and visual characteristics (for a complete description see Goudie, 2004).

Before the introduction of Digital Elevation Models (DEMs), landforms were only manually identified by means of surveys when available through interpretation of aerial photographs (Garbrecht and Martz, 2000). During the 19th and early 20th century, the study of the landforms aimed at the production of physiography maps (Thornbury, 1965; Graf, 1987). Since World War I, aerial

photographs were extensively used to give a view of the enemy's area, aiming at a description of the landforms (Pavlopoulos *et al.*, 2009). During this period, aircraft were equipped with cameras to record enemy movements. Technological advances in the photogrammetry and aerial photography interpretation were made in this time frame and boosted the geomorphological field of research.

In the late 20th century when the computing capacity of processors improved, DEMs have regularly been used for geomorphological investigations. Research was conducted on the effect of spatial resolution and the quality of the obtained derivatives (Chairat and Delleur, 1993; Zhang and Montgomery, 1994; Garbrecht and Martz, 2000; Pike, *et al.*, 2008).

Through TA, it is possible to achieve a qualitative and quantitative description of landforms (Florinsky, 2012). Therefore achieving a morphographical characterization of the terrain (Klimaszewski, 1982) that represents the appearance and shape of the features. Within the last forty years, physical

geography research has used TA techniques for the topographic analysis and visualization of the land surface features. This includes: drainage pattern extraction; river morphology (e.g. Peucker and Douglas, 1975; O'Callaghan and Mark, 1984; Skidmore, 1989; Smith *et al.*, 1990; Band, 1993); watershed delineation (e.g. Jenson and Domingue, 1988; Band, 1986); surface roughness assessment (e.g. Grohmann *et al.*, 2010); monitoring the slope movements (e.g. Wiecek and Snyder, 2009); predicting the spatial distribution of gully erosion, and soil texture (e.g. Zakerinejad and Maerker, 2014).

TA represents an established application that allows for the extraction of landforms and their qualitative and quantitative spatial assessment (Wilson, 2012). In particular, the availability of worldwide medium-resolution DEMs and the rapid development of different digital stereo photogrammetry software has allowed the extraction of detailed DEMs. Consequently, software (outlined in Table 1), tailored workflows and user-friendly toolboxes have been developed (Abera *et al.*, 2014), allowing geographers and geomorphologists their complete use.

Table 1: Software for Digital Terrain Analysis

Software	Information	Source
SAGA GIS	Free and open source, runs on Windows and Linux Focuses raster processing, support of various formats Highly specialized DTA	http://www.saga-gis.org
QGIS	Free and open source, user friendly Focuses both raster and vector processing Basic DTA, but growing pool of user-scripted extensions	http://www.qgis.org
MicroDEM	Freeware, runs on Windows Focuses raster processing, recommended for advanced users Many applications related to DEMs	http://www.usna.edu/Users/occano/pguth/website/microdem/microdem.htm
ERDAS Imagine	Commercial, one of the leading products in remote sensing Very user friendly and stable processing Terrain module including basic and advanced DTA	http://www.hexagongeospatial.com/products/remote-sensing/erdas-imagine
VisualSFM	Freeware, runs on Windows, Linux and Mac OSX Allows the 3D reconstruction of point sources and images Limited documentation (still in progress)	http://ccwu.me/vsfm/
Agisoft PhotoScan	Commercial, but leading software for photogrammetric use Allows the 3d reconstruction of various input source data Large documentation and support	http://www.agisoft.com

Landform recognition

Different studies have proposed various definitions of landforms (Speight, 1990; Milne *et al.*, 1995). Those converge in the definition of landforms as a *uniform element* in terms of the geomorphometric similarity (Minár *et al.*, 2013). Geomorphic thresholds are at the base of any landform recognition (Schumm, 1979). Mathematically, hillslopes have been investigated as a continuous surface made up by neighboring objects. Those objects or units have been classified from geographer in two ways: i) based on their geometry (classical approach); and, ii) based on their semantics, i.e. classes chosen and named by the scientist's contexts and aims (Dehn *et al.*, 2001).

With the arrival of DEMs at a middle resolution (5-30 cell size), attempts have been made towards an automated and hierarchical subdivision of the terrain into homogenous units (Dikau, 1989). The latter defined the *facets* as landform units defined by a homogenous profile and plan curvature. The concept of landform classification has been established, assigning rules (thresholds) to the DEM first and second derivatives

To improve Dikau's (1989) methodology (based on a square nearest neighbor research), Brabyn (1998) developed a methodology based on a circle nearest neighbor pixel search. He suggested the use of a circular-searching window, rather than a square to remove the problem related to the

microrelief representation. In this overview, the work of Evans (2012) on the specific geomorphometry of glacial cirques and drumlins should also be mentioned. His pioneering work was about programming computers on the automatic calculation of morphometric features, such as slope and other parameters using both manual measures and DEM-based attributes. Despite these landforms not being present in the study area, there are visual distinctive landscape units and have been studied for paleo-glaciology issues (Smith *et al.*, 2009).

Regarding Karst landforms, McIlroy de la Rosa (2012) gives an extensive overview of digital classifications at different scales, as well as suitable techniques. At the macroscale level remote sensing, photogrammetry or LiDAR are named as best options for DEM generation. Examples for the classification of fluvial landforms are given by Gilvear *et al.* (1999), Ramasamy and Paul (2005), Large and Heritage (2012), and Javernick *et al.* (2014).

Erosion landforms have gained a growing interest in the last decades due to their large off-site impacts on cultivated and residential lands (Valentin *et al.*, 2005) and the potential for low-cost monitoring (Marzloff and Poesen, 2009). The study of erosion landform development has brought researchers to

define functional approaches to predict their future activity and spatial distribution (Märker *et al.*, 2011), the latter attempt to derive Erosion Response Units (Sidorchuk *et al.*, 2003) from sophisticated machine learning and statistical approaches.

Prima *et al.* (2006) have carried out an example of regional scale automated landform classification in a volcanic environment. The methodology employed was a supervised classification of morphometric derivatives (slope and topographic openness) to achieve quantitative landform classification.

The remainder of the chapter will discuss the data capture and pre-processing that need to be undertaken, the terrain attributes methods available and then present a GIS case study highlighting how these are implemented.

Data capture (DEM sources) and pre-processing

There are several software options for TA, these are summarized in Table 1. This article will focus on the System for Automated Geo-Analysis (SAGA, open source GIS).

Table 2: Sources of DEMs at various scales

Dataset	Spatial resolution	Use in geomorphology	Suitable scale	Source and price	Notes
SRTM C	1 arc-sec (90 m)	Regional overview of landforms, landscape type, regional tectonic structures (main faults)	1:100,000	USGS/EROS Data Center, CGIAR (void-filled version). Free of charge	Two versions available original and filled void.
SRTM C (2014)	3 arc-secs (30 m)	Landforms, drainage pattern	1:25,000	EROS / NASA/ JPL. Free of charge.	Released in 2014
SPOT	20 m	Landforms, drainage pattern	1:25,000	ISIS/CNES 10€/km ² (Scene)	Processed on-demand.
ALOS PRISM	10 m	Landforms, drainage pattern, surface roughness, volumes calculation	1:15,000	Image triplet from ESA (Third Part Mission). ESA grant (50 euro) scene about 1125 km ²	Processed in Leica Photogrammetry-suite.
Aerial Photos	2 m	Landforms, detailed drainage pattern, surface roughness, fault scarps, volume calculation, gullies, riverine morphology	1:5,000	Local mapping agency or authorities. Costs may vary.	Processed by Structure from Motion (SfM). Mostly acquired with metric cameras

Several DEMs have been used for TA research (summarized in Table 2). Valid examples that can support TA and landform recognition are represented by the global DEM, such as SRTM-C DEM 90 m spatial resolution (± 20 m error at 90% confidence) and SRTM-C 30 m spatial resolution (± 16 m at 90% confidence) (Farr *et al.*, 2007). At a regional scale, ALOS PRISM DEM with resolution 10 m spatial resolution (± 5 m at 90% confidence) based on satellite stereo photogrammetry (Tadono *et al.*, 2014), and, the local DEM with 2 m spatial resolution (± 3 m at 90% confidence) resolution, based on aerial stereo photogrammetry Structure from Motion (SfM) (Westoby *et al.*, 2012).

Independent of the DEM's source, pre-processing is a required step to avoid artificial spikes and isolated pixels that will affect the generated surface (Wang and Liu, 2006).

Automatically generated DEMs often have artifacts, which represent local alteration of the land surfaces, arising from different factors such as feature matching techniques, coarse spatial resolution or reconditioning by anthropic structures, i.e. buildings or bridges. These are often just a few pixels large, but they can cause problems with hydrological modeling. Wang and Liu (2006) have provided an accurate description of the nature/behaviour of surface depressions. Therefore, an algorithm called "Fill Sinks" can be applied to a DEM to prevent faulty results.

However, especially for high resolution DEMs, artificial structures such as bridges can serve as unwanted drainage barriers and represent causes of large sinks (Figure. 1.) In this case, manual editing of the DEM has to be performed according to the workflow illustrated in Figure 2

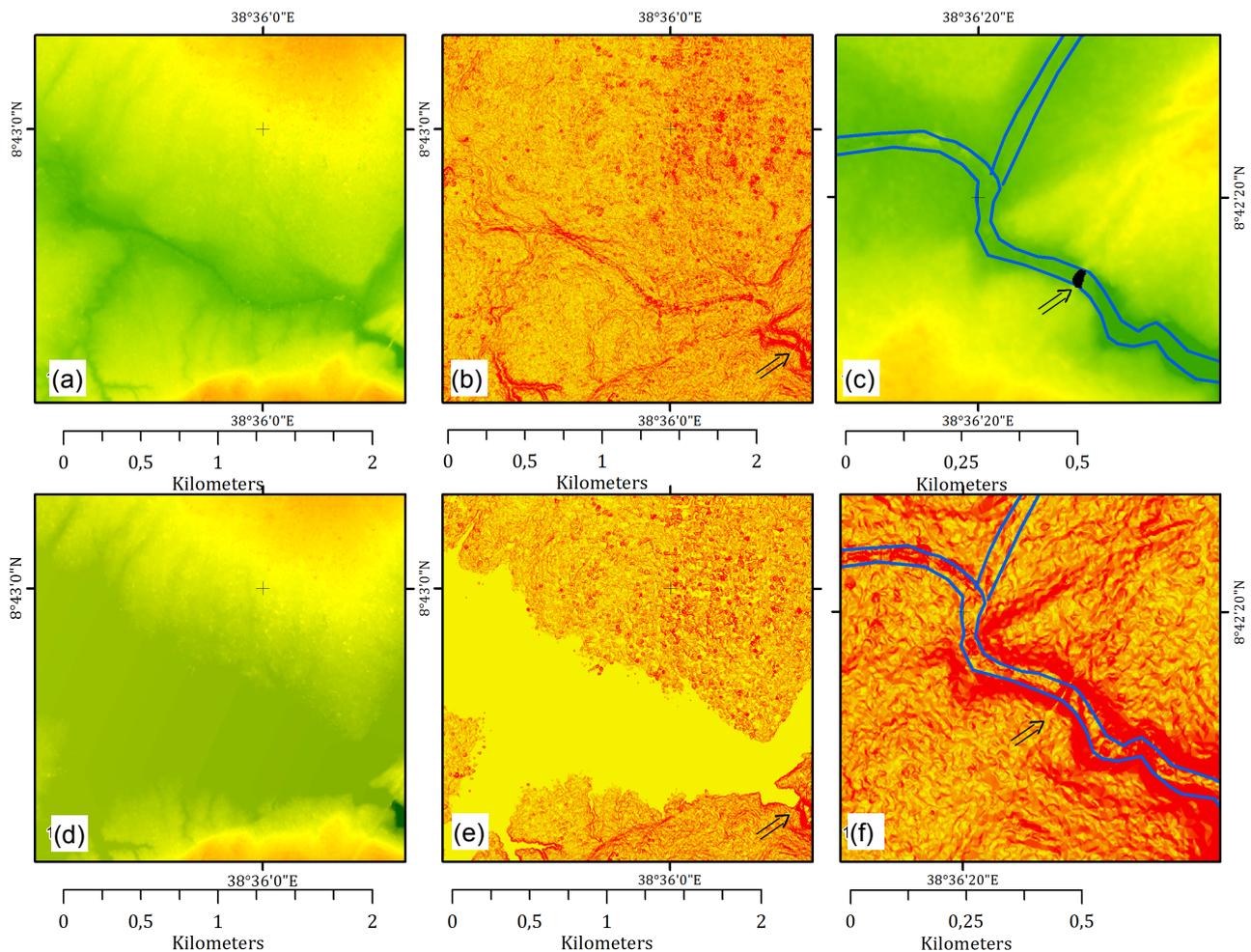


Figure 1: Pre-processing of the high resolution DEM: (a) the original SfM DEM, (b) the derived slope map, (c) shows the unwanted object, in this case represented by a bridge, that has to be removed to avoid (d) the SfM after Fill sink (Wang and Liu, 2006) and (e) the slope map derived with a huge "dam effect". Finally, (f) shows the result of the Fill sink after the tool "Burn stream network" in the DEM (SAGA).

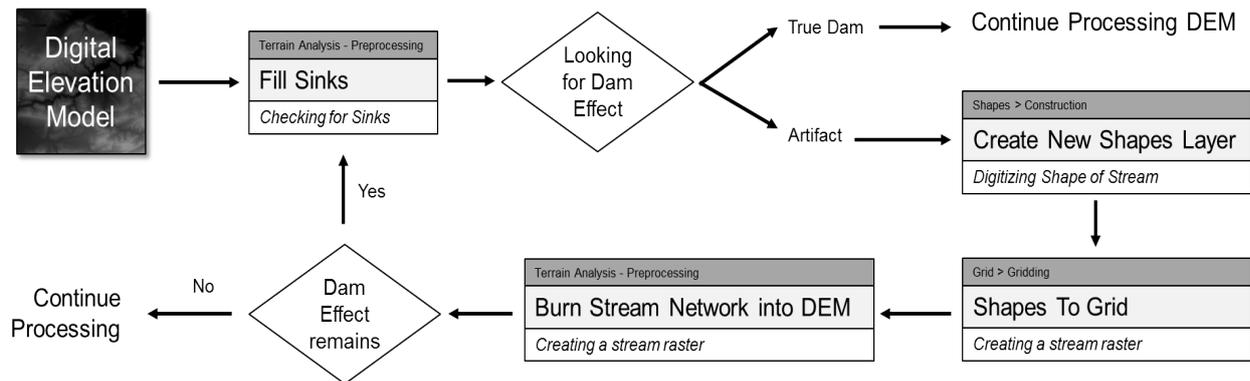


Figure 2: Workflow for the elimination of the "dam effect" in SAGA GIS.

In a first step to the removal of unwanted barriers, the erroneous barrier has to be identified in the DEM. Comparing the slope of the terrain before and after the removal of sinks (Figure 1) can show if errors were caused by artificial structures. These occur as extensive plains, mostly along valleys. In this case, the user has to digitize a new line (or polygon), which intersects the unwanted barrier. After converting the digitized vector into a grid, the SAGA module "Burn Stream Network into DEM" can be executed. It decreases the topography along the digitized line by a certain height. This decrease in altitude is expressed by the variable 'epsilon' in SAGA and has to be chosen by the user. Afterwards, the algorithm "Fill Sinks" will not interpret the area upstream of this structure as a basin any longer.

Overview of the most used terrain analysis attributes and landforms recognition methods

This section introduces the different domains of TA that are of interest to geomorphologists. This will start with basic topographic derivatives, determinants and terrain parameters, and later will proceed to more complex indices and hydrological parameters.

Digital topography

The topographic surface can be described as a continuous function, where z is the elevation and x and y are the Cartesian coordinates:

$$z = f(x, y) \quad \text{Eq. 1}$$

The simplest measure of topography is the **altitude** itself. It is important for the visibility

of topographic features and the distribution of temperature, rainfall or vegetation.

To enhance the interpretability of the elevation raster, a **shaded relief** or hillshade, is often generated first. Shaded relief is a visually pleasing representation of the terrain (Marston and Jenny, 2015). It assumes an illumination of the raster surface given at a defined direction and angle of the light source. Before the digital era, it was performed manually by darkening shaded areas in a map. Its application for analytical reasons has several possible calculations, which are outlined by Lukas and Weibel (1995).

Contour lines represent another possibility of visualizing the spatial distribution of elevations in the DEM. They are therefore often used in maps to give the reader an impression of the terrain without using too many colours or signatures. They can be generated based on a DEM by the minimum curvature approach suggested by Fogg (1985). However, the scale at which contour lines are generated is crucial as they significantly under-represent the areas between the chosen intervals, especially in areas of low relief (Hutchinson and Gallant, 1999). As a rule of thumb, the spacing between the contour lines should be at least twice the pixel size of the DEM (Hengl *et al.*, 2003).

Hydrology

One of the main questions in hydrology addresses the interaction of surface water and topography. Speight (1974) proposed a method for the digital delineation of watersheds. His concept of catchment area

(CA) calculation strongly influenced hydrological modelling; using slope and curvature to generate potential flow directions at the Earth's surface. These directions were then used to estimate borders of separate catchments. By adding up potential water flow (flow accumulation), a drainage network can be derived (Zeuberger and Thorne, 1987). This was already seen as a measure for landscape dissection by Horton (1945) and can be expressed as drainage density, indicating the total length of the streams related to an area (m/m^2).

These findings were later used to calculate hierarchical river systems, proposed by Strahler (1957) as the **Strahler order**. They are used to define stream size, based on a scaling hierarchy of tributaries. An indicator for the potential water content and horizon depth of the soil is the **Topographic Wetness Index (TWI)** by Beven *et al.* (1979), used for the quantification of topographic control on hydrological processes (Sørensen *et al.*, 2006). This is calculated by the following equation:

$$TWI = \ln \frac{CA}{\tan(G)}, \quad \text{Eq. 2}$$

where CA is the local upslope area draining through the cell and where $\tan(G)$ is the local slope in radians.

Morphometry

The two most important morphometric parameters are slope inclination and slope aspect; these are the basic flow attributes of the Earth's surface that are used to derive more complex features (Florinsky, 2012). **Slope gradient** (often referred to as *slope*) is the angle G between the tangent plane P and the horizontal plane S at a given point of the topographical surface (Lehmann, 1816). It is determined by the following equation:

$$G = \arctan \sqrt{p^2 + q^2}, \quad \text{Eq. 3}$$

Where $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$

Equation 3 is calculated in degrees, however percentage values are also common. However, it has to be kept in mind that 100% are corresponding to 45° because this is the angle in which vertical and horizontal differences are equal.

Slope aspect (referred to from now on as *aspect*) is a clockwise angle A from north to a projection of the external normal n to the horizontal plane S at a given point of the Earth's surface (Marida 1972). It determines measures of insulation, temperature, vegetation, soil characteristics and moisture.

$$G = -90(1 - \text{sgn}(q))(1 - |\text{sgn}(p)|) + 180[1 + \text{sgn}(p)] - \frac{180}{\pi} \text{sgn}(p) \arccos\left(\frac{-q}{\sqrt{p^2+q^2}}\right), \quad \text{Eq. 4}$$

It is measured in degrees, whereby 0° is equal to North and 180° is equal to South.

Local curvature is a measure of the surface roundness of an area. It can be divided into horizontal curvature k_h (*plan curvature*) and vertical curvature k_v (*profile curvature*) (Wilson and Gallant, 2000). **Plan curvature** is a measure of flow convergence ($k_h < 0$) and divergence ($k_h > 0$) and determines soil water or the deposition of particles. **Profile curvature** controls water flow acceleration ($k_h > 0$) and deceleration ($k_h < 0$) and therefore the erosion potential of an area (Shary, 1991). They can be described by the following equations:

$$k_h = \frac{p^2 r + 2pqs + q^2 t}{(p^2 + q^2)\sqrt{1+p^2+q^2}} \quad \text{Eq. 5}$$

$$k_v = \frac{q^2 r - 2pqs + p^2 t}{(p^2 + q^2)\sqrt{(1+p^2+q^2)^3}}, \quad \text{Eq. 6}$$

where $t = \frac{\partial^2 z}{\partial y^2}$ and $s = \frac{\partial^2 z}{\partial x \partial y}$.

Many other derivatives of curvatures exist (mean curvature, excess curvature, ring curvature or tangential curvature, for example) but they all rely on the same principles (Florinsky, 2012).

Morphology

DEMs can also be used to distinguish between different landforms. One way is to describe the **surface roughness** of areas. This parameter refers to the variability in elevation within a defined radius and is therefore very sensitive to the selected scale. It is used for the identification of coherent structures or underlying processes (Hobson, 1972). Many approaches have been developed to calculate surface roughness and each of them fits into a certain application. Grohmann *et al.* (2010) compare five different types of roughness: area ratio,

vector dispersion, the absolute standard deviations (SD) of elevation, slope and profile curvature. This was completed for different DEM resolutions to extract topographic information. They conclude that area ratio operates most independently of scale, therefore being more suitable when coarser resolution DEMs are applied.

Besides the relative comparison of surfaces according to their roughness, there are also algorithms, which directly address whole types of landforms. For example, the **Fuzzy landform element classification** is an unsupervised continuous classification method (McBratney and De Grujter, 1992). It was first implemented by Zadeh (1965). Since a landform's behaviour can be defined as continuous, the fuzzy classification algorithm can satisfactorily describe them (Zadeh, 1965). This classification creates clusters that can have a partial membership to the nearest cluster. Several parameters

are needed in order to perform the fuzzy classification algorithm, and, the presence of the elevation avert the misclassification of flat ridges from valleys (Irvin *et al.*, 1997). The identification of proper fuzzy rules for an unbiased mapping of terrain forms was shown by Haider *et al.* (2015) on example of mapping of peneplains.

Case study

The workflow presented in the case study will be extended to the most used terrain derivatives calculation Figure. 3, used for geomorphology investigations. The general TA workflow introduced previously showed what can be achieved in terms of interpretation of relief forms, feature extraction and hence the calculation of the basic terrain attributes for hydrological modelling in a friendly, used environment, such as SAGA.

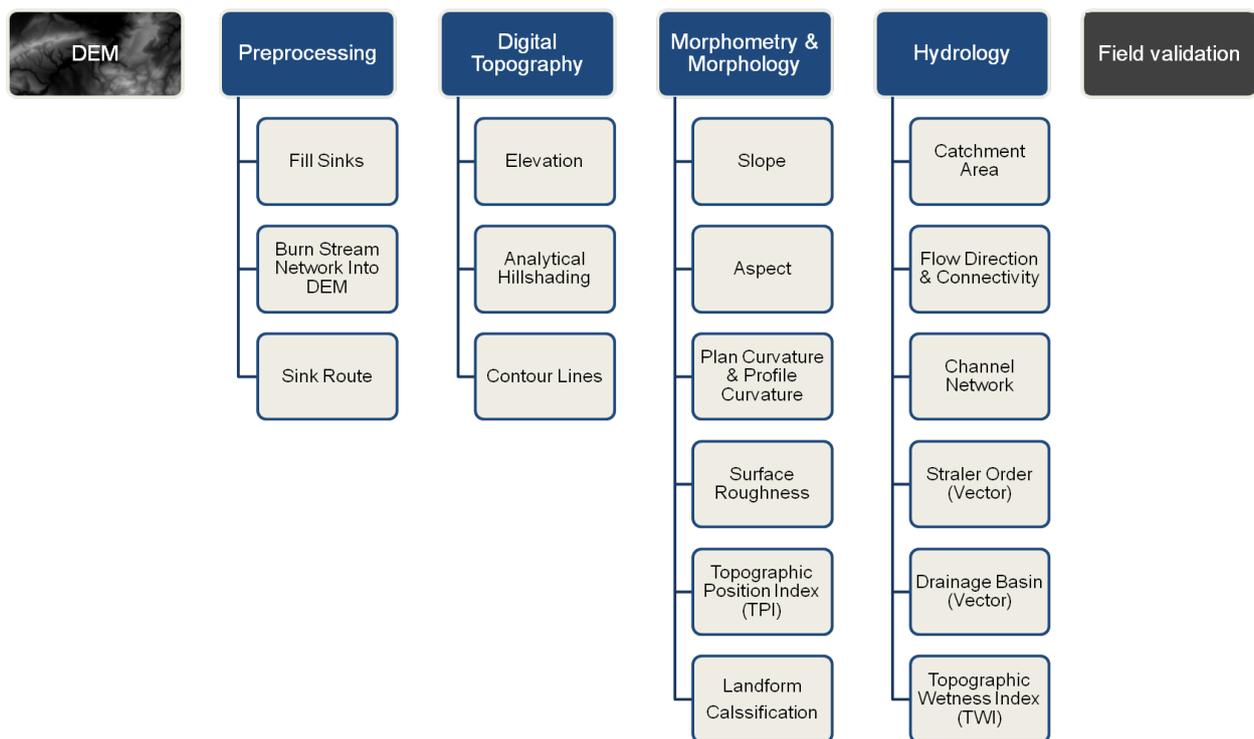


Figure 3: Workflow in SAGA GIS.

Three DEM sources were compared to assess their suitability for the automatic delineation of landforms and the terrain features at certain scales. Based on the case study a discussion of the workflow and calculation of various land surface parameters with the following DEMs:

- i. SRTM - 3 arcsec, 90 m
- ii. SRTM C - 1 arcsec, 30 m
- iii. SfM DEM - 2 m

The three DEMs were selected based on representative scales used in geomorphological mapping (Figure 1).

Furthermore, we intended to provide a visual comparison of 10 terrain derivatives.

34 aerial images taken in 1972 were used for this case study, obtained from the Ethiopian Mapping Agency EMA, as scanned photographs (1200 dpi) and were processed with the software PhotoScan Agisoft, (Professional Edition). With this methodology, camera positions and orientation were solved automatically. The resulting DEM allowed coverage of the whole study area.

The selected study area is located in the Upper Awash River Basin, Southwest of Addis Ababa, Ethiopia (Figure 4). This site was chosen due to: i) the form and features diversity (heterogeneity); and, ii) the

availability of a wide range of spatial scales DEM (from the middle to the high resolution). The study site is approximately 64 km² in area, with the altitude varying between 1850 and 2200 m above sea level.

Software

All the analyses in this article were performed in SAGA GIS; a software package licensed under the General Public Licence (GNU). The first release of SAGA was in 2004 at the Department for Physical Geography at the University of Göttingen. Two volume manuals have been provided by the SAGA working group (Cimmery, 2010a, 2010b), which give an insight to the software utilities

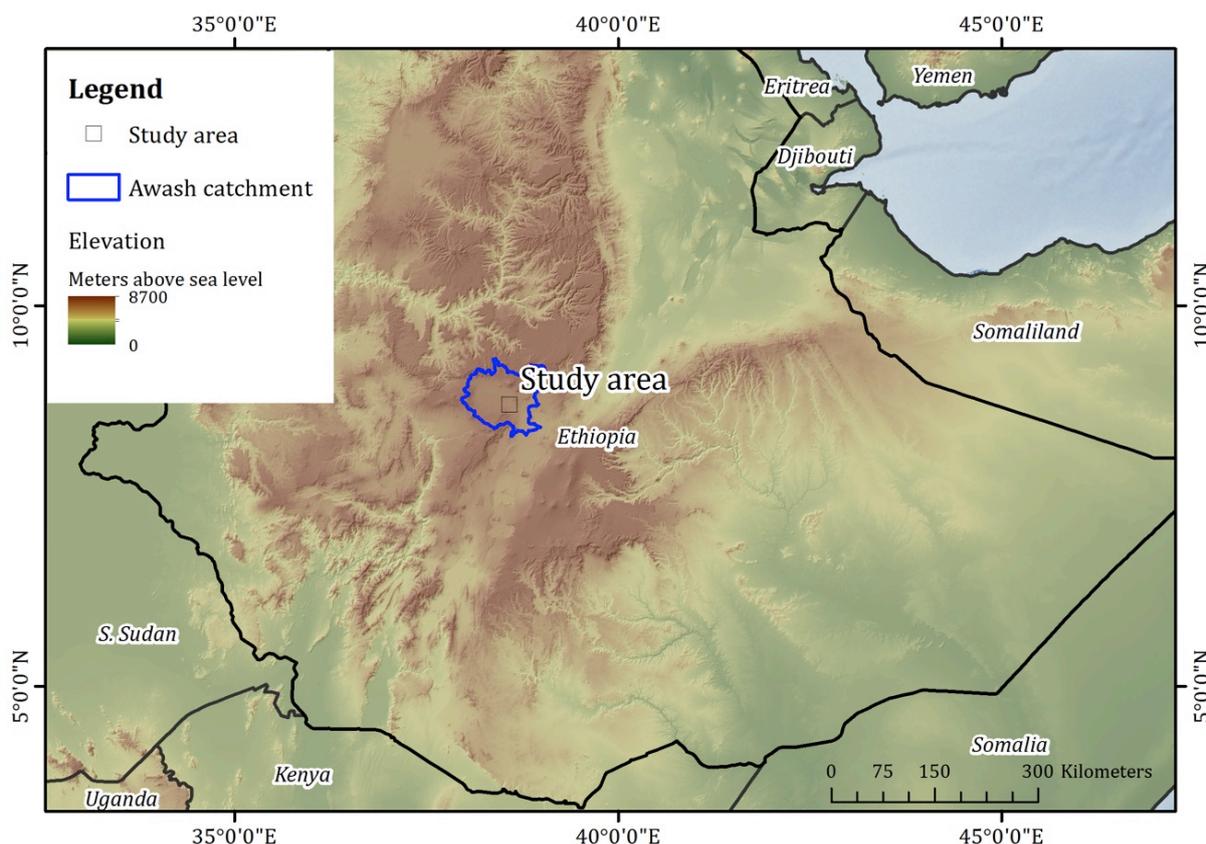


Figure 4: Study area of the case study

Results

Hillshade and contour lines

Terrain Analysis > Lighting > Analytical Hillshading
Shapes - Grid > Contour Lines from Grid

When performing a hillshade or slope function in any GIS, the elevation units are presumed to be the same as the horizontal

units. Therefore, working in a projected coordinate system based on meter units (e.g. UTM) is compulsory (Florinsky, 2012).

Current scientific publications often feature an automatically generated hillshade map (analytical hillshade) to quickly demonstrate relief features. Hereby a new raster is created illustrating lighter and darker areas, depending on the topography, which is easily recognizable as a 3D surface for the human eye

In contour lines, 10 m line spacing has been chosen. Figure 5 shows the shaded relief of the study area with contour lines; there is an increase in spatial resolution in Figure 5 (a, b and c).

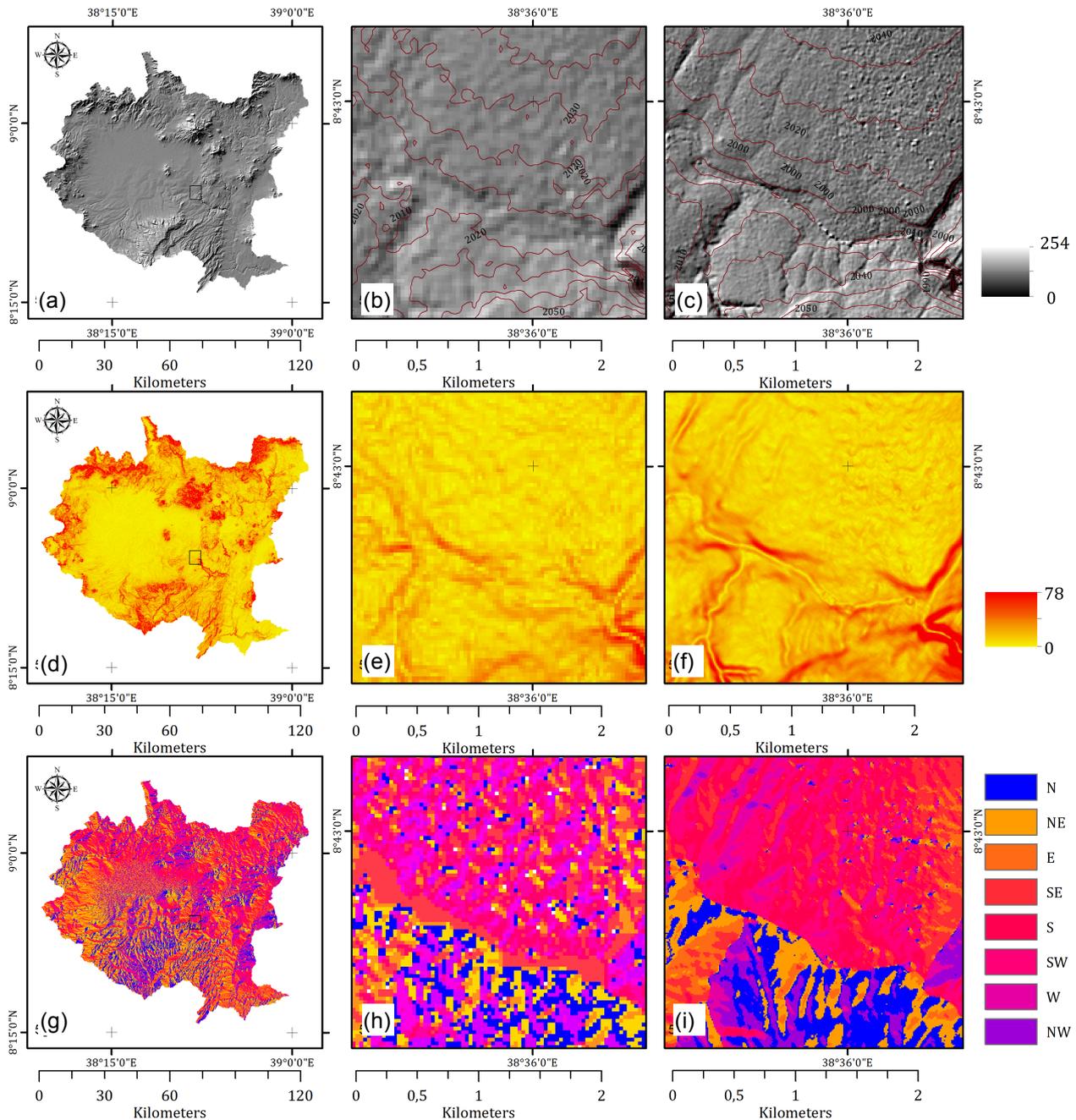


Figure 5. Hillshade maps of the Upper Awash catchment: (a) 90 m spatial resolution SRTM DEM; (b) SRTM-C (2014) 30 m spatial resolution with contour lines; (c) SfM DEM 2 m spatial resolution with contour lines. Slope maps: (d) 90 m spatial resolution SRTM DEM; (e) SRTM-C (2014) 30 m spatial resolution; (f) SfM DEM 2 m spatial resolution. Aspect maps: (g) 90 m spatial resolution SRTM DEM; (h) SRTM-C (2014) 30 m spatial resolution; (i) SfM DEM 2 m spatial resolution.

Slope

Terrain Analysis > Morphometry > Slope, Aspect, Curvature

Figures 5a, 5b and 5c show the gain in resolution of slope gradient with the increase in the DEM spatial resolution, with slope maps of the same resolutions shown in Figures 5d, 5e and 5f. It is an important

parameter for models that are used to predict water flows, flooding, erosion, construction, geology, insulation or soil depth.

Aspect

Terrain Analysis > Morphometry > Slope, Aspect, Curvature

The aspect map was reclassified following the division of 360 degrees into quadrants and sub dials:

N	:	0 - 22.5°
NE	:	22.5 - 67.5°
E	:	67.5 - 112.5°
SE	:	112.5 - 157.5°
S	:	157.5 - 202.5°

SW	:	202.5 - 247.5°
W	:	247.5 - 292.5°
NW	:	292.5 - 337.5°
N	:	337.5 - 359.5°

The latter produce 8 possible orientation classes (Figure 5g, 5h and 5i).

Plan and profile curvature

Terrain Analysis > Morphometry > Slope, Aspect, Curvature

Plan Curvature is the horizontal curvature, intersecting with the XY plane. Negative values coincide with concave features. These morphologies are normally produced and/or used for the water to flows. Figures 6a, 6b and 6c clearly shows the river network where blue (or negative curvature values) have been calculated.

Profile curvature is the curvature intersecting with the plane defined by the Z-axis and maximum gradient direction. Positive values describe convex profile curvature, with negative values representing a concave profile, and units are expressed in [1/m]. Figure 6d, 6e and 6f shows the river flowing where curvatures assume negative values, or alternatively where curvature is concave.

SD of elevation (Roughness)

Modules->Grid >Filter > Multi Direction Lee Filter > Standard deviation

Figure 6 (6g, 6h and 6i) demonstrates the roughness calculated from the elevation; it assumes high negative values for the incisions and lower values for flat terrains. The SD of elevation detects breaks in terrain

steepness and is therefore most suitable at regional scales while the SD of slope enhances noise in the data (such as caused by forest stands)

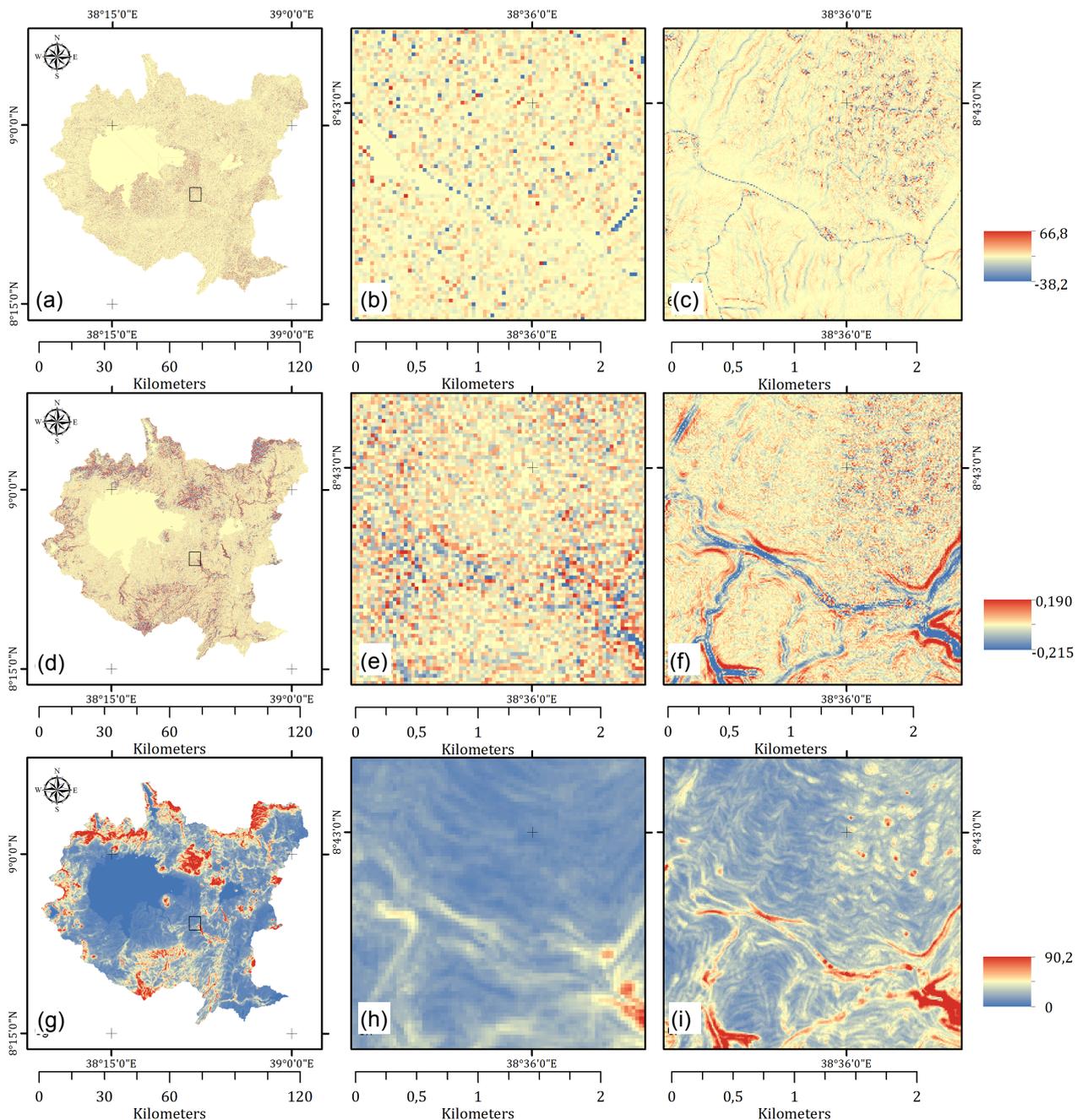


Figure 6. Plan curvature of the Upper Awash catchment at: (a) 90 m spatial resolution based on SRTM DEM; (b) SRTM-C DEM (2014) at 30 m spatial resolution; (c) SfM DEM at 2 m spatial resolution. Profile curvature at: (d) 90 m spatial resolution based on SRTM DEM; (e) SRTM-C DEM (2014) at 30 m spatial resolution; (f) SfM DEM at 2 m spatial resolution. Standard Deviation of elevation (roughness) maps based on: (g) 90 m spatial resolution based on SRTM DEM; (h) SRTM-C DEM (2014) at 30 m spatial resolution; (i) SfM DEM at 2 m spatial resolution.

Catchment area

Terrain Analysis>Hydrology > Catchment Area

Figure 7 (a, b and c) shows the potential flow directions and soil moisture distribution of the study area. The flow tracing algorithm was used, which traces the flow of each cell in a DEM separately until it finally leaves the DEM or ends in a sink. Since the amount of

moisture along the hillslope tends to increase from upslope to downslope, additional moisture contributed from upslope as the catchment area increases to the bottom of the valleys. Moisture distribution usually follows a logarithmic scale.

Strahler order

Terrain Analysis > Channel > Strahler order

Figure 7d, 7e and 7f) shows the Strahler order classification. Strahler scheme begins with the smallest channels being classified as 1st order streams. Under such a system the highest order generated from the classification is used to classify the drainage basin or the area of interest, for example the

Upper-Awash River can be considered as a seventh order. Good results in terms of perennial streams identification have been achieved for both the 90 m and 30 m DEMs (Figures 7d and 7e), whilst the headwater perennial streams are not well identified within the 2 m SfM DEM (Figure. 7f).

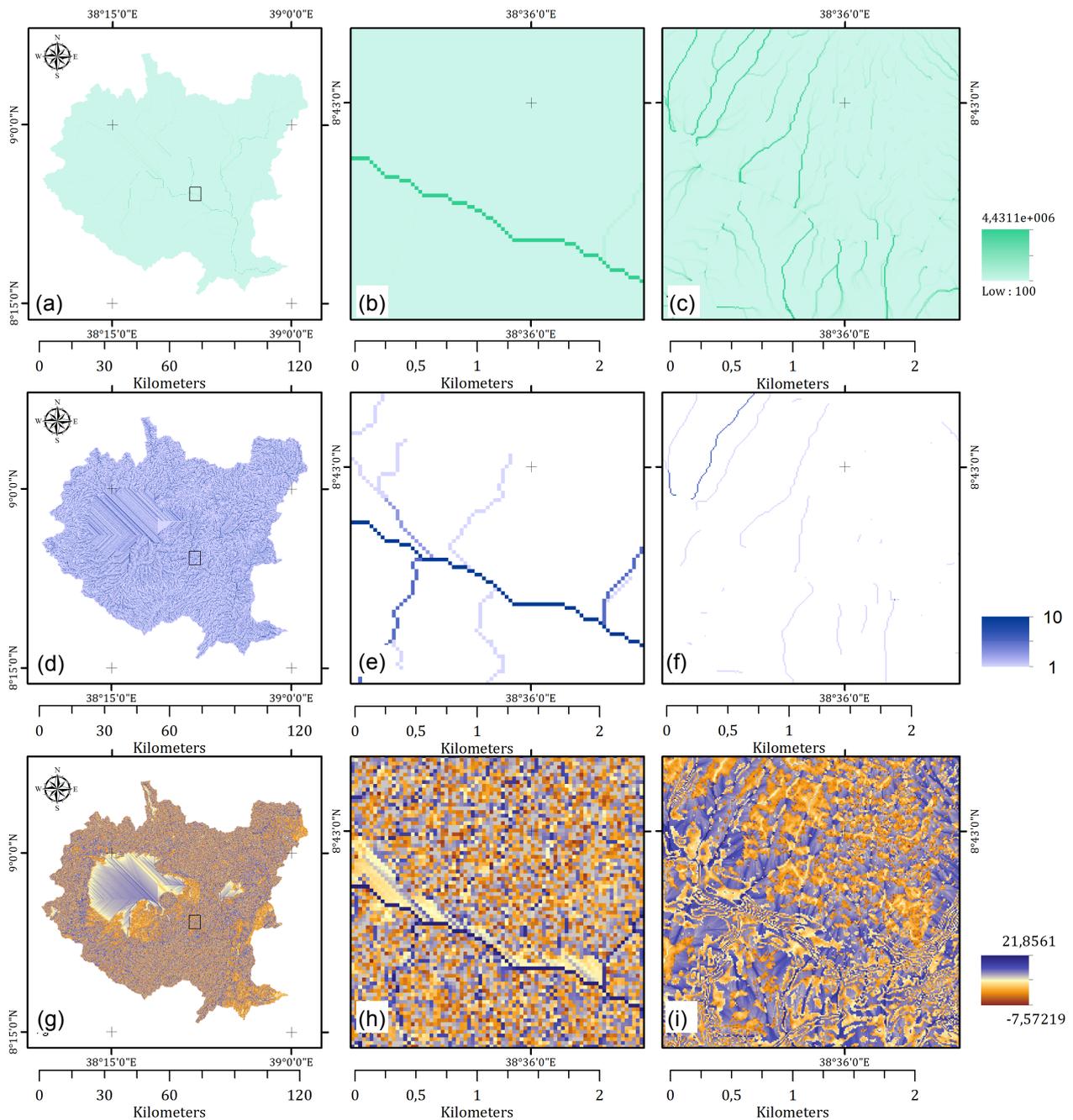


Figure 7. Catchment area of the Upper Awash catchment at: (a) 90 m spatial resolution based on SRTM DEM; (b) SRTM-C DEM (2014) at 30 m spatial resolution; (c) SfM DEM at 2 m spatial resolution. Strahler network at: (d) 90 m spatial resolution based on SRTM DEM; (e) SRTM-C DEM (2014) at 30 m spatial resolution; (f) SfM DEM at 2 m spatial resolution. Topographic wetness index TWI: (g) 90 m spatial resolution based on SRTM DEM; (h) SRTM-C DEM (2014) at 30 m spatial resolution; (i) SfM DEM at 2 m spatial resolution.

Topographic Wetness Index (TWI)

Terrain Analysis > Hydrology > Topographic Wetness Index

Figures 7g, 7h and 7i show the spatial distribution of wetness conditions that result from the contributing areas overland flow, the area drained per unit contour length (and the slope gradient, Eq. 2). It is used to quantify flow intensity and accumulation potential. The TWI index has to be used carefully to predict the distribution of dynamic phenomena, mainly because surface saturation is a threshold process and this diverge from the assumption of a parallel piezometric head

gradient (that dictates the direction of subsurface flow) to the land surface. In the reality, surface saturation measurement, would be useful to validate the calculation results.

The study area has been better characterized by the middle resolution SRTM 1arc sec topography, because the mostly flat terrain and the meandering water streams.

Fuzzy Element Landform Classification

Terrain Analysis > Terrain Classification > Fuzzy Element Landform Classification

The fuzzy landform classification is based on the continuous classification method, this groups each pixel according with the grade of membership. It is scale-dependent and identifies different morphometric features (peaks, ridges, passes, channels, pits and planes). The obtained landform classification map successfully shows the distribution of plains, peaks and steep slopes. Attributes used for the classification were elevation, slope, plan curvature, profile curvature, tangential curvature and both maximal and

minimal curvature. The continuous classification created 9 classes for the 90 m spatial resolution SRTM C DEM (Figure 8a), 11 classes for the 30 m spatial resolution SRTM C (2014) DEM (Figure 8b) and 2 m SfM DEM (Figure 8c). These are: Back Slope, Spur, Foot Slope, Foot Hollow, Foot Spur, Shoulder Slope, Shoulder Hollow, Shoulder Spur, Plain, Pit and Peak. The maximum membership function provided the degree of landforms characterization into a particular class. (see Figure 8).

Summary

Examples of a general TA workflow (Figure 3) shows a possible approach to calculate the primary and secondary terrain attributes. Selected attributes are presented below at the three chosen spatial resolutions to demonstrate the benefits in landform representation detail:

- i. SRTM 90 m shows its strength in the watershed delineation performance and in terms of regional landforms characterization;
- ii. SRTM 30 m improved significantly the TA resolution detail because of the better vertical resolution and led to an accurate determination of drainage slopes and precise locations of channels and ridges; and,
- iii. TA derived from the 2 m (SfM) DEM shows the fine terrain relief, and the smallest landforms can be derived

and offer the best solution for soil erosion modelling issues. It is also possible to delineate sub catchments and define minor drainage networks.

Conclusion

In this chapter, an overview of the particular fields of geomorphology in which TA is of crucial contribution was provided. This is only possible due to the increased availability of near-global DEMs on medium resolution (SRTM 1 arcsec, the forthcoming ALOS PRISM 5 m). The technological progress in terms of sensors and processing techniques (e.g. SfM) allow the transfer of already existing concepts to detailed scales, as well as the potential to develop new ones.

The availability of processing tools improved considerably within the last forty years. Tools

based on sophisticated algorithms are now in the freeware domain. They represent a user-friendly platform that can be used by geomorphologists with a limited knowledge in informatics. Fuzzy landform classification offers non-experts a method to provide information about landforms in terms of their quantitative distribution and degree of uncertainty. It demonstrates how meaningful, spatially coherent landform classes can be achieved from high resolution DEMs.

Temporal variations are encountered in all of the indexes calculations, especially for the highly dynamic areas, (e.g. Rift Valley, fluvial environments, glaciers, deserts). The TA

workflow proposed, provides the framework for the assessment of the geomorphological dynamics and can be applied to many heterogeneous environments and multiscale datasets.

It is of crucial importance for geomorphologists to pay close attention to the parameter that influences the hillslope stability, and more broadly to landscape evolution. Both the digital terrain analysis approach and the landform recognition required careful consideration, and even though they are computer based (digital), a significant part of the work has to be carried out in the field to validate the results.

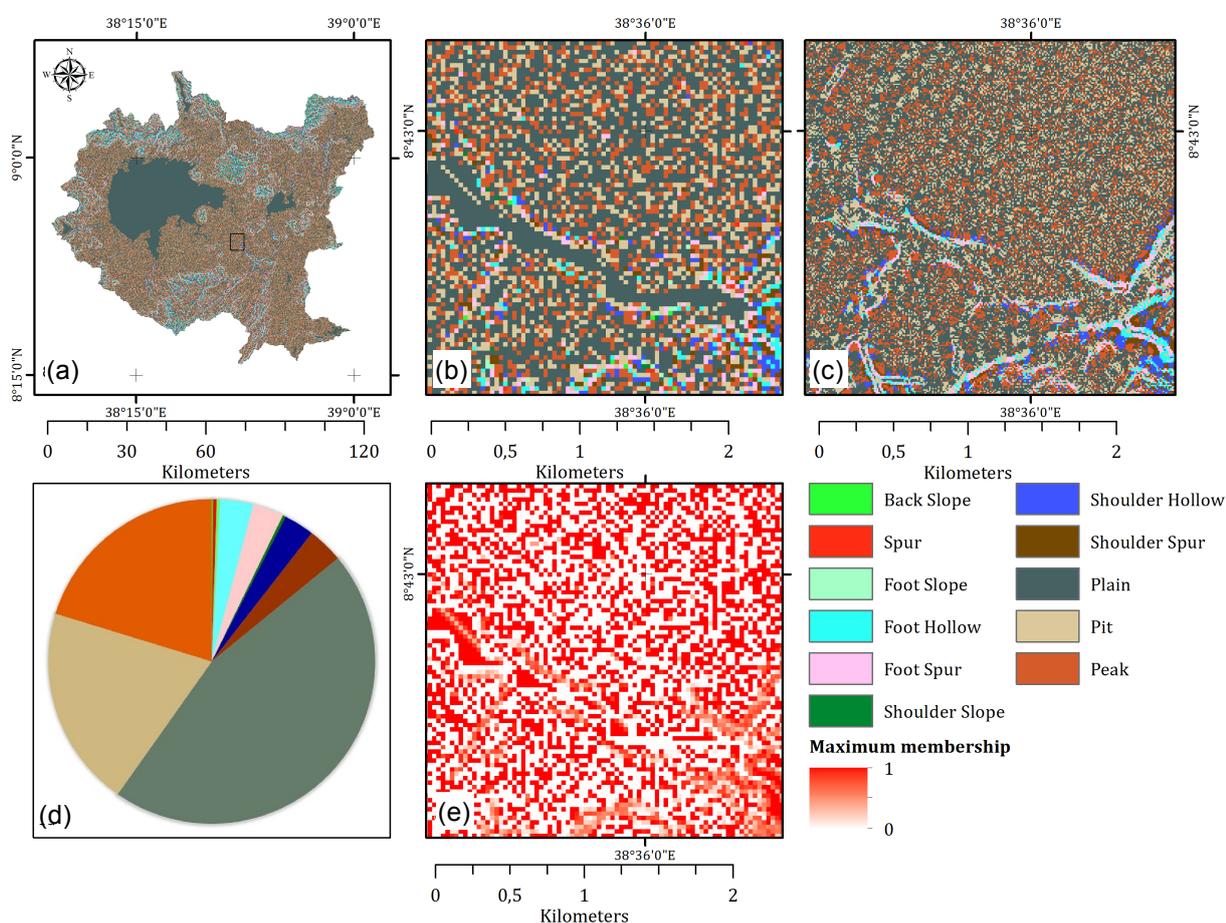


Figure 8. The Fuzzy landform classification of the Upper Awash catchment at: (a) 90 m spatial resolution (SRTM C); (b) SRTM-C DEM (2014) at 30 m spatial resolution; (c) SfM DEM 2 m spatial resolution; (d) pie chart showing the landform classification results expressed in percentage; and, (e) map of maximum membership - this map expresses the uncertainty of the classification, with this approach no pixels are left unclassified.

Acknowledgments

The authors thank the anonymous reviewers that helped us to refine this chapter. We also

express our gratitude to the Editors of this Book Dr Lucy Clarke. Data collection was carried out by Dr Jan Kropáček and Calogero Schillaci, in the frame of the project

"Integrated assessment of geomorphological processes dynamics on different spatio-temporal scales in the Ethiopian Highlands using remote sensing and advanced modelling approaches" financed by the DFG-Deutsche Forschungsgemeinschaft (Grant number: HO-1840/11-1). Sincere thanks are also given to Prof. Volker Hochschild, Dr. Michael Märker and our Computer administrator Christian Bick, at the Geosciences Department, University of Tuebingen.

References

- Abera W, Antonello A, Franceschi S, Formetta G, Rigon R. 2014. Section 2.4.1: The uDig Spatial Toolbox for hydro-geomorphic analysis. In: Clarke L, Nield J. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology: London, UK.
- Band LE. 1993. Extraction of channel networks and topographic parameters from digital elevation data In: Beven K, Kirkby MJ (Eds.) *Channel Network Hydrology*. Wiley: New York, USA, 13-42.
- Band LE. 1986. Topographic partition of watersheds with digital elevation models. *Water Resources Research* **22**: 15-24.
- Beven KJ, Kirkby MJ, Seibert J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin* **24**: 43-69.
- Brabyn L. 1998. GIS analysis of macro landform. *The 10th Colloquium of the Spatial Information Research Centre: University of Otago, New Zealand*, 35-48.
- Chairat S, Delleur JW. 1993. Integrating a physically based hydrological model with GRASS. In: Kovar K, Nachtnebel HP. (Eds.) *HydroGIS 93: Application of Geographical Information Systems in Hydrology and Water Resources*. International Association of Hydrological Sciences, IAHS Publ. No. 211: Vienna, 143-150.
- Cimmery V. 2010a. *SAGA User Guide, updated for SAGA version 2.0.5*. Volume 1: An introduction to the graphical user interface of SAGA. Online common licenced.
- Cimmery V. 2010b. *SAGA User Guide, updated for SAGA version 2.0.5*. Volume 2: 'How To' information on many SAGA modules, functions, and GIS applications. Online common licensed.
- Dehn M, Gartner H, Dikau R. 2001. Principles of semantic modeling of landform structures. *Computers and Geosciences* **27(8)**: 1005-1010.
- Dikau R. 1989. The application of a digital relief model to landform analysis in geomorphology. In: Raper J. (Ed.) *Three Dimensional Application in Geographic Information Systems*. Taylor & Francis: London, UK, 51-77.
- Evans IS. 2012. Geomorphometry and landform mapping: What is a landform? *Geomorphology* **137(1)**: 94-106. DOI:10.1016/j.geomorph.2010.09.029.
- Farr TG, Rosen PA, Caro, E, Crippen R, Duren R., Hensley S, Kobrick, M, Paller, M, Rodriguez, E, Roth, L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D. 2007. The Shuttle Radar Topography Mission. *Review of Geophysics* **45**: RG2004. DOI:10.1029/2005RG000183.
- Florinsky I. 2012. *Digital Terrain Analysis in Soil Science and Geology*. Elsevier Academic Press: Oxford, UK.
- Fogg GA. 1985. Contour to rectangular grid conversion using minimum curvature. *Computer Vision, Graphics, and Image Processing* **28**: 85-91.
- Garbrecht, J, Martz LW. 2000. Digital elevation model issues in water resources modeling. In: Maidment D, Djokic D. (Eds.) *Hydrologic and Hydraulic Modeling Support with Geographic Information Systems*. ESRI Press: Redlands, USA, 1-28.
- Gilvear DJ, Bryant RG, Hardy TB. 1999. Application of remote sensing to the study of channel morphology. *Progress in Environmental Science* **1(3)**: 257-284.
- Goudie AS. 2004. *Encyclopaedia of Geomorphology*. Routledge: London, UK.
- Graf WL. 1987. *Geomorphic systems of North America*. Geological Society of America: Boulder, USA.
- Grohmann CH, Smith MJ, Riccomini C. 2010. Multiscale Analysis of Topographic Surface Roughness in the Midland Valley, Scotland. *Geoscience and Remote Sensing* **49(4)**: 1-14.

- Haider, VL., Kropáček, J, Dunkl, I, Wagner, B, von Eynatten, H. 2015. Identification of peneplains by multi-parameter assessment of digital elevation models. *Earth Surface Processes and Landforms*, (in print). doi: 10.1002/esp.3729
- Hengl TG, Gruber S, Shrestha DP. 2003. *Digital Terrain Analysis in ILWIS*. Lecture Notes and User Guide: https://www.itc.nl/library/Papers_2003/misc/hengl_digital.pdf
- Hobson RD. 1972. Surface roughness in topography: quantitative approach. In: Chorley RJ (Ed.) *Spatial Analysis in Geomorphology*. Methuen: London, UK, 225-245.
- Horton RE. 1945. Erosional development of streams and their drainage basins, hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin* **56**: 275-370.
- Hutchinson, MF, Gallant, JC, 1999. Representation of terrain. In: Longley PA, Goodchild MF, Maguire DJ, Rhind DW. (Eds.) *Geographical Information Systems: Principles, Technical Issues, Management Issues and Applications (2nd Edition)*. Wiley: New York, USA, 105-124.
- Irvin BJ, Ventura SJ, Slater BK. 1997. Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. *Geoderma* **77**: 137-154.
- Javernick L, Brasington J, Caruso B. 2014. Modelling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology* **213**: 166-182. DOI: 10.1016/j.geomorph.2014.01.006.
- Jenson SK, Domingue JO. 1988. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. *Photogrammetric Engineering and Remote Sensing* **54(11)**: 1593-1600.
- Klimaszewski M. 1982. Detailed geomorphological maps. *ITC Journal* **1982-3**: 265-271.
- Large A, Heritage G. 2012. Ground based LiDAR and its application to the characterisation of fluvial forms. In: Carbonneau PE, Piégay H. (Eds.) *Fluvial Remote Sensing for Science and Management*. John Wiley & Sons Ltd: Chichester, UK. DOI: 10.1002/9781119940791.ch14
- Lehmann JG. 1816. Die Lehre der Situation-Zeichnung, oder Anweisung zum richtigen Erkennen und genauen Abbilden der Erdoberfläche in topographischen Karten und Situation-Planen.
- Lukas K, Weibel R. 1995. Assessment and improvement of methods for analytical hillshading. *Proceedings of the 17th International Cartographic Conference* Barcelona, Spain: 2231-2240: http://icaci.org/files/documents/ICC_proceedings/ICC1995/PDF/Cap430.pdf
- McIlroy de la Rosa J. 2012. Section 6.1: Karst Landform Classification Techniques In: Clarke LE, Nield J.M. (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology: London, UK.
- McBratney AB, De Grujter JJ. 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *Journal of Soil Science* **43**: 159-175.
- Marida KV. 1972. Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)* **37(3)**: 349-393.
- Märker M, Pelacani S, Schröder B. 2011. A functional entity approach to predict soil erosion processes in a small Plio-Pleistocene Mediterranean catchment in Northern Chianti, Italy. *Geomorphology* **125**: 530-540.
- Marston BE, Jenny B. 2015 Improving the representation of major landforms in analytical relief shading. *International Journal of Geographical Information Science*: DOI:10.1080/13658816.2015.1009911.
- Marzoff I, Poesen J. 2009. The potential of 3D gully monitoring with GIS using high-resolution aerial photography and a digital photogrammetry system. *Geomorphology* **111**: 48-60.
- Milne JD, Clayden B, Singleton PL, Wilson AD. 1995. *Soil Description Handbook*. Manaaki Whenua Press, Landcare Research: Lincoln, New Zealand.
- Minár J, Evans IS, Krcho J. 2013. Geomorphometry: Quantitative Land-Surface Analysis. In: Schroder JF (Ed.) *Treatise on Geomorphology*. Academic Press: San Diego, USA, 22-34. DOI:10.1016/B978-0-12-374739-6.00370-5.
- O'Callaghan JF, Mark DM. 1984. The extraction of drainage networks from digital

- elevation data. *Computer Vision, Graphics, and Image Processing* **28**: 323-344.
- Pavlopoulos K, Evelpidou N, Vassilopoulos A. 2009. *Mapping Geomorphological Environments*. Springer: Berlin, Germany.
- Peucker TK, Douglas DH. 1975. Detection of surface-specific points by local parallel processing of discrete terrain elevation data. *Computer Graphics and Image Processing* **4**: 375-387.
- Pike, RJ, Evans IS, Hengl T. 2008. Geomorphometry: a Brief Guide. In: Hengl, T. and Reuter, H.I. (Eds), *Geomorphometry: Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, vol. 33, Elsevier, 1-28 pp.
- Prima OD, Echigo A, Yokoyama R, Yoshida T. 2006. Supervised landform classification of Northeast Honshu from DEM-derived thematic maps. *Geomorphology* **78**: 373-386.
- Ramasamy SM, Paul MA. 2005. Aid of remote sensing in fluvial geomorphic mapping. In: Ramasamy SM. (Eds.) *Remote Sensing in Geomorphology*. New India Publishing Agenda: New Dehli, India, 98-106.
- Schumm SA. 1979. 'Ergodic' reasoning in geomorphology: time for a review of the term? *Transactions of the Institute of British Geographers* **4(4)**: 485-515.
- Shary PA. 1991. The second derivative topographic method. In: Stepanov I. (Ed.) *The Geometry of the Earth Surface Structures*. Pushchino Research Centre Press: Pushchino, USSR.
- Sidorchuk A, Märker M, Moretti S, Rodolfi G. 2003. Gully erosion modelling and landscape response in the Mbuluzi River catchment of Swaziland. *Catena* **50**: 507-525.
- Skidmore AK. 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographical Information Systems* **3(4)**: 323-334.
- Smith MJ, Rose J, Gousie MB. 2009. The Cookie Cutter: A method for obtaining a quantitative 3D description of glacial bedforms. *Geomorphology* **108(3-4)**: 209-218. DOI:10.1016/j.geomorph.2009.01.006.
- Smith TR, Zhan C, Gao P. 1990. A knowledge-based, two-step procedure for extracting channel networks from noisy DEM data. *Computer Geoscience* **16**: 777-786.
- Sørensen RZ, Zinko U, Seibert J. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences* **10**: 101-112.
- Speight JG. 1990. *Australian Soil and Land Survey: Field Handbook* (2nd Edition). Inkata Press: Melbourne, Australia.
- Speight JG. 1974. A parametric approach to landform regions. In: Brown EH, Waters RS. (Eds.) *Progress in Geomorphology: Papers in honour of D.L. Linton*. Institute of British Geographers: London, UK, 213-230.
- Strahler AN. 1957. Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* **33**: 913-920.
- Tadono T, Ishida H, Oda F, Naito S, Minakawa K, Iwamoto H. 2014. Precise global DEM generation by ALOS PRISM. *ISPRS Annals of the Photogrammetry and Remote Sensing Spatial Information Science II-4*: 71-76.
- Thornbury WD. 1965. *Regional Geomorphology of the United States*. Wiley: New York, USA.
- Valentin C, Poesen J, Yong Li. 2005. Gully erosion: Impacts, factors and control. *Catena* **63 (2-3)**: 132-153.
- Wang L, Liu H. 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *International Journal of Geographical Information Science* **20**: 193-213.
- Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 2012. Structure-from-Motion" photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **179**: 300-314.
- Wieczorek GF, Snyder JB. 2009. Monitoring slope movements. In: Young R, Norby L (Eds.) *Geological Monitoring*. Geological Society of America: Boulder, USA, 245-271.
- Wilson JP. 2012. Digital terrain modeling. *Geomorphology* **137**: 107-121.
- Wilson JP, Gallant JC. 2000. *Terrain Analysis: Principles and Applications*. John Wiley and Sons: New York, USA.
- Zadeh LA. 1965. Fuzzy Sets. *Information and Control* **8**: 338-358.

Zakerinejad R, Märker M. 2014. Prediction of gully erosion susceptibilities using detailed terrain analysis and maximum entropy modeling: a case study in the Mazayejan Plain, Southwest Iran. *Geografia Fisica e Dinamica Quaternaria* **37**: 67-76.

Zeveloff LW, Thorne CR. 1987. Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* **12**: 47-56.

Zhang W, Montgomery DR. 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations, *Water Resources Research* **30**: 1019-1028.

Sitography

Following a list of useful URLs:

CGIAR (Consortium for Spatial Information):
<http://srtm.csi.cgiar.org/>

EROS (NASA): <http://eros.usgs.gov/>

ESA/AIRBUS: <http://www.geo-airbusds.com/>

NASA: <http://www2.jpl.nasa.gov/srtm/>

USGS (Earth Explorer):
<http://earthexplorer.usgs.gov/>

Geomorphological mapping

Jan-Christoph Otto¹ and Mike J. Smith²

¹Department of Geography and Geology, University of Salzburg, Austria (jan-christoph.otto@sbg.ac.at)

²Department of Geography, Geology and Environment, Kingston University London (michael.smith@kingston.ac.uk)



ABSTRACT: Geomorphological mapping is regarded as a fundamental technique of the discipline producing valuable base data for geomorphological and environmental research and practice. Geomorphological maps can be considered graphical inventories of a landscape depicting landforms and surface as well as subsurface materials. Geomorphological mapping is a preliminary tool for land management and geomorphological risk management, also providing baseline data for other sectors of environmental research such as landscape ecology, forestry or soil science. The widespread distribution and extended graphical capabilities of GIS-software as well as the availability of high-resolution remote sensing data such as aerial and satellite imagery or digital elevation data has led to a rejuvenation of the method. This chapter outlines the history, creation and dissemination of geomorphological maps; it provides a brief overview of field and digital mapping techniques, as well as introductory information on cartographic principles applied to geomorphological mapping. Finally, there are a range of key references that can aid the reader by providing more in-depth information on selected topics.

KEYWORDS: map, GIS, cartography, DEM

Introduction

Geomorphological maps can be considered graphical inventories of a landscape depicting landforms and surface as well as subsurface materials. Sketches and maps of landscapes and landforms (e.g. Dykes, 2008) have been fundamental methods to analyse and visualise Earth surface features ever since early geomorphological research. The widespread distribution and extended graphical capabilities of geographic information systems (GIS) as well as the availability of high-resolution remote sensing data such as aerial and satellite imagery and digital elevation models (DEMs) has led to the recent rejuvenation of the method (Lee, 2001, Paron and Claessens, 2011, Smith et al., 2011). Geomorphological maps can act as a preliminary tool for land management and geomorphological and geological risk management, as well as providing baseline data for other applied sectors of environmental research such as landscape ecology, forestry or soil science (Cooke and

Doornkamp, 1990, Dramis et al., 2011, Paron and Claessens, 2011).

Geomorphological maps can be categorised as *basic* or *analytical* and *derived* or *specialised*. Whilst basic maps represent the observed features of a landscape, derived maps are focused on a specific theme or application. One example of derived maps are geomorphological hazard maps that depict risk-causing phenomena and their magnitude and frequency (Dramis, et al., 2011). Basic geomorphological maps may either focus on selected landscape features, for example only depicting the morphology of active processes, or deliver a full view on the landscape composition and its evolution (Knight et al., 2011, Verstappen, 2011).

In contrast to other types of geoscientific maps, very little international standardisation exists for geomorphological mapping legends. In the second half of the 20th century different legend systems developed in

various European countries (Barsch and Liedtke, 1980, Brunsden et al., 1975, Embleton and Verstappen, 1988, Evans, 1990, Klimaszewski, 1982, Pellegrini, 1993, Schoeneich et al., 1998, Tricart, 1965). The choice of mapping symbols (or legend system) is determined by the purpose of the map, the message to communicate and the targeted group of users. It is therefore important to analyse and plan these conditions before starting a mapping campaign. The application of complex symbols using GIS or graphic software may influence the selection of the legend system. This chapter outlines the techniques for the creation and dissemination of geomorphological maps; it provides a brief overview of field and digital mapping techniques, as well as introductory information on cartographic principles applied to geomorphological mapping. Finally, there are a range of key references that can aid the reader by providing more in-depth information on selected topics.

Field Mapping

Despite the growing offer of high resolution digital information on the Earth's land surface the visual impression gained from direct observation in the field provides unprecedented detail and access. Field observation, and subsequent mapping, allows the most direct way to appreciate a landscape's character and enables a basis for terrain assessment and geomorphological analysis. Though field mapping by its nature is subjective and affected by the skills (experience) of the mapper, it allows the mapper to become familiar with the landscape (Cooke and Doornkamp, 1990). This is a crucial aspect as it enables exploratory investigation through direct observation of surface morphologies and subsurface exposures; this allows the development of a "mental model" of the landscape which can be incorporated in to later interpretation and analysis.

The production of good and purposeful maps in the field requires a clear definition of the aims and objectives. The mapping procedure includes (i) pre-field preparatory steps, (ii) the actual field work and (iii) post mapping activities that result in the creation of a publishable final map (*Table 1*).

Table 1. Workflow for undertaking geomorphological field mapping (modified from: Knight et al. 2011)

Activity

Pre-Mapping:

- Identify region of interest
- Identify purpose or goal of mapping
- Obtain remote sensing data
- Obtain geological and soils mapping information
- Design and create a GIS database
- Compose a field mapping protocol
- Map major morphological forms using remote sensing data
- Create draft map at a suitable scale for field mapping
- Prepare legend systems and symbols
- Obtain permission for access to the mapping region
- Conduct risk assessment for the planned activities
- Obtain weather forecast

During the field campaign:

- Conduct field mapping following the protocol, including walking the area
- Use hand-held GPS to mark tracks or waypoints
- Write notes and take photos, positioned using GPS
- Adhere to health and safety issues and/or update the risk assessment

Post-Mapping:

- Download and integrate GPS data with the existing GIS database
 - Compare field and remote sensing mapping in order to validate remotely sensed observations
 - Write up notes and integrate notes with photos
 - Produce final geomorphological map in the GIS software
 - Draw final map using analogue or digital cartographic symbols either using the GIS software, or graphic design software
 - Write and present explanatory notes accompanying the map
 - Publish the map
-

Typical preparation includes the gathering of additional information on the area (location, accessibility, history, and previous work), the choice of mapping scale and symbols, as well as planning the mapping campaign (Knight, et al., 2011). The analysis of aerial imagery or DEMs helps to familiarise the observer with the field area and should always be performed before mapping. This results in a draft map that includes major landforms mapped from the remote sensing data. The draft map is taken into the field for verification, addition of detail and delineation of exact boundaries. The creation of a field mapping protocol that details the steps of the mapping campaign is important as it allows the organised collection of additional information on landforms and the creation of a database for the data mapped in the field (Knight, et al., 2011).

The mapping scale used for data recording is usually larger than the scale of the final map in order to collect as many details in the field as possible. Typical detailed field maps are created at a mapping scale of 1:3000 to 1:25,000, whilst detailed geomorphological maps have scales between 1:5000 and 1:50,000. Consequently, map production requires generalisation of the field data. It is worth noting that medium and small scale geomorphological maps exist, ranging from 1:100,000 to > 1:1,000,000 (Dramis, et al., 2011).

Key equipment

Paper based (analogue) field mapping is usually performed using tracing paper, a topographic map of the area of interest, a clipboard and a pencil. The topographic map is used for orientation and location and should be enlarged to the mapping scale. Aerial photos and/or relief shaded (see below) images at the same scale are also useful. Other tools include a global positioning system (GPS) receiver for the collection of waypoints or the exact positioning of boundaries and objects. Additionally, binoculars are useful particularly in mountain areas to investigate inaccessible terrain. In order to have all the required symbols at hand, the legend system of the symbols should be printed out and taken into the field.

When using digital devices such as mobile field computers, mapping is generally performed using field mapping software. This allows for the digitisation of vector data directly onto the screen. The base data for mapping (topographic map, aerial photo, DEM) are stored on the field device. For exact positioning the field computer should be connected to a GPS. The application of a digital mapping device requires the preparation of the GIS data base structure prior to the field campaign (Figure 1).

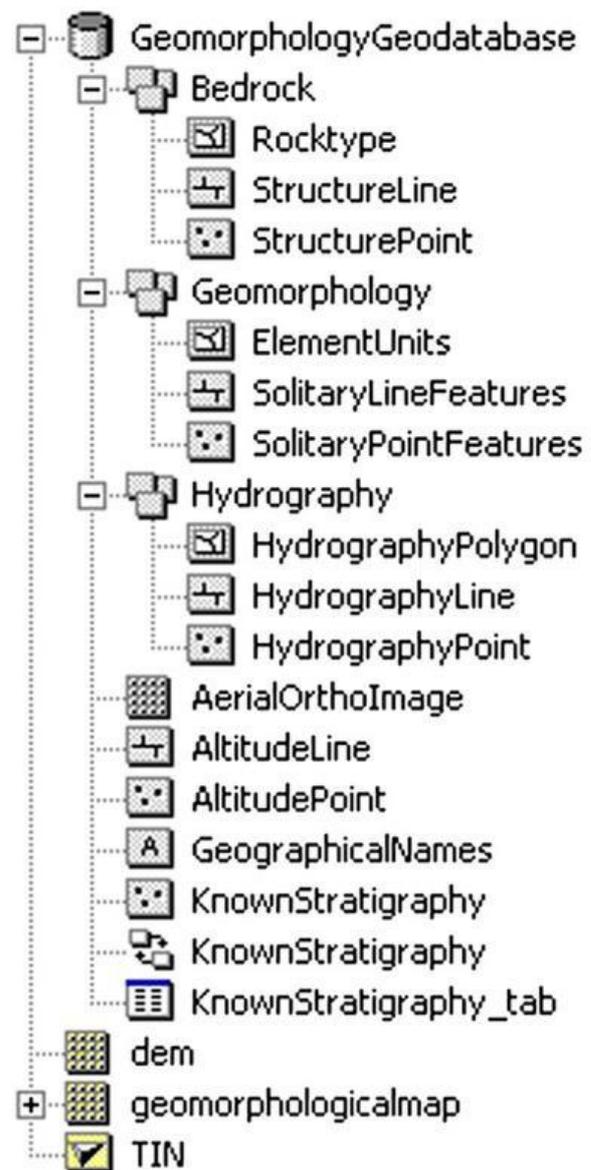


Figure 1. A typical GIS data base structure for a geomorphological mapping project (after Gustavsson et al., 2008).

Mapping procedure

A typical field mapping campaign starts with a 'walk through' to generate an overview of the area. This enables an estimate of the time for mapping and planning of the route through the study area in accordance with the field protocol. The mapping itself starts with the marking of breaks of slopes or a verification of the draft map contents. One method for avoiding overload of the field map sheets is to use separate sheets of tracing paper for each layer of information; for example morphology, process, and surface material. Mapping of processes usually requires interpretation from surface form and material composition since few active processes will be observed 'live' in the field. Mapping of surface and subsurface material is usually performed by visual interpretation or rapid field sampling. The goal of material mapping lies in the identification of material patterns rather than detailed sediment analysis (Dackombe and Gardiner, 1983).

Mapping speed depends upon the experience of the observer and the ground conditions. Even though mapping in mountainous regions may be more time consuming due to the rough terrain, mountains/hills often provide a better view of the landscape compared to flat terrain that needs to be walked through completely in order to see all features. An experienced mapper is able to map up to 2-3 km²/day, depending on local conditions and the complexity of terrain.

Digital Mapping

Where field mapping is not practical, it is common practice to utilise digital data sources (Smith and Clark, 2005, Smith and Pain, 2009, Oguchi et al., 2011, Smith, 2011) to compile mapping within the framework of a GIS. Primary data sources include aerial photography, satellite imagery and DEMs - care should be taken in the selection of data so that it is at an appropriate scale and provides sufficient coverage of the study area. The user then has two options for performing mapping: (1) manual and (2) automated or semi-automated.

Manual Mapping

Manual approaches rely on the experience and skill of an individual mapper using visual heuristics to identify landforms of interest (Colwell, 1983); accuracy is generally high, the method is simple and rapid to deploy, however this is at the expense of an objective and repeatable approach. The *representation* of a landform on an image is dependent upon: (i) the landform itself, (ii) the data source and (iii) the method used to process it (visualisation method).

For satellite imagery a range of standard image enhancement techniques can be applied; these include false colour composites, band ratios, convolution filtering and contrast stretches. However there are three main controls on landform representation (Smith and Wise, 2007): (i) relative size: the size of the landform relative to the spatial resolution, (ii) azimuth biasing: the orientation of the landform with respect to solar azimuth and (iii) landform signal strength: the tonal/textural differentiation of the landform. For any specific image used for mapping there is a minimum resolvable landform size and a range of landform orientations (which is more pronounced the greater the linearity) at which they can be represented. The definition of these landforms is then determined by their relative reflectance in comparison to surrounding features. Smith and Wise (2007) recommend acquisition of imagery with solar elevation angles <20° in order to enhance the topographic "signal" through shadowing.

For DEMs only the issue of relative size holds true and as a result they should offer more refined visualisations for mapping (Hillier and Smith, 2008, Smith and Clark, 2005, Smith et al., 2013). DEM processing is a subject area of itself known as geomorphometry (Hengl and Reuter, 2008), with land surface parameters (LSPs) the processed "derivatives" of the raw elevation data; these are used to present visualisations for manual mapping.

The most common method is *relief shading* which mimics a satellite image by placing an artificial sun in the sky and calculating the shadowing (Figure 2a). It is intuitive to view and highlights subtle topographic features,

however it reintroduces the above controls on representation, particularly azimuth biasing. It is therefore desirable to utilise methods that are free from azimuth biasing, however these may not improve identification rates of landforms (Smith and Clark, 2005); in addition, if interpreters are not familiar with the output from these techniques, then they may find landform identification more difficult.

Gradient is perhaps the most common method (Figure 2b) and measures the steepness of slope (rate of change of elevation); as landforms often have steep sides they can be readily identified. The rate of change of gradient is *curvature* and is comprised of three components (Schmidt et al., 2003): profile, planform and tangential.

Profile curvature is particularly pertinent as it measures downslope curvature, helping identify breaks-of-slope. Two other common methods include *local contrast stretch (LCS)* and *residual relief separation (RRS)*. LCS (Smith and Clark, 2005) uses the concept of relative elevation and that a landform is spatially distinct from neighbouring features - a standard linear contrast stretch is applied to a region of a specified size, thereby providing

a localised increase in contrast. RRS (Hillier and Smith, 2008) takes a different approach and begins with the conceptual understanding that landscapes are comprised of different elevation elements that are "stacked" on top of one another - these often occur at different width-scales and if the regional-scale relief is removed, then the small-scale "remainder" (or residual) can be extracted. These residuals ideally only contain features of interest, however as this is a scale-based technique it will contain *all features* at that scale and this can include anthropogenic features.

Semi-Automated Mapping

Given the benefits of objective and repeatable methods, automated and semi-automated techniques are highly desirable. They utilise algorithms to quantitatively process input data in order to identify landforms, however the complexity of the process often results in low or moderate levels of accuracy when compared to manual methods - as a result this remains an active topic of research that is evolving rapidly (Seijmonsbergen et al., 2011). The area borrows heavily from the classification of

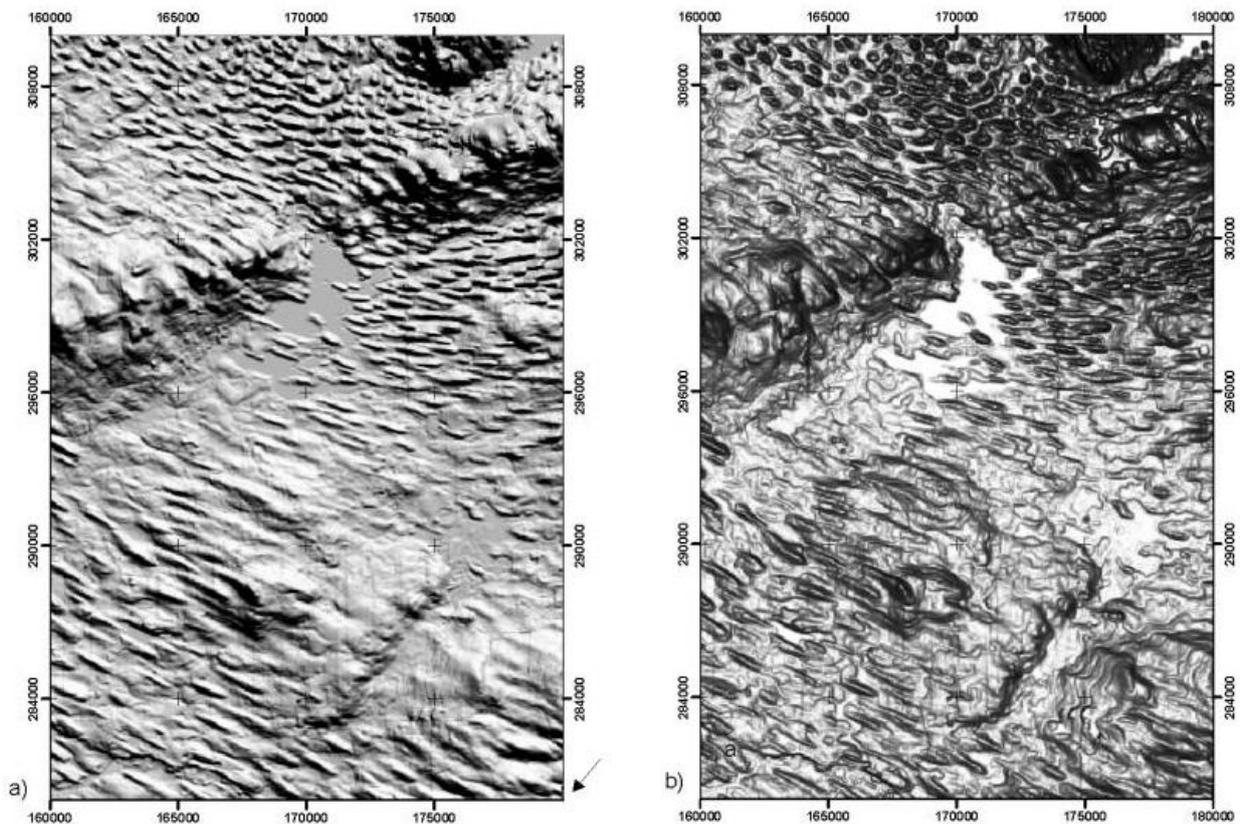


Figure 2 Illustrating the effect of relief shading (a) and gradient (b; darker tones are steep slopes) on the detectability of drumlins from high resolution DEMs of the same area

satellite imagery (Mather, 2004) where feature extraction techniques are used to identify surface features using reflectance at different wavelengths. For mapping, different wavelength inputs are replaced with different LSPs and then processed using either *supervised* or *unsupervised* techniques (Lillesand et al., 2008). Supervised techniques require interaction with a mapper and include methods such as rule-based object classification, cluster analysis (e.g. maximum likelihood) and regression analyses. Unsupervised techniques involve unsupervised object based classification, clustering (e.g. k-means), and machine learning. Perhaps the most promising approach is rule-based object classification (Blaschke, 2010), more commonly termed object-based image analysis (OBIA). Unlike more "traditional" remote sensing based classification techniques that operate on a pixel-by-pixel approach, OBIA operates on the fundamental principle that landforms are conglomerations of pixels. The two-stage process begins by using input LSPs to segment the image into clusters of pixels using a multi-scale algorithm; once complete a set of classification rules is then applied to the objects. The segmentation is controlled by parameters restricting the size and shape of clusters, with the size particularly important for landforms - recent developments have seen automated techniques for optimising the scale parameter for landform extraction (Anders et al., 2011, Dragut et al., 2010).

It is worth noting that testing mapping techniques is a difficult task as it requires a pre-existent knowledge of landforms to determine if an identification of individual features is successful. A higher spatial resolution dataset or manual mapping under "controlled" conditions are often used. Hillier and Smith (2012) take a different approach and place synthetic landforms in a real landscape allowing *a priori* knowledge of the landforms present. This can allow unequivocal testing of the landform mapping method.

Digital Output

Geomorphological maps are a specific kind of thematic map that use complex and illustrative symbolisation. The challenge of mapping is to portray a three-dimensional

landscape with all its elements on a two-dimensional sheet of paper. To deliver this complex information, geomorphological maps make full use of a variety of elements of cartographic design. Different kinds of symbols and colours need to be arranged and composed carefully in order to generate a readable map that clearly expresses the content and message of the project.

Before starting the process of map design it is necessary to review the following questions (Otto et al., 2011):

- What is the purpose, message, and central theme of the map?
- Who has commissioned the map?
- Who will be using the map?
- How will the reader use the map (office, field)?

Applications of geomorphological maps range from simple descriptions of a field site, for example accompanying a journal publication or construction site report, to specialised land system analyses, for example for land management or natural hazard assessment. It is equally important to consider the production process and dissemination of the final product. Is it a paper map? Is the map produced in colour or black and white? Is the map accompanying a journal publication? Will it be published online? These issues strongly influence how to compile and arrange the data, which symbols are used, how the various map items are composed and whether colour can be used or not (Otto, et al., 2011).

Communication with maps differs significantly from other types of human communication. Maps are visual media and evoke visual stimuli that cause different reactions in people than books or conversations. Graphic communication, like maps, delivers all information at once. That means information is not perceived sequentially, but instantaneously with respect to the location and relative position on the map sheet or screen. Thus, the appearance and composition of graphical elements need to be considered thoughtfully. On a map, all information is spatially related and should be considered holistically. The composition of map items determines if and how the reader

understands the message, with perception and understanding occurring subconsciously. To allow map-users to understand the meaning of the map, a visual link to the symbols and their attributes that correspond to the intention of the cartographer needs to be assigned (Robinson et al., 1995).

The basis for most geomorphological maps is a topographic base map presenting contour lines and the general layout of the hydrography and infrastructure. The basic representations of objects on maps are the symbol primitives: point, line, and area (also referred to as marker, line and polygon symbols) (Robinson et al., 1995). Whether a linear feature in nature is represented by a line symbol on the map is primarily a question of scale. For example, a river could be depicted by a blue line. On larger maps (with increasing size of the map items) the river would be depicted using an area symbol of the river shoreline. The map scale also determines if a landform is depicted by a point symbol, or if it is split into its morphological components. Rock glaciers for example could be represented by a single point symbol on small scale maps, or by the assemblage of line and area symbols that differentiate the step height of the rock glacier front, furrows and ridges and the accumulation of boulders and blocks on top of the rock glacier, if the map scale increases (Otto et al., 2011).

Line and pattern symbols (or shading) are commonly used for illustrating gradient and morphology. Some guidelines for the creation of geomorphological maps have been established by the International Geographical Union (IGU) Subcommittee of Geomorphological Survey and Mapping (Gilewska, 1968).

Geomorphological legend systems

A large number of symbols and legends have been developed since the onset of geomorphological mapping in the early 20th century (Passarge, 1914). The impressive diversity of concepts and cartographic conventions created in different scientific communities throughout Europe and the US is related to the terrain configuration of the surveyed region and the scientific focus and aim of the map. The more complex the

terrain, for example mountain areas, the greater the diversity of symbols and colours required (Verstappen, 2011).

In general geomorphological maps and their legend systems can be differentiated into maps showing a single aspect of geomorphology, most commonly morphology, and analytical maps that encompass the full information of a landscape including processes, morphogenesis and even lithology (Knight et al., 2011, Verstappen, 2011). An overview of different legend systems is provided by Otto et al. (2011), in part highlighting how different countries have developed their own systems, either formally or informally, to suit their own cultural needs.

In Britain most geomorphological maps focus on the representation of form producing morphological maps (Cooke and Doornkamp, 1990, Evans, 1990) with significant breaks in slope the most dominant information. The purely morphometric information can be extended by morphogenetic and material information to complete the geomorphological representation of landscapes.

Most maps are generated using GIS or graphic design software, but only a few symbols sets exist that have been developed for digital application. Many legend symbols are so complex that automated drawing by GIS/graphic design software is not possible. Among the few exemptions are the "Mapping Symbols for High Mountains" (Otto, 2008, Otto and Dikau, 2004) and the legend for geomorphological mapping developed by the Institute de Géographie de l'Université de Lausanne (IGUL) (Maillard et al., 2011, Schoeneich et al., 1998) which can be used with Esri ArcGIS. Figure 3 shows a geomorphological map of a high alpine valley (Käferal, Austria) produced in ArcGIS using the symbol set of Otto (2008).

Online maps

Digitally produced geomorphological maps hold the potential for online publication. The possibilities range from simple publication of raster maps or single vector layers as overlays for Google Earth to the generation of a genuine webGIS application containing each single layer of information of the geomorphological map. While the former is

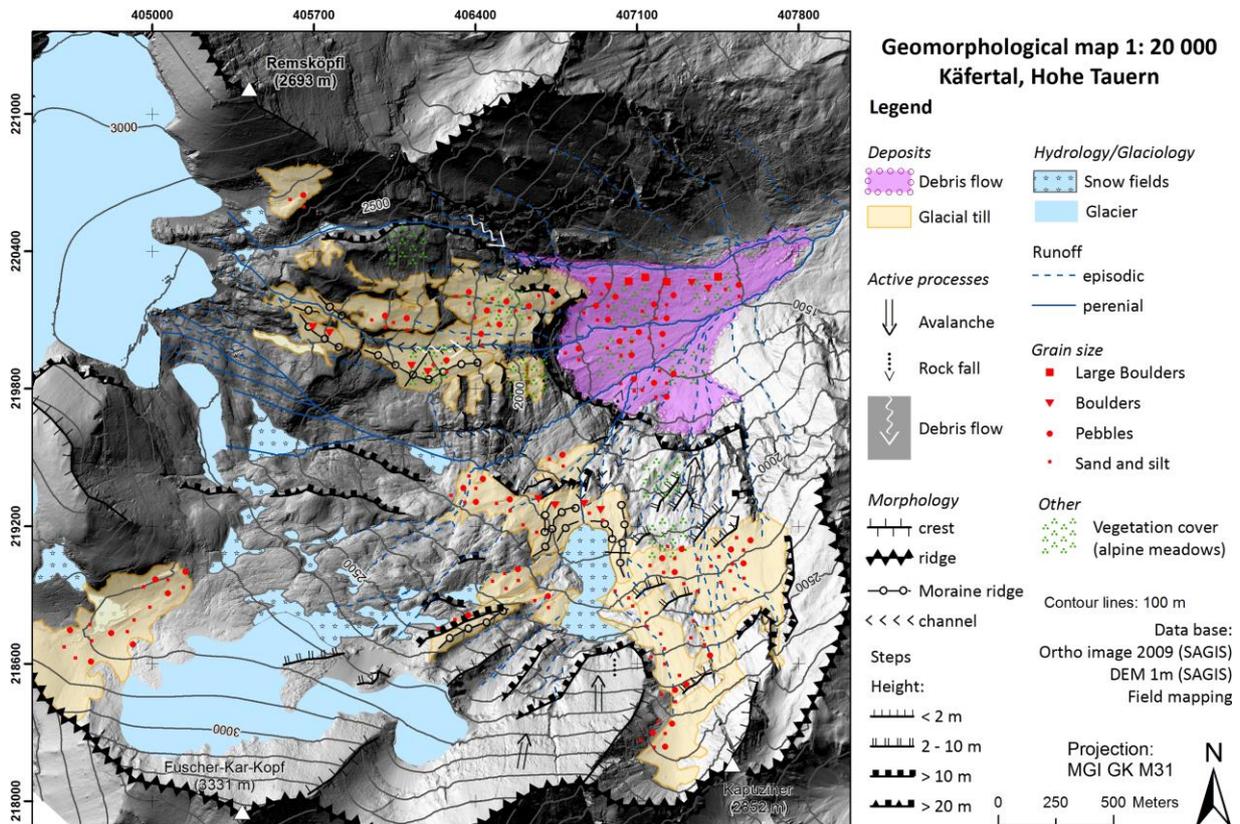


Figure 3: Geomorphological map of the Käfertal, Hohe Tauern, Austrian Alps using the "Mapping Symbols for High Mountains" (Otto, 2008, Otto and Dikau, 2004).

easily performed by export of maps or data layers into the standardised KML/KMZ format, the latter requires a web server application (webGIS software) for the dissemination of the map online. Otto, et al. (2011) and Smith, et al. (2013) provide an introduction to these methods for online publication.

Recently, the standard PDF (Portable Document Format) has been extended into a GeoPDF for display and dissemination of referenced map data. Geospatial functionality of a GeoPDF includes scalable map display, layer visibility control, access to attribute data, coordinate queries, and spatial measurements. Some geospatial data providers such as the United States Geological Survey (USGS; <http://store.usgs.gov>) and the Australian Hydrographic Service (AHS; <http://www.hydro.gov.au/>) currently publish interactive maps using the GeoPDF format.

Conclusions

Geomorphological mapping is undergoing a renaissance - with geomorphology proving to be a cross-cutting discipline, integrating academic and professional applications, the positive impacts upon society are large. Knowing *where* a landform is, *why* it is there, *what* it is made of and *how* it has changed is an incredibly powerful position from which to manage society's interaction with surface (and near-surface) features.

The re-emergence of geomorphological mapping as the pre-eminent paradigm for such studies is a direct result of the proliferation of spatial data (aerial and satellite imagery, DEMs) and IT systems to manage them (GIS). A clear framework for geomorphological spatial data management has been established. Whilst the 1960s and 1970s saw initial academic development, application was limited and as geomorphology as a discipline shifted to field-based studies interest declined. Indeed the mantle of geomorphological mapping shifted to engineering geology during the 1980s and

1990s as the benefits for direct application were recognised. The forefront of geomorphology has seen the integration of process-level understanding with regional scale modelling. GIS provides the underlying framework; maps and mapping are an integral part of this process and the skills, knowledge and understanding of the methods for data collection, processing and output are a key element. Geomorphological mapping is the underpinning domain as geomorphology moves centre-stage.

References

- Anders NS, Seijmonsbergen AC, Bouten W. 2011. Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. *Remote Sensing of Environment* **115**: 2976-2985
- Barsch D, Liedtke H. 1980. Principles, scientific value and practical applicability of the geomorphological map of the Federal Republic of Germany at the scale of 1 : 25,000 (GMK 25) and 1 : 100,000 (GMK 100). *Zeitschrift für Geomorphologie Suppl.* **36**: 296-313
- Blaschke T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* **65** (1): 2-16
- Brunsdon D, Doornkamp JC, Fookes PG, Jones DKC, Kelly JMH. 1975. Large scale geomorphological mapping and highway engineering design. *Quarterly Journal of Engineering Geology* **8**: 227-253
- Colwell RN. 1983. *Manual of Remote Sensing*. American Society of Photogrammetry: Washington
- Cooke RU, Doornkamp JC. 1990. *Geomorphology in environmental management, A New Introduction*. Clarendon Press: Oxford
- Dackombe RV, Gardiner V. 1983. *Geomorphological field manual*. George Allen & Unwin: London
- Dragut L, Tiede D, Levick S. 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* **24** (4): 859-871
- Dramis F, Guida D, Cestari A. 2011. Nature and Aims of Geomorphological Mapping. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 39-74
- Dykes AP. 2008. Geomorphological maps of Irish peat landslides created using hand-held GPS. *Journal of Maps* **2008**: 258-276
- Embleton C, Verstappen HT. 1988. The nature and objectives of applied geomorphological mapping. *Zeitschrift für Geomorphologie, Supplementband* **68**: 1-8
- Evans IS. 1990. Cartographic techniques in geomorphology. In *Geomorphological techniques*. Goudie A (eds.). Unwin Hyman: London, 97-108
- Gilewska SK, M. 1968. Project of the Unified key to the geomorphological map of the World. *Folia Geographica, Ser. Geographica-Physica* II. Polska Akademia Nauk, Kraków:
- Gustavsson M, Seijmonsbergen AC, Kolstrup E. 2008. Structure and contents of a new geomorphological GIS database linked to a geomorphological map - With an example from Liden, central Sweden. *Geomorphology* **95** (3-4): 335-349.
- Hengl T, Reuter HI. 2008. *Geomorphometry: Concepts, Software, Applications*. Elsevier:
- Hillier JK, Smith M. 2008. Residual relief separation: digital elevation model enhancement for geomorphological mapping. *Earth Surface Processes and Landforms* **33** (14): 2266-2276.
- Hillier JK, Smith MJ. 2012. Testing 3D landform quantification methods with synthetic drumlins in a real digital elevation model. *Geomorphology* **153-154**: 61-73
- Klimaszewski M. 1982. Detailed geomorphological maps. *ITC Journal* **3**: 265-271
- Knight J, Mitchell W, Rose J. 2011. Geomorphological Field Mapping. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 151-188
- Lee EM. 2001. Geomorphological mapping. *Geological Society Special Publication* (18): 53-56

- Lillesand T, Kiefer RW, Chipman J. 2008. Remote Sensing and Image Interpretation. John Wiley & Sons: New York
- Maillard B, Lambiel C, Martin S, Pellitero Ondicol R, Reynard E, Schoeneich P. 2011. The ArcGIS version of the geomorphological mapping legend of the University of Lausanne. Department of Geography, University of Lausanne: Lausanne
- Mather PM. 2004. Computer Processing of Remotely-Sensed Images. Wiley Blackwell: London
- Oguchi T, Hayakawa Y, Wasklewicz T. 2011. Data Sources. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 189-224
- Otto JC. 2008. Symbols for geomorphologic mapping in high mountains for ArcGIS. Available at <www.geomorphology.at>
- Otto JC, Dikau R. 2004. Geomorphic system analysis of a high mountain valley in the Swiss Alps. *Zeitschrift für Geomorphologie N.F.* **48** (3): 323-341
- Otto JC, Gustavsson M, Geilhausen M. 2011. Cartography: design, symbolisation and visualisation of geomorphological maps. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 253-296
- Paron P, Claessens L. 2011. Makers and users of geomorphological maps. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 75-106
- Passarge S. 1914. Morphologischer Atlas. Erläuterungen zu Lief. 1, Morphologie des Messtischblattes Stadtremba (1:25.000). Mitt. Geogr. Gesell. Hamburg
- Pellegrini GB. 1993. A proposal of a legend for applied geomorphology. *Proposta di legenda geomorfologica ad indirizzo applicativo* **16** (2): 129-152
- Robinson AH, Morrison JL, Muehrcke PC, Kimerling AJ, Guptill SC. 1995. Elements of Cartography. Wiley: Chichester
- Schmidt J, Evans IS, Brinkmann J. 2003. Comparison of polynomial models for land surface curvature calculation. *International Journal of Geographic Information Science* **17** (8): 797-814
- Schoeneich P, Reynard E, Pierrehumbert G. 1998. Geomorphological mapping in the Swiss Alps and Prealps. In *Hochgebirgskartographie Silvretta '98*. Kriz K (eds.). Geographisches Institut: Wien, 145-153
- Seijmonsbergen AC, Hengl T, Anders NS. 2011. Semi-automated identification and extraction of geomorphological features using digital elevation data. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 297-336
- Smith MJ. 2011. Digital Mapping: visualisation, interpretation and quantification of landforms. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 225-251
- Smith MJ, Clark CD. 2005. Methods for the visualisation of digital elevation models for landform mapping. *Earth Surface Processes and Landforms* **30** (7): 885-900
- Smith MJ, Wise SM. 2007. Problems of bias in mapping linear landforms from satellite imagery. *International Journal of Applied Earth Observation and Geoinformation* **9**: 65-78
- Smith MJ, Pain C. 2009. Applications of Remote Sensing in Geomorphology. *Progress in Physical Geography* **33** (4): 568-582
- Smith MJ, Griffiths J, Paron P (eds.). 2011. Geomorphological Mapping: methods and applications. Elsevier: London
- Smith MJ, Hilier J, Otto JC, Geilhausen M. 2013. Geovisualisation In *Treatise on Geomorphology*. Shroder JF (eds.). Academic Press Elsevier: San Diego, 299-325
- Tricart J. 1965. Principes et methodes de la geomorphologie. Masson: Paris
- Verstappen HT. 2011. Old and New Trends in Geomorphological and Landform Mapping. In *Geomorphological Mapping: methods and applications*. Smith MJ, Paron P, Griffiths J (eds.). Elsevier: London, 13-38

Shoreline Geometry: DSAS as a Tool for Historical Trend Analysis

Temitope D. T. Oyedotun^{1,2}

¹ Coastal and Estuarine Research Unit, Department of Geography, University College London (temitope.oyedotun.10@ucl.ac.uk)

² Department of Geography and Planning Sciences, Adekunle Ajasin University, Akungba-Akoko, Ondo State of Nigeria (oyedotuntim@yahoo.com)



ABSTRACT: Shoreline geometry remains one of the key parameters in the detection of coastal erosion and deposition and the study of coastal morphodynamics. The Digital Shoreline Analysis System (DSAS) as a software extension within the Environmental System Research Institute (ESRI) ArcGIS© has been used by many researchers in measuring, quantifying, calculating and monitoring shoreline rate-of-change statistics from multiple historic shoreline positions and sources. The main application of DSAS is in utilisation of polyline layers as representation of a specific shoreline feature (e.g. mean high water mark, cliff top) at a particular point in time. A range of statistical change measures are derived within DSAS, based on the comparison of shoreline positions through time. These include Net Shoreline Movement (NSM), Shoreline Change Envelope (SCE), End Point Rate (EPR), Linear Regression Rate (LRR) and Weighted Linear Regression Rate (WLR). Despite the inability of this tool to determine the forcing of morphodynamics, it has been shown to be effective in facilitating an in-depth analysis of temporal and historical movement of shoreline positions and cliff geometry.

KEYWORDS: GIS, Historical Trend Analysis, DSAS, Shoreline Changes, transects.

Introduction

Coastal shorelines, the interface between land and sea, change variably in response to one or more factors, which may be morphological, climatological or geological in nature. Shoreline geometry depends on the interactions between and among waves, tides, rivers, storms, tectonic and physical processes. Erosion (landward retreat) and deposition (advance and growth through accretion) can both present challenges to coastal communities, coastal infrastructures and the adjacent estuarine systems (e.g. Benumof *et al.*, 2000; Moore and Griggs, 2002; Collins and Sitar, 2008; Katz and Mushkin, 2013). Knowledge and assessment of the changes in shoreline position have proved crucial in the overall understanding of dynamics in coastal areas and the morphodynamic processes driving the change. Shorelines are vulnerable to change driven by sea-level rise, varying coastal

climates including marine, astronomical or other meteorological factors (Lisitzin, 1974; Cowell and Thom, 1994; Pugh, 1996, 2004; Paskoff and Clus-Auby, 2007; Pardo-Pascual *et al.*, 2012; Thébaudeau *et al.*, 2013). Changes (whether short-term or long-term) in the position and geometry of shorelines are very important in the understanding of coastal dynamism and the management of coastal areas (Esteves *et al.*, 2009; Rio *et al.*, 2013). Prompt and intensive short-term field research is mostly used to evaluate shoreline responses to the coastal forcings (Collins and Sitar, 2008; Young *et al.*, 2009; Quinn *et al.*, 2010; Brooks *et al.*, 2012). However, the wider range of longer time marine or terrestrial events which drive shoreline responses may not be appropriately linked or accounted for. Therefore, quantitative analysis of shoreline changes at different timescales is very important in understanding and establishing the processes driving

erosion and accretion (Sherman and Bauer, 1993; Esteves *et al.*, 2011; Katz and Mushkin, 2013), computing sediment budgets (Zuzek *et al.*, 2003), identification of hazard zones (Lawrence, 1994; Al Bakri, 1996), as a basis for modelling of morphodynamics (Maiti and Bhattacharya, 2009), and for coastal management and interventions (Esteves *et al.*, 2009).

Shoreline geometry and position are perhaps the most basic indicators with which to evaluate changes in coastal regions. Digital Shoreline Analysis System (DSAS) has therefore been used in investigating the dynamics of shoreline movements and changes at both shorter (e.g. Brooks and Spencer, 2010) and longer / historical (e.g. González-Villanueva *et al.*, 2013) time scales. However, it is only by including longer time periods (decades to centuries) of investigation that a wider range of past coastal events, magnitudes and frequencies can be linked with the shoreline morphodynamics (Wolman and Miller, 1960; Brooks *et al.*, 2012). Historical Trend Analysis (HTA henceforth) is, therefore, one of the key approaches used in the analysis of change over historical (decade to century) timescales (Blott *et al.*, 2006; HR Wallingford *et al.*, 2006). This 'bottom-up' approach is used to identify past behaviour in order to predict future trends. The focus of this article is the application of HTA to analyse changes in shoreline positions, channel morphology, identification of areas of 'cut' and 'fill', or erosion and deposition over historical time. Specifically the application of the GIS-based Digital Shoreline Analysis System to HTA is explored.

DSAS

The Digital Shoreline Analysis System (DSAS henceforth) is a GIS tool that can be used in HTA to examine past or present shoreline positions or geometry. One of the main benefits of using DSAS in coastal change analysis is its ability to compute the rate-of-change statistics for a time series of shoreline positions. The statistics allow the nature of shoreline dynamics and trends in change to be evaluated and addressed. DSAS has

been developed as a freely available extension to Environmental System Research Institute (ESRI)'s ArcGIS (Thieler *et al.*, 2009). It has been updated and upgraded over time, so multiple versions exist allowing its use with ArcView 3.2 through to ArcGIS v10. In 2013, a web-based version (DSASweb) was released (USGS, 2013). Download and further information on the installation and use of DSAS can accessed at <http://woodshole.er.usgs.gov/project-pages/dsas/>. Instructions, usage of the software and the configuration of input and output parameters are well documented in Thieler and Danforth (1994a, 1994b) and Thieler *et al.* (2009).

There are numerous examples of the use of DSAS in the study of coastal behaviour and shoreline dynamics. Table 1 reviews examples of recent studies that have utilised DSAS in Historical Trend Analysis, coastal system dynamics, shorter time shoreline changes, cliff geometry modelling and estimations. More broadly, DSAS in HTA can be used to undertake:

- i. The mapping of historic configurations of shoreline position over the period covered by available spatial data (e.g. maps, aerial photographs);
- ii. The evaluation of historic changes and trends of individual or selected transects (discrete alongshore positions). Within DSAS, shoreline change is calculated at specific transects, and the time-series of change at specific locations can be evaluated using the DSAS output;
- iii. The analysis of shoreline geometry, including foreshore steepening (using the distance between mean high and low water marks (after Taylor *et al.*, 2004) and orientation (for example, to examine rotational tendencies (e.g. Nebel *et al.*, 2012);
- iv. To predict patterns of shoreline behaviour using the derivation of historical rate of change trends as an indicator of future trends assuming continuity in the physical, natural or anthropogenic forcing which have forced the historical change observed at the site.

Table 1. Recent studies on shoreline geometry that made use of DSAS

Coastline feature studied	Articles
Historical record of coastline dynamics	Carrasco <i>et al.</i> , 2012; Montreuil and Bullard, 2012; González-Villanueva <i>et al.</i> , 2013; Jabaloy-Sánchez <i>et al.</i> , 2014
Shoreline variation, shoreline erosion and short-time coastal changes	Houser <i>et al.</i> , 2008; Brooks and Spencer, 2010; Houser and Mathew, 2011; Restrepo, 2012; Beetham and Kench, 2014; Hapke <i>et al.</i> , 2013; Rio <i>et al.</i> , 2013
Gully development and evolution	Leyland and Darby, 2008; Draut <i>et al.</i> , 2011
Cliff retreat and erosion	Rio and Gracia, 2009; Brooks <i>et al.</i> , 2012; Katz and Mushkin, 2013; Young <i>et al.</i> , 2014
Shoreline / cliff measurement and modelling	Hackney <i>et al.</i> , 2013; Thébaudeau <i>et al.</i> , 2013

Shoreline digitisation and data quality consideration

The main stages in the shoreline analysis workflow, as undertaken using DSAS within ArcGIS, are detailed in Thieler *et al.*, 2009. Shoreline positions are important features defined in DSAS analysis. Specific features of interest should be extracted through digitisation, and these may include the Mean Low Water (MLW) and Mean High Water (MHW) marks. These shoreline positions are explicitly indicated on Ordnance Survey (OS) and other national mapping agency publications which make for simple digitisation and analysis within a GIS, thereby reducing some complications associated with automatic shoreline detections in other sources (Ryu *et al.*, 2002; Loos and Niemann, 2002; Maiti and Bhattacharya, 2009). Other sources from which shoreline positions can be derived include satellite imageries, digital orthophotos, aerial photographs, global positioning-system field surveys (GPS/ and dGPS), historical coastal-survey maps, or extracted from LiDAR surveys (e.g. Stockdon *et al.*, 2002; Weber *et al.*, 2005; Himmelstoss *et al.*, 2011; Hapke *et al.*, 2011; Hapke *et al.*, 2013). Accurate and careful digitisation of shoreline position, possibly with constant reference to the same feature, is however, recommended when digitising from these other sources. DSAS form of analysis is not immune to the usual limitations associated with digitisation and synthesis of variable quality and resolution data derived from various sources as a result of irregular time sampling interval.

For example, reliance on Ordnance Survey (OS) mapping relies on the accurate and consistent interpretation of surveyors and cartographers over decades and centuries (Fenster *et al.*, 1993; Burningham and French, 2006). Older surveys were usually land-based whilst later ones are often derived from aerial photography (Fenster *et al.*, 1993). Care must therefore be undertaken to ensure that accurate digitisation and critical review of features are considered in these source materials. The calculated measures of change provided by DSAS are only as reliable as the sampling and measurement accuracy associated with the source materials. For example, mapping errors can be estimated as +/- 10m for the pre-2000 data and +/-5m for post-2000 data (Anders and Byrnes, 1991; Crowel *et al.*, 1991; Thieler and Danforth, 1994; and, Moore, 2000).

Any form of spatial or laboratory analysis is always aimed at finding solutions to certain spatial problems or to understand certain processes (Uluocha, 2007). In order to achieve this focus, the significance of data quality cannot be over-stressed. The issue of data quality must therefore be given prominence. Care should be taken to determine the integrity, quality and relevance of any data to be used in any analyses (Uluocha, 2007). Data quality must not be compromised in shoreline analysis using DSAS. The indices which can be used in data quality check include logical consistency, completeness, positional accuracy and precision, scale, spatial resolution and

currency (temporal accuracy and precision), (Faiz and Boursier, 1996; Jones, 1997; Dobson, 1992; Uluocha, 2007). It is therefore important that mapping uncertainties introduced and associated with shoreline mapping procedures be identified and included in shoreline change analysis so as to reflect the long-term trend which are not based on short-term variability (Romine *et al.*, 2009).

Shoreline analysis and interpretation

The DSAS approach calculates shoreline rates of change based on the measured differences between the shoreline positions associated with specific time periods. The following statistical measures (from Thieler *et al.*, 2009) are possible in DSAS:

- (i) Shoreline Change Envelope (SCE): a measure of the total change in shoreline movement considering all available shoreline positions and reporting their distances, without reference to their specific dates.
- (ii) Net Shoreline Movement (NSM): reports the distance between the oldest and the youngest shorelines.
- (iii) End Point Rate (EPR): derived by dividing the distance of shoreline movement by the time elapsed between the oldest and the youngest shoreline positions.
- (iv) Linear Regression Rate (LRR): determines a rate-of-change statistic by fitting a least square regression to all shorelines at a specific transects. Further statistics associated with LRR include Standard Error of Linear Regression (LSE), Confidence Interval of Linear Regression (LCI) and R-Squared of Linear Regression).

Other standard DSAS statistical parameters are Weighted Linear Regression Rate (WLR) and associated measures (Standard Error of Weighted Linear Regression (WSE), Confidence Interval of Weighted Linear Regression (WCI), R-squared of Linear Regression (WR2)) and Least Median of Squares (LMS). All parameters can be illustrated to show the spatial patterns of change in shoreline change statistics. The objectives of the analysis to be investigated and the characteristics of the datasets are major determinants of the choice of statistical method.

Case Study: Shoreline morphodynamics at Crantock Beach, southwest England

Crantock Beach lies between Pentire Point East and Pentire Point West cliffs, on the north coast of Cornwall, southwest England (Figure 1). The coastline is macrotidal (mean spring tide range 6.4 m), and is exposed to a predominantly westerly wave climate with a 10% annual exceedance wave height of 2.5-3m, and a 1 in 50 year extreme offshore wave height of 20m. The beach lies at the seaward extent of the Gannel estuary, a ria estuarine system comprising sandy intertidal flats (c.70% of the valley is intertidal Davidson *et al.*, 1991). The estuarine valley is widest at the mouth where the large sandy beach-dune system of Crantock Beach lies. Saltmarshes have infilled the estuary at the landward extent (Oyedotun *et al.*, 2012). Previous research on the system has focused mainly on the Gannel estuary, on the impacts of mining on sediments / sedimentation (Reid and Scrivenor, 1906; Bryan *et al.*, 1980; Pirrie *et al.*, 2000; Pirrie *et al.*, 2002) and sedimentary environments and sedimentology (Oyedotun *et al.*, 2012). The case study presented here describes the changes in shorelines at Crantock Beach, and low tide channel position within the inlet of the Gannel estuary through the analysis of historical maps using DSAS in ArcGIS. Temporal and spatial variability in coastal change, the geomorphic sensitivity and likely processes forcing the morphological behaviour are explored.

DSAS Method

The investigation of changes in shoreline positions in the vicinity of the sandy, macrotidal Crantock Beach and Gannel estuary is carried out using the Ordnance Survey historical mapping archive, available for the period between 1888 and 2012 (Table 2). Movements of both the Mean Low Water (MLW) and Mean High Water (MHW) here are investigated in GIS using the DSAS extension developed by the USGS (Thieler *et al.*, 2009). The historical mapping is available as georeferenced GeoTiffs from Digimap. Shorelines were digitised from each map, and the standard DSAS shoreline change measures - Net Shoreline Movement (NSM) and End Point Rate (EPR) - were calculated. Net Shoreline Movement (NSM): reports the

distance between the oldest (1888) and the youngest (2012) shorelines, which presents the overall change in shoreline position for the 124 year period. The End Point Rate (EPR) converts this net shoreline movement

into an annual rate of shoreline change, i.e. dividing the distance of shoreline movement from the earliest to most recent shorelines by the time period passed.

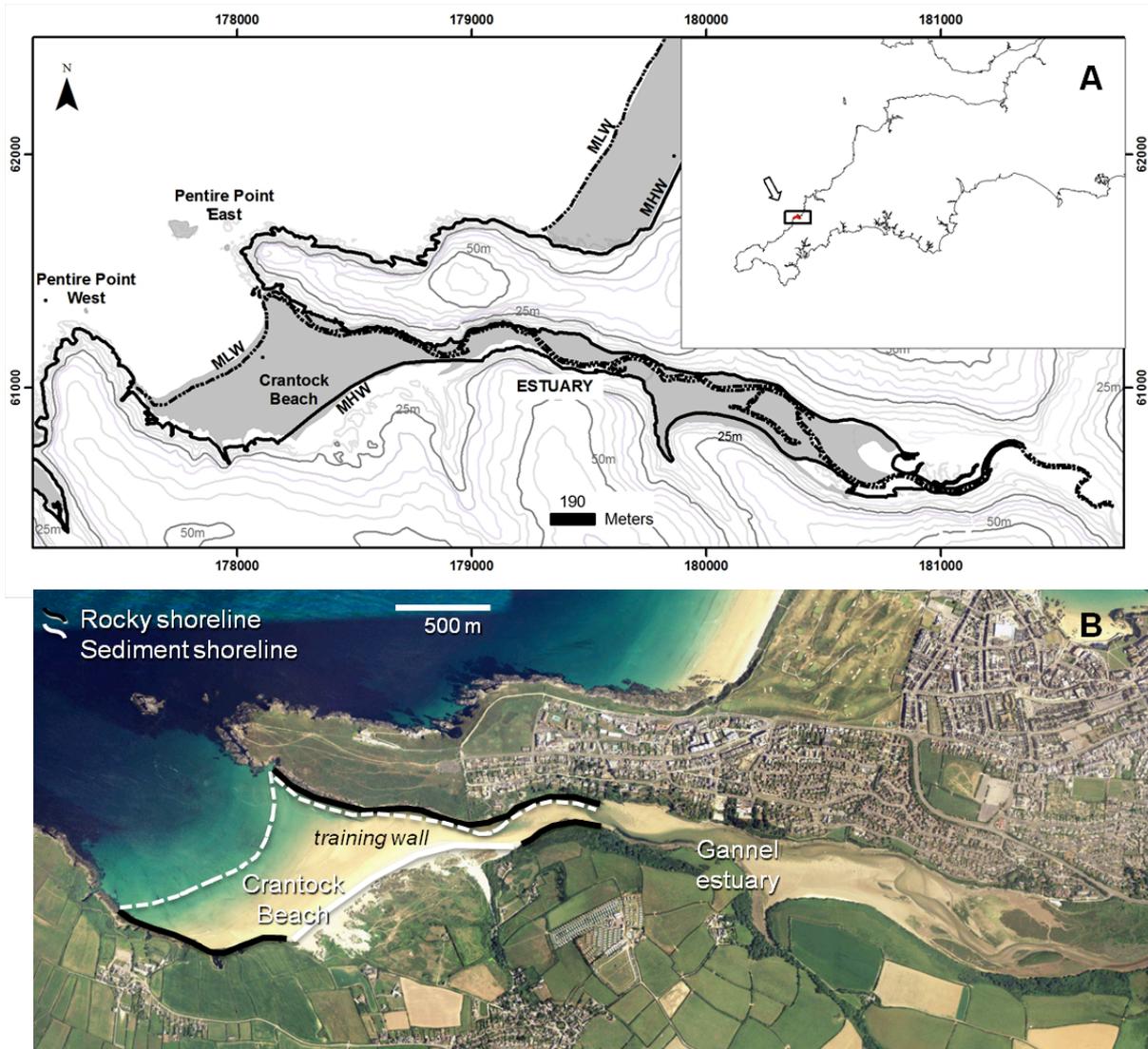


Figure 1: (A) Crantock Beach and the Gannel estuary, Cornwall, located in southwest England and (B) the area photograph showing the rocky and sediment shoreline.

Table 2. Ordnance Survey Data for this study.

Map date	Data	Scale	Resolution/accuracy
1888	County Series	1:10,560	+/-10m
1901	County Series	1:10,560	+/-10m
1977	National Survey	1:2,500	+/-10m
1996	1 st Metric Edition	1:10,000	+/-10m
2012	MasterMap	1:2,000	+/-5m

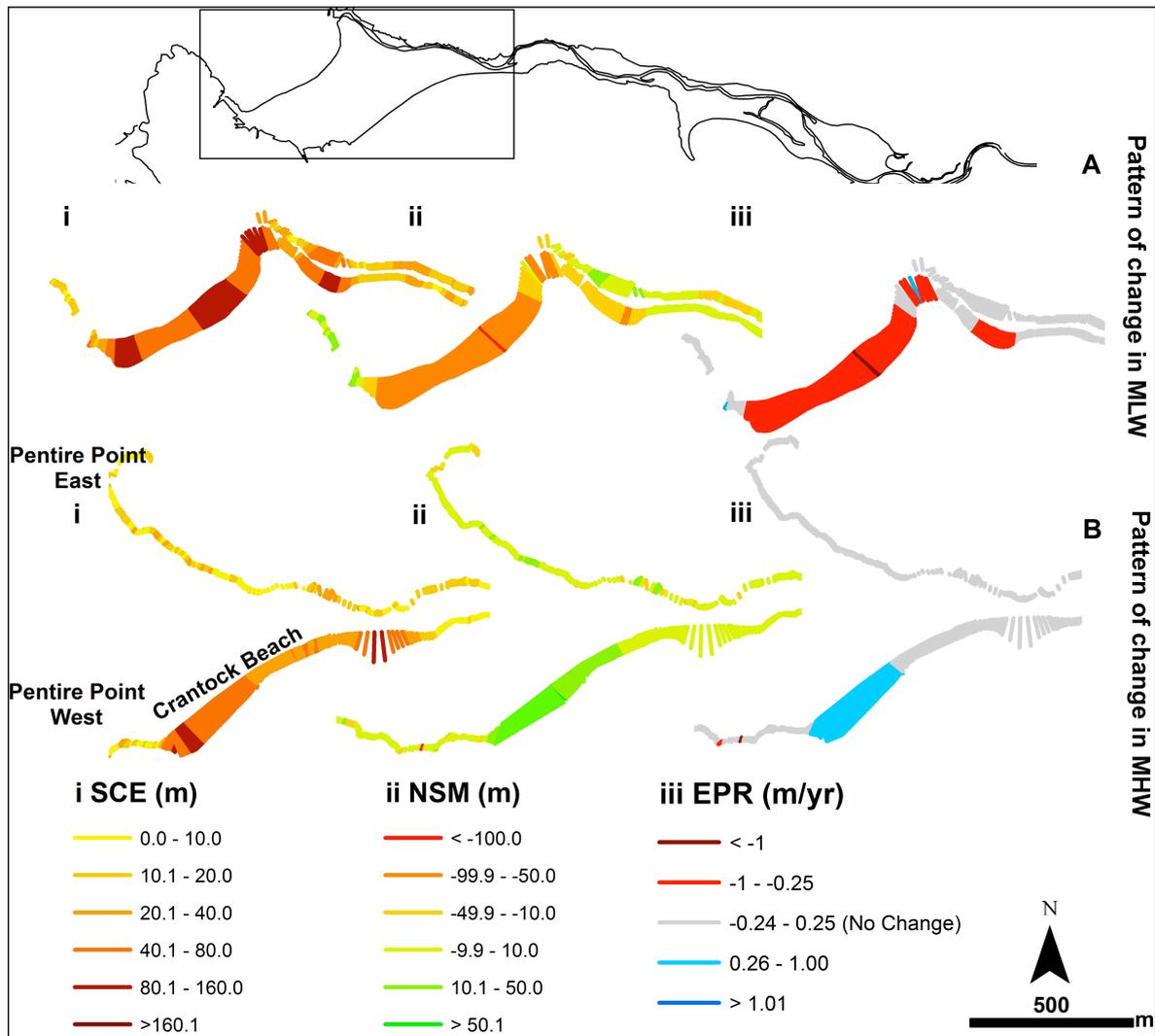


Figure 2. Shoreline Change Envelope (SCE), annual rate of change (EPR) and Net Shoreline Movement (NSM) for Mean Low Water (MLW) and Mean High Water (MHW) at Crantock Beach

Historical Shoreline Changes

Figure 2 summarises the scales and rates of change in shoreline position at Crantock Beach. Scales of change in the position of MLW are maximum in the centre of the bay and minimum in the inlet (Figure 2Ai). These changes are almost entirely the consequence of recession (landward movement) of the low water shoreline at a rate of $0.1\text{--}0.8\text{ myr}^{-1}$ (Figure 2Aii). The only place in the lower foreshore where significant erosion is not taking place is along the ebb channel margins within the inlet. Comparison of the SCE with NSM shows that the majority of change in the position of MLW change exhibited here is equivalent to the difference between the earliest (1888) and most recent (2012) shorelines.

The scales of change in the position of high water are somewhat reduced by comparison. First, it is clear that the rock-dominated shorelines along Pentire Points West and East have changed very little ($<20\text{ m}$) over this 124 year period (Figure 2B). Second, the high water shoreline of Crantock Beach has shifted in position by up to around 140 m , but is largely characterised by change of the order of $30\text{--}60\text{ m}$. Interestingly, shifts in MHW are the product of shoreline advance (deposition), and there is evidence that gross change (SCE) is greater than net change (NSM). In the absence of significant change in tidal regime, the product of a retreating low water and advancing high water is steepening of the intertidal profile.

Sites of specific change can be explored in detail through the examination of discrete

transect data (Figure 3). Comparison of the time series of changing shoreline position shows how important it is to consider both the envelope of variability (SCE) and the net change (NSM and EPR). Transects at Pentire Points West (Figure 3A) and East (Figure 3D) show consistency in shoreline positions, principally because they represent the rocky shorelines (indicated in aerial photograph of Figure 1). Changes along the Pentire Points are small, but the inlet shoreline shows the minimum change across the system, and these changes are likely well within the accuracy margin (Figure 3C). At Crantock Beach (Figure 3B) however, the high water shoreline has advanced substantially (c. 25 m), but during the entire 124 year period, also experienced an episode of erosion during the late 19th century. All the plots show that the most recent shoreline is a significant departure from those from previous years. In most cases, the change between the late 1990s and 2012 is greater than change at any other time.

Annual rates of change in MHW and MLW at each transect are summarised in Table 3. The analyses show that 42% of MHW transects and 71% of MLW transects are erosion-dominated (i.e. these have experienced a net landward retreat in position). Furthermore, 10% of MHW transects remained unchanged (with no appreciable change in MHW position). Rates of change in MLW are significantly greater than rates of change in MHW. Fewer transects show evidence of deposition, with only 27% experiencing advance in the MLW and 48% showing seaward shifts in the MHW.

The scenario observed here can be compared to what Taylor *et al.* (2004) referred to as “lateral landward retreat through non equilibrium profile” (Page 181). The patterns of change shown indicate an overall dominance of erosion, and rates of retreat and advancement are occurring at unequal levels between MHW and MLW. This leads to a change in foreshore geometry. Here, the landward shift in MLW, and seaward advance in MHW has produced a narrower and steeper intertidal zone. This has occurred at some locations in the tidal inlet, but mainly along Crantock Beach. Where MLW and MHW both retreat or

advance at similar rates, the foreshore is maintaining a consistent intertidal profile. The shifts in shoreline positions vary throughout the bay, and regions of stability can be found in close proximity to areas of significant change.

Shoreline change statistics (SCE, NSM and EPR) presented for the case study has been able to show the large-scale patterns of retreat and growth of the case study's shorelines. Other statistical methods are also effective in indicating the pattern of spatial and temporal movements. However, the choice of SCE is because of its capacity is providing the envelope of variability and it is the only measure which take into account all shorelines. The NSM and EPR only reflect change between the first and the most recent surveys. The choice of DSAS statistical parameters in the case study has been able to explore the temporal and spatial dynamics of the coastal change and the geomorphic variability along the beach because of their ability in making use of all shoreline positions (SCE), the cumulative shoreline movement (NSM) and the time variations (EPR) which encapsulate the rate-range of the historical dataset.

The consideration of utilising five historical maps in the case study underlies the importance of the availability and accessibility of archival datasets in HTA. The maps utilised here are the five complete historical dataset accessible at the Ordnance Survey. Full and detailed opportunities for decadal and annual scale analyses are mostly infrequent in many analyses because the complete and up-to-date archival records have to exist for such analyses in being able to: better understand process drivers, determine barriers to recession and conditions for erosion / deposition etc. (Bray and Hooke, 1997; Brooks *et al.*, 2012). More archival data can, therefore, be sourced from other sources like area photographs and other imagery, to add further understanding of the contemporary morphodynamics of the case study (which is not the focus of this current article). However, the sampled analysis executed using only the Ordnance Survey data has shown that DSAS can yield valuable information on shoreline dynamics at both historical and non-historical timescale.

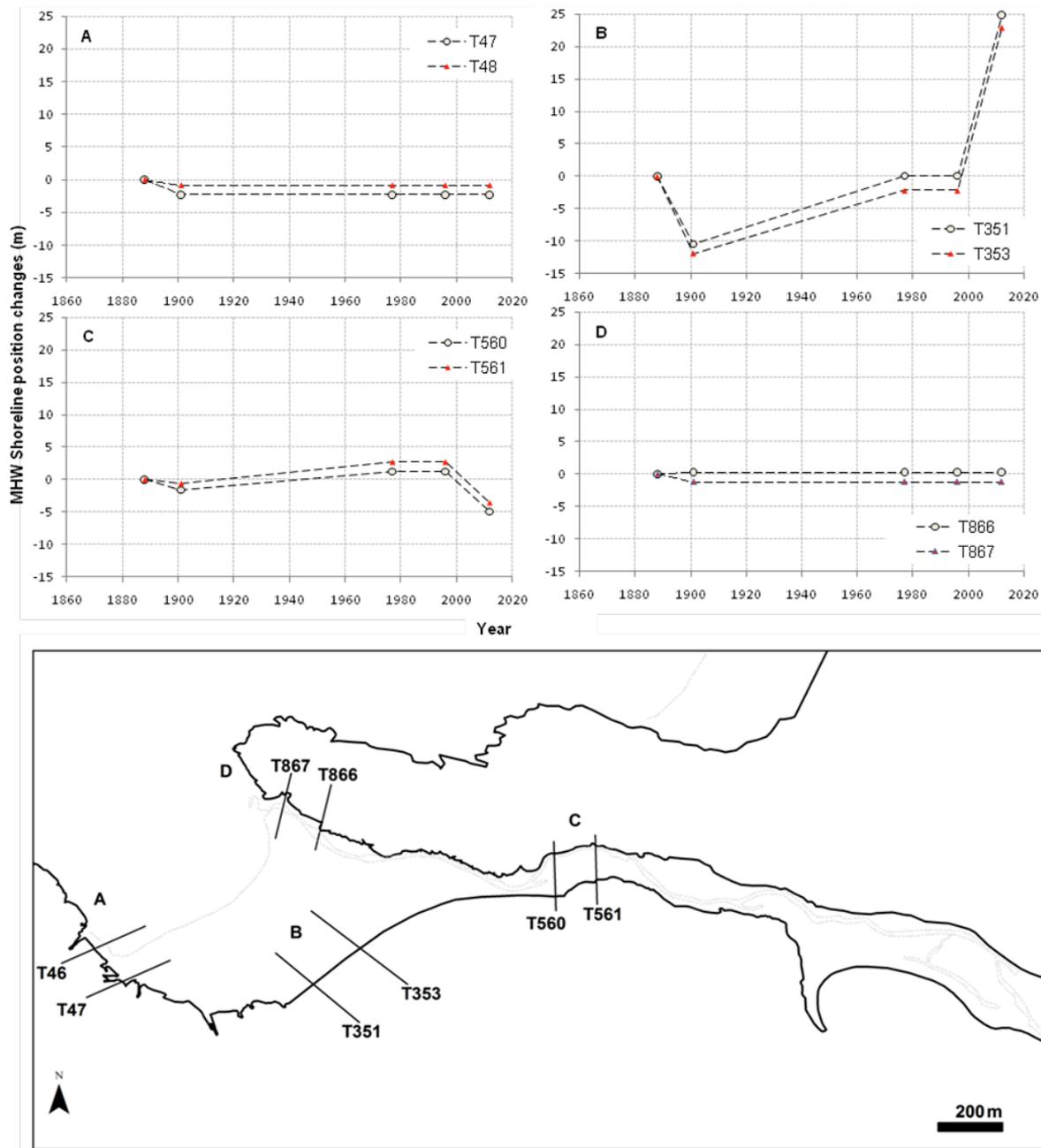


Figure 3. Cumulative change in shoreline position at example transects (A) Pentire Point West, (B) Crantock Beach, (C) Inlet and (D) Pentire Point East.

Table 3 Summary of MHW and MLW movements and trends in the beach

Shoreline	No of Transects	% of Transects	Change rate (m yr ⁻¹)	Total Erosion Change (m) Minimum to Maximum
MLW Retreat	1141	71	-0.01 to -5.96	-0.17 to -738.67
MLW Advance	157	27	+0.01 to +6.51	+0.1 to +707.7
No MLW Movement	8	1.4	-	-
MHW Retreat	382	42	-0.01 to -0.75	-0.01 to -93.64
MHW Advance	437	48	+0.01 to +1.99	+0.01 to +248.26
No MHW Movement	87	10	-	-

Conclusion

The historical evolution and temporal morphodynamics of shoreline position and geometry are of significant importance in evaluating the spatial dynamics of the coastal system behaviour. The ability of DSAS within ArcGIS not only enhances the functionality of the software but also enables the calculation of scales and rates of change statistics from multiple historic shoreline positions and sources. The potential and application of the DSAS extension has been explored in this article. DSAS, however, is only useful to address features that can be represented as lines at a particular point in time, with the accuracy of the results being dependent on the accuracy of the input data. Despite the inability of the DSAS extension to indicate the forcing driving the observed dynamics and changes in coastal environment, it is effective in (i) historic measurement of movement of shoreline geometry, and, (ii) the mapping and identification of coastline erosion and accretion on the coastal environment. Changes at the annual, decadal and historical time scale on the coastal environment are easily and simply measurable in GIS through the DSAS extension within the ArcGIS environment.

The case study presented here has shown that DSAS can yield valuable information on the morphodynamic behaviour of shorelines in terms of shifting shoreline position and changes in foreshore geometry, in the identification of areas of erosion and deposition, and in the variation of planimetric properties of the coastal environment.

References

- Al Bakri D. 1996. A geomorphological approach to sustainable planning and management of the coastal zone of Kuwait. *Geomorphology*, 17, 323-337. DOI: 10.1016/0169-555X(96)00009-8.
- Anders FJ, Byrnes MR. 1991. Accuracy of shoreline change rates as determined from maps and aerial photographs. *Shore and Beach*, 59, 17-26.
- Beetham EP, Kench PS. 2014. Wave energy gradients and shoreline change on Vabbinfaru platform, Maldives. *Geomorphology*, 209, 98 – 110. DOI: 10.1016/j.geomorph.2013.11.029.
- Benumof BT, Storlazzi CD, Seymour RJ, Griggs GB. 2000. The relationship between incident wave energy and sea cliff erosion rates: San Diego County, California. *Journal of Coastal Research*, 17, 1162-1178.
- Blott SJ, Pye K, van der Wal D, Neal A. 2006. Long-term morphological change and its causes in the Mersey Estuary, NW England. *Geomorphology*, 81, 185-206. DOI: 10.1016/j.geomorph.2006.04.008.
- Brooks SM, Spencer T. 2010. Temporal and spatial variation in recession rates and sediment release from soft rock cliffs, Suffolk coast, UK. *Geomorphology*, 124, 26-41. DOI: 10.1016/j.geomorph.2010.08.005.
- Brooks SM, Spencer T, Boreham S. 2012. Deriving mechanisms and thresholds for cliff retreat in soft-rock cliffs under changing climates: Rapidly retreating cliffs of the Suffolk coast, UK. *Geomorphology*, 153/154, 48-60. DOI: 10.1016/j.geomorph.2012.02.007.
- Bray MJ, Hooke JM. 1997. Prediction of soft-cliff retreat with accelerating sea level rise. *Journal of Coastal Research*, 13, 453-467.
- Bryan GW, Langston WJ, Hummerstone LG. 1980. The use of biological indicators of heavy-metal contamination in estuaries. *Marine Biological Association, Occasional Publications*, No 1, p. 73.
- Burningham H, French J. 2006. Morphodynamic behaviour of a mixed sand- gravel ebb-tidal delta: Deben estuary, Suffolk, UK. *Marine Geology*, 225, 23-44. DOI: 10.1016/j.margeo.2005.09.009.
- Carrasco AR, Ferreira Ó, Matias A, Freire P. 2012. Natural and human-induced coastal dynamics at a back-barrier beach. *Geomorphology*, 159/160, 30-36. DOI: 10.1016/j.geomorph.2012.03.001.
- Collins BD, Sitar N. 2008. Processes of coastal bluff erosion in weakly lithified sands, Pacifica, California, USA. *Geomorphology*, 97, 483-501. DOI: 10.1016/j.geomorph.2007.09.004.
- Cowell PJ, Thom BG. 1994. Morphodynamics of coastal evolution. In: Carter RWG, Woodroffe CD (Eds.) *Coastal Evolution*. Cambridge University Press, Cambridge, pp. 33-86.
- Crowel M, Leatherman SP, Buckley MK. 1991. Historical shoreline change- error analysis and mapping accuracy. *Journal of Coastal Research*, 7, 839-852.
- Davidson, N. C., Laffoley, D. D., Way, L. S., Key, R., Drake, C. M., Pienkowski, M. W., Mitchell, R., & Duff, K. L. 1991, *Nature conservation and estuaries in Great Britain*,

- Nature Conservancy Council, Peterborough, UK. DOI: 10.1002/aqc.3270010109.
- Dobson JE. 1992. Global challenges demand more precision, better generalization. *GIS World*, 68-69.
- Draut AE, Logan JB, Mastin MC. 2011. Channel evolution on the dammed Elwha River, Washington, USA. *Geomorphology*, 127, 71-87. DOI: 10.1016/j.geomorph.2010.12.008.
- Esteves LS, William JJ, Brown JM. 2011. Looking for evidence of climate change impacts in the eastern Irish Sea. *Natural Hazards and Earth System Sciences*, 11, 1641-1656. DOI: 10.5194/nhess-11-1641-2011.
- Esteves LS, William JJ, Nock A, Lymbery G. 2009. Quantifying shoreline changes along the Sefton coast (UK) and the implications for research-informed coastal management. *Journal of Coastal Research*, S156, 602-606.
- Faiz S, Boursier P. 1996. Geographic data quality: From assessment to exploitation. *Cartographica*, 33(1), 33-40. DOI: 10.3138/66R3-U1RO-1652-4Q06.
- Fenster MS, Dolan R, Elder JF. 1993. A new method for predicting shoreline positions from historical data. *Journal of Coastal Research*, 9(1), 147-171.
- González-Villanueva R, Costas S, Pérez-Arlucea M, Jerez S, Trigo RM. 2013; Impact of atmospheric circulation patterns on coastal dune dynamics, NW Spain. *Geomorphology*, 185, 96-109. DOI: 10.1016/j.geomorph.2012.12.019.
- Hackney C, Darby SE, Leyland J. 2013. Modelling the response of soft cliffs to climate change: A statistical, process-response model using accumulated excess energy. *Geomorphology*, 187, 108-121. DOI: 10.1016/j.geomorph.2013.01.005.
- Hapke CJ, Kratzmann MG, Himmelstoss EA. 2013. Geomorphic and human influence on large-scale coastal change. *Geomorphology*, 199, 160-170. DOI: 10.1016/j.geomorph.2012.11.025.
- Hapke CJ, Himmelstoss EA, Kratzmann M, List J, Thieler ER. 2011. National assessment of shoreline change. Historical shoreline change along the New England and Mid-Atlantic Coasts. *US Geological Survey Open-file Report* 2010-1118: <http://pubs.usgs.gov/of/2010/1118/>.
- Himmelstoss EA, Kratzmann M, Hapke CJ, Thieler ER, List J. 2011. The national assessment of shoreline change — a GIS compilation of vector shorelines and associated shoreline change data for the New England and Mid-Atlantic coasts. *US Geological Survey Open-File Report* 2010-1119: <http://pubs.usgs.gov/of/2010/1119/>
- Houser C, Mathew S. 2011. Alongshore variation in foredune height in response to transport potential and sediment supply: South Padre Island, Texas. *Geomorphology*, 125, 62-72. DOI: 10.1016/j.geomorph.2010.07.028.
- Houser C, Hapke C, Hamilton S. 2008. Controls on coastal dune morphology, shoreline erosion and barrier island response to extreme storms. *Geomorphology*, 100, Volume, 100, 223-240. DOI: 10.1016/j.geomorph.2007.12.007
- HR Wallingford, Pmer AB, Pethick J. 2006. *Review and formalisation of geomorphological concepts and approaches for estuaries*. R&D Technical Report FD2116/TR2. Joint DEFRA EA Flood and Coastal Erosion Risk Management R&D Programme: http://www.estuary-guide.net/pdfs/FD2116_TR2.pdf.
- Jabaloy-Sánchez A, Lobo FJ, Azor A, Martín-Rosales W, Pérez-Peña JV, Bárcenas P, Macías JM, Fernández-Salas LM, Vázquez-Vílchez M. 2014. Six thousand years of coastline evolution in the Guadalfeo deltaic system (southern Iberian Peninsula). *Geomorphology Vol. 206*, 374 – 391. DOI: 10.1016/j.geomorph.2013.08.037.
- Jones C. 1997. *Geographical Information Systems and Computer Cartography* Essex, Addison Wesley Longman Limited.
- Katz O, Mushkin A. 2013. Characteristics of sea-cliff erosion induced by a strong winter storm in the eastern Mediterranean. *Quaternary Research*, 80, 20-32. DOI: 10.1016/j.yqres.2013.04.004.
- Lawrence, PL. 1994. Natural hazards of shoreline bluff erosion: a case study of Horizon View, Lake Huron. *Geomorphology*, Volume 10, Issues 1 – 4, 65 – 81. DOI: 10.1016/0169-55X(94)90008-6.
- Leyland J, Darby SE. 2008. An empirical-conceptual gully evolution model for channelled sea cliffs. *Geomorphology*, 102, 419-434. DOI: 10.1016/j.geomorph.2008.04.017.
- Lisitzin E. 1974. *Sea level changes. Oceanography Series*, 8, Amsterdam: Elsevier, 286 pp.
- Loos EA, Niemann KO. 2002. Shoreline feature extraction from remotely sensed imagery. *IEEE*

- International*, 6(24-28), 3417-3419. DOI: 10.1109/IGARSS.2002.1027201.
- Maiti S, Bhattacharya AK. 2009. Shoreline change analysis and its application to prediction: a remote sensing and statistics based approach. *Marine Geology*, 257, 11-23. DOI: 10.1016/j.margeo.2008.10.006.
- Montreuil A-L, Bullard JE. 2012. A 150-year record of coastline dynamics within a sediment cell: Eastern England. *Geomorphology*, 179, 168-185. DOI: 10.1016/j.geomorph.2012.08.008.
- Moore LJ. 2000. Shoreline mapping techniques. *Journal of Coastal Research*, 16, 111-124.
- Moore LJ, Griggs GB. 2002. Long-term cliff retreat and erosion hotspots along the central shores of the Monterey Bay National Marine Sanctuary. *Marine Geology*, 181, 265-283. DOI: 10.1016/S0025-3227(01)00271-7.
- Nebel SH, Trembanis AC, Barber DC. 2012. Shoreline analysis and barrier island dynamics: Decadal scale patterns from Cedar Island, Virginia. *Journal of Coastal Research*, 28(2), 332-341. DOI: 10.2112/JCOASTRES-D-10-00144.1.
- Oyedotun TDT, Burningham H, French JR. 2012. Characterisation of estuary and adjacent beach sediments in the Gannel Estuary, South-West England. *Geoscience in South-West England*, 13, 70-76.
- Pardo-Pascual JE, Almonacid-Caballer J, Ruiz LA, Palomar-Vázquez J. 2012. Automatic extraction of shorelines from Landsat TM and ETM+ multi-temporal images with subpixel precision. *Remote Sensing of Environment*, 123, 1-11. DOI: 10.1016/j.rse.2012.02.024.
- Paskoff R, Clus-Auby C. 2007. *L'érosion des plages. Les causes, les remèdes*. Institut Océanographique Éditeur: Paris, 184pp.
- Pirrie D, Power MR, Wheeler PD, Cundy A, Bridges C, Davey G. 2002. Geochemical signature of historical mining: Fowey Estuary, Cornwall, UK. *Journal of Geochemical Exploration* 76, 31-43. DOI: 10.1016/S0375-6742(02)00203-0.
- Pirrie D, Power MR, Payne GS, Wheeler PD. 2000. Impact of mining on sedimentation: The Camel and Gannel estuaries, Cornwall. *Geoscience in South-West England*, 10, 021-028.
- Price W A, Kendrick MP. 1963. Field and model investigation into the reasons for siltation in the Mersey Estuary. *Proceedings of the Institution of Civil Engineers*, Vol. 24, pp. 473-518. DOI: 10.1680/iicep.1963.10702.
- Pugh D. 2004. *Changing Sea Levels. Effects of Tides, Weather and Climate*. Cambridge University Press, 280pp.
- Pugh D. 1996. *Tides, Surges and Mean Sea Level*. Chichester, UK: John Wiley & Sons, 472pp.
- Quinn JD, Rosser NJ, Murphy W, Lawrence JA. 2010. Identifying the behavioural characteristics of clay cliffs using intensive monitoring and geotechnical numerical modelling. *Geomorphology*, 120, 107-122. DOI: 10.1016/j.geomorph.2010.03.004.
- Reid C, Scrivenor JB. 1906. The geology of the country near Newquay. *Memoir of the Geological Survey, England and Wales: Explanation of Sheet 346*.
- Restrepo AJD. 2012. Assessing the effect of sea-level and human activities on a major delta on the Pacific coast of northern South America: The Patía River. *Geomorphology*, 151/152, 207-223. DOI: 10.1016/j.geomorph.2012.02.004.
- Rio LD, Gracia JF, Benavente J. 2013. Shoreline change patterns in sandy coasts. A case study in SW Spain. *Geomorphology*, 196, 252-266. DOI: 10.1016/j.geomorph.2012.07.027.
- Rio LD, Gracia JF. 2009. Erosion risk assessment of active coastal cliffs in temperate environments. *Geomorphology*, 112, 82-95. DOI: 10.1016/j.geomorph.2009.05.009.
- Romine BM, Fletcher CH, Frazer LN, Genz AS, Barbee MM, Lim SC. 2009. Historical shoreline change, southeast Oahu, Hawaii; applying polynomial models to calculate shoreline change rates. *Journal of Coastal Research*, 25(6), 1236-1253. DOI: 10.2112/08-1070.1.
- Ryu J-H, Won J-S, Min KD. 2002. Waterline extraction from Landsat TM data in a tidal flat: a case study in Gosmo Bay, Korea. *Remote Sensing of the Environment*, 83, 442-456. DOI: 10.1016/S0034-4257(02)00059-7.
- Sherman DJ, Bauer BO. 1993. Coastal geomorphology through the looking glass. *Geomorphology*, 7, 225-249. DOI: 10.1016/B978-0-444-89971-2.50014-6.
- Stockdon HF, Sallenger AH, List JH, Holman RA. 2002. Estimation of shoreline position and change from airborne topographic LiDAR data. *Journal of Coastal Research*, 18, 502-513.

- Taylor, J.A., Murdock, A.P., & Pontee, N.I. 2004. A macroscale analysis of coastal steepening around the coast of England and Wales. *The Geographical Journal*, 170, (3) 179-188. DOI: 10.1111/j.0016-7398.2004.00119.x.
- Thébaudeau B, Trenhaile AS, Edwards RJ. 2013. Modelling the development of rocky shoreline profiles along the northern coast of Ireland. *Geomorphology*, 203, 66-78. DOI: 10.1016/j.geomorph.2013.03.027.
- Thieler ER, Danforth WW. 1994a. Historical shoreline mapping (1) Improving techniques and reducing positioning errors. *Journal of Coastal Research*, 10, 549-563.
- Thieler ER, Danforth WW. 1994b. Historical shoreline mapping (II) Application of the Digital Shoreline Mapping and Analysis Systems (DSMS/DSAS) to shoreline change mapping in Puerto Rico. *Journal of Coastal Research*, 10(3), 600-620.
- Thieler ER, Himmelstoss EA, Zichichi JL, Ergul A. 2009. The Digital Shoreline Analysis System (DSAS) Version 4.0 - An ArcGIS Extension for Calculating Shoreline Change. Open-File Report. US Geological Survey Report No. 2008-1278: <http://woodshole.er.usgs.gov/project-pages/dsas/version4/>
- Uluocha NO. 2007. *Elements of Geographic Information Systems* Lagos. Sam Ironusi Publications.
- USGS. 2013. DSASweb: <http://marine.usgs.gov/dsasweb>.
- Weber KM, List JH, Morgan KM. 2005. An operational mean high water datum for determination of shoreline position from topographic LiDAR data. *US Geological Survey Open-File Report* 2005-1027: <http://pubs.usgs.gov/of/2005/1027/html/intro.html>.
- Wolman MG, Miller JP. 1960. Magnitude and frequency of forces in geomorphic processes. *Journal of Geology*, 68, 54-74.
- Young AP, Flick RE, O'Reilly WC, Chadwicj DB, Crampton WC, Helly JJ. 2014. Estimating cliff retreat in southern California considering sea level rise using a sand balance approach. *Marine Geology*, 348, 15-26. DOI: 10.1016/j.margeo.2013.11.007.
- Young AP, Guza RT, Flick RE, O'Reilly WC, Gutierrez R. 2009. Rain, waves and short-term evolution of composite sea cliffs in southern California. *Marine Geology*, 267, 1-7. DOI: 10.1016/j.margeo.2009.08.008.
- Zuzek PJ, Nairn RB, Thieme SJ. 2003. Spatial and temporal consideration for calculating shoreline change rates in the Great Lakes Basin. *Journal of Coastal Research*, 38, 125-146.

Surf Zone Hydrodynamics: Measuring Waves and Currents

Kris W. Inch¹

¹Coastal Processes Research Group, School of Marine Science and Engineering, University of Plymouth (kris.inch@plymouth.ac.uk)



ABSTRACT: Wave breaking is the most significant energy input into wave-dominated coastal environments and is responsible for driving a number of interrelated, hydrodynamic processes in a region known as the surf zone. Surf zone hydrodynamics occur over a range of spatial and temporal scales and are constantly changing. As a result, obtaining measurements of waves and currents in the surf zone can be challenging and field experiments require detailed planning to ensure the collection of useful data. Measurements of surf zone hydrodynamics are divided into two general categories: (1) non-directional wave measurements; and (2) directional wave and current measurements. Near-bed pressure transducers are by far the most common method of measuring the non-directional properties of waves. Acoustic sensors are being used with increasing frequency for Eulerian measurements of surf zone currents, whereas Lagrangian surf zone drifters are most useful for measuring flow patterns in rip currents, surf zone circulation and alongshore currents. This article gives a summary of the various methods of measuring waves and currents in the surf zone. Additionally, simple data analysis techniques to study infragravity waves are demonstrated.

KEYWORDS: surf zone, breaking waves, pressure transducer, current meters, infragravity waves.

Introduction

The surf zone is that part of the shoreface extending from the seaward boundary of wave breaking to the swash zone; the part of the beach that is alternately covered and exposed by wave uprush and backwash. Breaking waves drive a number of interrelated surf zone processes such as the creation of turbulent bores, wave set-up, nearshore currents and low frequency motions. Together these processes force the entrainment and transport of sediment, which leads to morphological change. For a comprehensive overview of surf zone processes see Komar (1998), or the introductory texts of Aagaard and Masselink (1999), Woodroffe (2002), Davidson-Arnott (2010) and Masselink *et al.* (2011) for a more succinct summary.

The width and characteristics of the surf zone vary constantly, driven by changes in the tide elevation, incident wave height and direction, low-frequency motions and local wind speed. The beach slope also plays an important role

in determining the breaker type and dissipative characteristics of the surf zone. Three types of breaking waves are commonly recognised (Figure 1); spilling, plunging and surging (Galvin, 1968). Spilling waves typically occur on low gradient beaches and dissipate their energy gradually over a wide surf zone. With plunging breakers, the shoreward face of the wave steepens until it is vertical and the crest curls over and plunges forward and downward as an intact mass of water. Plunging waves are more energetic than spilling waves at the point of breaking and are normally associated with narrower surf zones and steeper beaches. Surging waves are found on steep, reflective beaches where there is often no clear surf zone as the waves slide up the beach without physically breaking. In reality, there is a continuum of breaker types blending from one to another and on a natural beach with a spectrum of wave heights and periods it is common to see a range of breaker types at a given time. A number of dimensionless parameters have been developed to predict

the breaker type and surf zone state. The most widely used of these is the Iribarren number ξ defined as

$$\xi = \frac{\tan\beta}{\sqrt{H_b/L_o}}$$

where β is the beach slope, H_b is the breaker height and L_o is the deep water wavelength given by linear wave theory. Spilling waves occur when $\xi < 0.4$, plunging waves occur when $0.4 < \xi < 2.0$, and surging waves occur when $\xi > 2.0$ (Battjes, 1974).

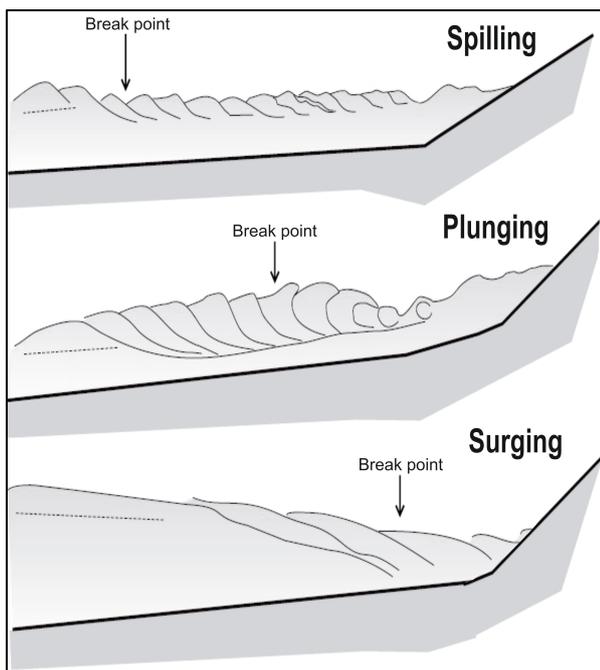


Figure 1: The three types of breaking wave. Note the different break point locations in relation to the shoreline. Modified from Davidson-Arnott (2010).

In the inner surf zone of low gradient, dissipative beaches, the incident wave height H is controlled by the local water depth (Thornton and Guza, 1982) and can be expressed by

$$H = \langle \gamma \rangle h$$

where γ is a coefficient ranging from 0.3-0.6 and increases with beach slope, and h is water depth. When the wave height is limited by the local water depth the surf zone is considered to be saturated. Under saturated conditions, to maintain a constant relationship between wave height and water depth, an increase in offshore wave height acts only to increase the width of the surf zone. However, on steep, sandy beaches plunging waves

break close to shore creating a narrow surf zone that can remain unsaturated.

The inner surf zone is of particular importance as the conditions here force the hydro and sediment dynamics in the swash zone, which is arguably the most dynamic part of the nearshore region (Masselink and Puleo, 2006). Water motion in the inner surf zone is typically dominated by infragravity waves, particularly on dissipative beaches. Infragravity waves are low-frequency (0.005-0.05 Hz) waves that approach the shoreline bound to incident wave groups but become free waves in the surf zone following incident wave breaking. Infragravity waves are an important research topic for coastal scientists as they play an important role in beach and dune erosion, especially during storms (Russell, 1993). Due to their long wavelengths, which prevent breaking, infragravity waves reflect from the shoreline and travel seawards giving rise to a cross-shore quasi-standing wave pattern (Guza and Thornton, 1985). Infragravity energy in the surf zone increases with increasing offshore wave height; however, recent studies have revealed that under very energetic wave conditions infragravity waves may also become saturated (e.g. Senechal *et al.*, 2011a; Guedes *et al.*, 2013; De Bakker *et al.*, 2014).

In addition to wave motion, three types of quasi-steady, wave-induced currents exist in the surf zone; bed return flow, alongshore currents and rip currents. These currents exist simultaneously and are driven by cross-shore and alongshore gradients in the mean water level caused by variations in the wave breaker height. The intensity of nearshore currents increases with increasing incident wave height. Thus, the strongest currents capable of transporting vast quantities of sediment occur during storms (Senechal *et al.*, 2011b). Additionally, rip currents pose a significant hazard to water users and are a major cause of lifeguard rescues around the world (Short and Brander, 1999; Scott *et al.*, 2007, 2008).

Measurements of surf zone processes are of vital importance for estimating potential storm damage, modelling shoreline evolution and designing shoreline management plans. Waves and currents are well documented in the coastal literature; advances in instrument

technology and analysis techniques have improved our knowledge considerably over the last few decades. However, certain aspects remain poorly understood, such as surf zone hydrodynamics during extreme storms. The purpose of this entry is to provide an overview of the current *in situ* methods of measuring surf zone waves and currents, and to give a simple demonstration of wave data analysis.

Surf Zone Measurements

For convenience, the following section on surf zone measurements is split into two main categories: (1) non-directional wave measurement; and (2) directional wave and current measurement. Non-directional wave measurements are measurements of the water surface elevation from which information about wave height and period can be obtained. Current measurements are measurements of the velocity vector and are used to study nearshore currents and the directional properties of waves.

Non-directional wave measurement

Field measurements of waves have previously been collected using a form of surface piercing wave staff (among others, Thornton and Guza, 1983; Davidson-Arnott and McDonald, 1989). Wave staffs are part of an electronic circuit that takes advantage of the conductivity of sea water by recording the change in electrical resistance or capacitance as waves pass the instrument. These recordings are then converted to measurements of water depth using calibrated signals from known water depths. The sensors can be attached to a temporary support driven into the sand or attached to a more permanent structure such as a pier. The advantage of using wave staffs is that they provide a direct measure of the sea surface elevation. The main disadvantage of wave staffs is that they are exposed to wave action when deployed in shallow water or in energetic surf zones. This can damage wave staffs or displace them from their intended vertical positions, thus causing error in the water depth measurements. Furthermore, mounting requirements restrict the use of wave staffs at certain locations and may limit the deployment of additional sensors at the same location.

Nowadays, almost all surf zone field experiments use bottom-mounted pressure transducers. Pressure transducers measure pressure variations associated with passing waves above and this pressure is converted (either by the sensor or in post-processing) into the equivalent water depth. Pressure transducers are most commonly attached to a temporary frame driven into the sand (Figure 2a), but can also be fixed to permanent features such as shore platforms or buried beneath the surface of the beach. An advantage of using buried pressure transducers is to avoid corruption of the signal caused by dynamic pressure variations from accelerating and decelerating flows (Austin *et al.*, 2014). Unlike wave staffs, pressure transducers are not directly exposed to wave action. They are also relatively easy to deploy, can be co-located with other sensors and are often self-logging eliminating the need for cables.

The main limitation of using pressure transducers is depth attenuation of the pressure signal. The extent of depth attenuation is frequency dependant; as depth increases high frequency signals are lost before low frequency signals. Depth attenuation can be corrected for during post-processing by applying a frequency dependant depth attenuation factor $K(f)$ derived from linear wave theory as:

$$K(f) = \frac{\cosh(kz)}{\cosh(kh)}$$

where k is the frequency dependant wave number ($2\pi/L$, where L is wavelength) and z is the height of the pressure sensor above the bed. A consequence of using this method to correct for depth attenuation is that it introduces high frequency noise, which can cause significant error at the data analysis stage. Therefore a high frequency cut-off should be chosen beyond which energy in the wave spectrum (see *Wave Data Analysis*) is discarded. The high frequency cut-off should be chosen carefully depending on factors such as the sensor depth and local wave climate. However, an alternative method of determining the high frequency cut-off ω_c is by

$$\omega_c = 0.564\pi\sqrt{(g/d)}$$

(Green, 1999; Aagaard *et al.*, 2002) where ω is radian frequency ($2\pi f$, where f is frequency), g is gravitational acceleration

(9.81 m s^{-2}) and d is the depth of the pressure sensor. Pressure transducers deployed in the surf zone tend to be located in relatively shallow water compared to open ocean deployments. Nonetheless, correcting for depth attenuation is an important process, especially in macrotidal regions and low fetch environments where high frequency wind waves prevail. In addition to depth attenuation, there is also the need to correct for atmospheric pressure. Some pressure transducers do this automatically using the mean atmospheric pressure at sea level (10.1325 dbar); however, this does not account for local variations in atmospheric pressure or short term fluctuations during an experiment, for which local pressure observations are required.

An additional source of error in water depth measurements provided by pressure transducers arises from the unknown height of the sensor above the bed. If this offset is measured at the start of an experiment and at every opportunity thereafter, a linear function may be used to correct for bed level change between measurements. However, the results of recent studies in the inner surf zone (e.g. Puleo *et al.*, 2014) suggest that net bed level change over a tidal cycle may not be linear but rather the result of just a few “large” waves. A way to overcome this is by co-locating the pressure transducer with a sensor, such as an ultrasonic distance meter,

providing continuous measurements of bed level elevation (e.g. Ridd, 1992; Saulter *et al.*, 2003; Arnaud *et al.*, 2009; Puleo *et al.*, 2010; *et al.*, 2014). Whilst the unknown height of the pressure transducer can cause error in measurements of total water depth, Ruessink (1999) suggests that the impact of these uncertainties on wave height calculations is $< 10\%$.

Directional wave and current measurement

Wave staffs and pressure transducers provide a time series of the water surface elevation, which is useful for studying wave height and period. However, in order to investigate directional wave properties and nearshore currents, the horizontal component of water motion also needs to be measured. Current data can broadly be divided into two categories: (1) Eulerian observations; and (2) Lagrangian observations. Eulerian methods measure fluid flow at a fixed location through time, whereas Lagrangian methods follow fluid parcels through space and time. Eulerian observations are collected by *in situ* current meters, often co-located with pressure transducers and referred to as PUV set-ups (Figure 2b) due to the three quantities measured; pressure (P) with cross-shore (U) and alongshore (V) velocities (Morang *et al.*, 1997).

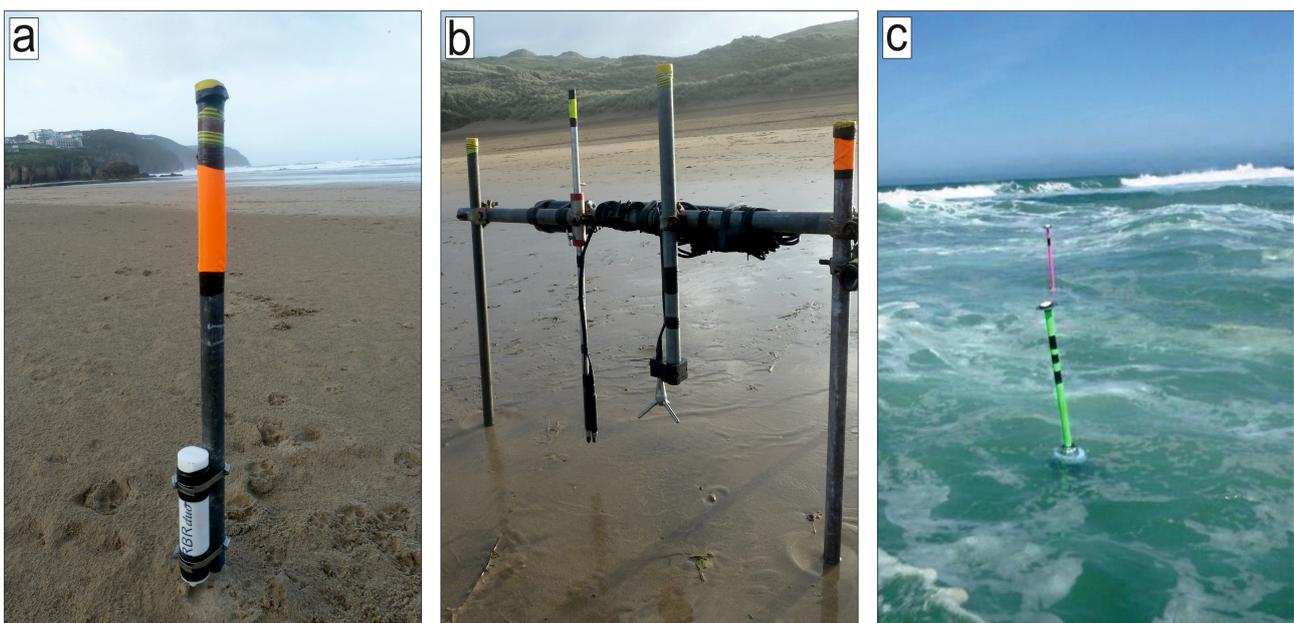


Figure 2: Examples of (a) a bottom-mounted pressure transducer, (b) a PUV rig equipped with an acoustic Doppler velocimeter, and (c) surf zone drifters. (Photos: K. Inch and E. Woodward)

These types of measurements are useful for studying directional wave properties and near-bed currents. Lagrangian observations are most useful in studies of rip currents, surf zone circulation and surf zone flushing where current magnitude and direction vary spatially in the surf zone.

Eulerian observations

Since their introduction in the 1970s, electromagnetic current meters (EMCMs) have been widely used to measure currents in the surf zone. EMCMs measure the voltage generated by waves passing through a fluctuating magnetic field produced by the sensor. Faraday's law is then used to convert the voltage into the proportional current velocity along two perpendicular axes (u,v). Over the past decade or so, acoustic Doppler velocimeters (ADV) have become a popular alternative to EMCMs. ADVs operate by emitting acoustic signals into the water column and recording the return signal backscattered by fine material a short distance from the sensor. Current velocity is calculated using the phase lag between successive return signals. ADVs are seen as an advancement on EMCMs as they measure three-dimensional (u,v,w) velocity components and are non-intrusive (MacVicar *et al.*, 2007). Also, EMCMs are prone to error if the sensor is located too close to the seabed or free surface, drift of the zero offset and electronic interference (Guza *et al.*, 1988). A significant problem with using ADVs in the surf zone, however, is their high sensitivity to bubbles and sediment which results in poor data (determined by the signal to noise ratio) being removed in quality control procedures (Elgar *et al.*, 2005; Feddersen, 2010). The lower sensitivity of EMCMs means that they remain a practical and much used instrument in surf zone studies and, under highly turbulent conditions, may outperform ADVs (e.g. Rodriguez *et al.*, 1999). Elgar *et al.* (2001) simultaneously deployed EMCMs and ADVs on the same frame in the surf zone to provide a direct and detailed comparison of their performance. They found that the data were highly correlated in velocities up to 3 m s^{-1} , confirming the ability of both instruments to perform well at measuring surf zone currents.

There has been increasing interest in recent years in using pulse-coherent acoustic

Doppler profilers (ADPs) to monitor surf zone hydrodynamics (e.g. Senechal *et al.*, 2011b). ADPs operate using the same basic principles as ADVs and are therefore also highly sensitive to bubbles and sediment. However, by measuring the return signal in much smaller time increments, ADPs determine the velocity vector for a series of discrete "bins" over a portion of the water column. This makes ADPs especially useful for studying boundary layer dynamics in the surf zone. For example, Puleo *et al.* (2012) used an ADP to investigate the inner surf zone boundary layer at a spatial resolution of 1 mm and a temporal resolution of 100 Hz on a microtidal beach in Florida, USA.

Lagrangian observations

A Lagrangian method which is particularly popular in Australian research is releasing a bright, inert dye into the surf zone and tracking its movement (e.g. Huntley *et al.*, 1988; Brander, 1999; Brander and Short, 2001). Dye tracking is useful in providing qualitative data on the location and path of nearshore currents and is a valuable tool in beach safety education. However, it does not provide any data on actual current velocities. An increasingly popular technique for measuring the flow patterns in nearshore currents is using GPS-tracked surf zone drifters (Figure 2c). Surf zone drifters are buoyant, PVC tubes equipped with on-board GPS data loggers which record the drifter's position in the surf zone at typically 0.5-1 Hz (Schmidt *et al.*, 2003). The drifters are equipped with a damping plate designed to allow broken waves to pass without rapidly transporting the drifter onshore (Schmidt *et al.*, 2003). Accuracies of $< 0.4 \text{ m}$ in position and $< 0.01 \text{ m s}^{-1}$ in velocity can be achieved by post-processing the raw GPS data from a static base position (MacMahan *et al.*, 2009). The combined tracks of multiple drifters provide detailed information on mean nearshore flow patterns and velocities (e.g. Austin *et al.*, 2010, 2014). Surf zone drifters have been observed to closely follow simultaneous dye releases as well as provide velocity measurements in good agreement with those from *in situ* current meters, thus confirming their ability to make valid Lagrangian observations of surf zone currents (Schmidt *et al.*, 2003; Johnson and Pattiaratchi, 2004; MacMahan *et al.*, 2009). A recent study by McCarroll *et al.* (2014) used

34 surf zone drifters to investigate rip current behaviour on an embayed beach in New South Wales, Australia. They were able to make 293 individual drifter deployments over a single ebbing tidal cycle, thus providing a detailed observation of the nearshore flow patterns (Figure 3).

Remote sensing

It is beyond the scope of this entry to discuss remote sensing methods in detail; however, it should be acknowledged that remote sensing techniques are being used with increasing frequency to study nearshore waves and currents (for a review, see Holman and Haller, 2013). These methods include video

imagery (e.g. Holman *et al.*, 2006; Holman and Stanley, 2007; De Vries *et al.*, 2011), X-band radar (e.g. Ruessink *et al.*, 2002; Catalan *et al.*, 2014; Haller *et al.*, 2014) and LIDAR (e.g. Blenkinsopp *et al.*, 2012). Possible advantages of using remote sensing methods include better synoptic measurements, lower maintenance costs, improved robustness and longer deployments including during storms. A potential disadvantage of remote sensing is that measurements are often based on a proxy rather than a direct measurement; therefore, the accuracy of the measurements depends on how well the proxy represents the variable being studied.

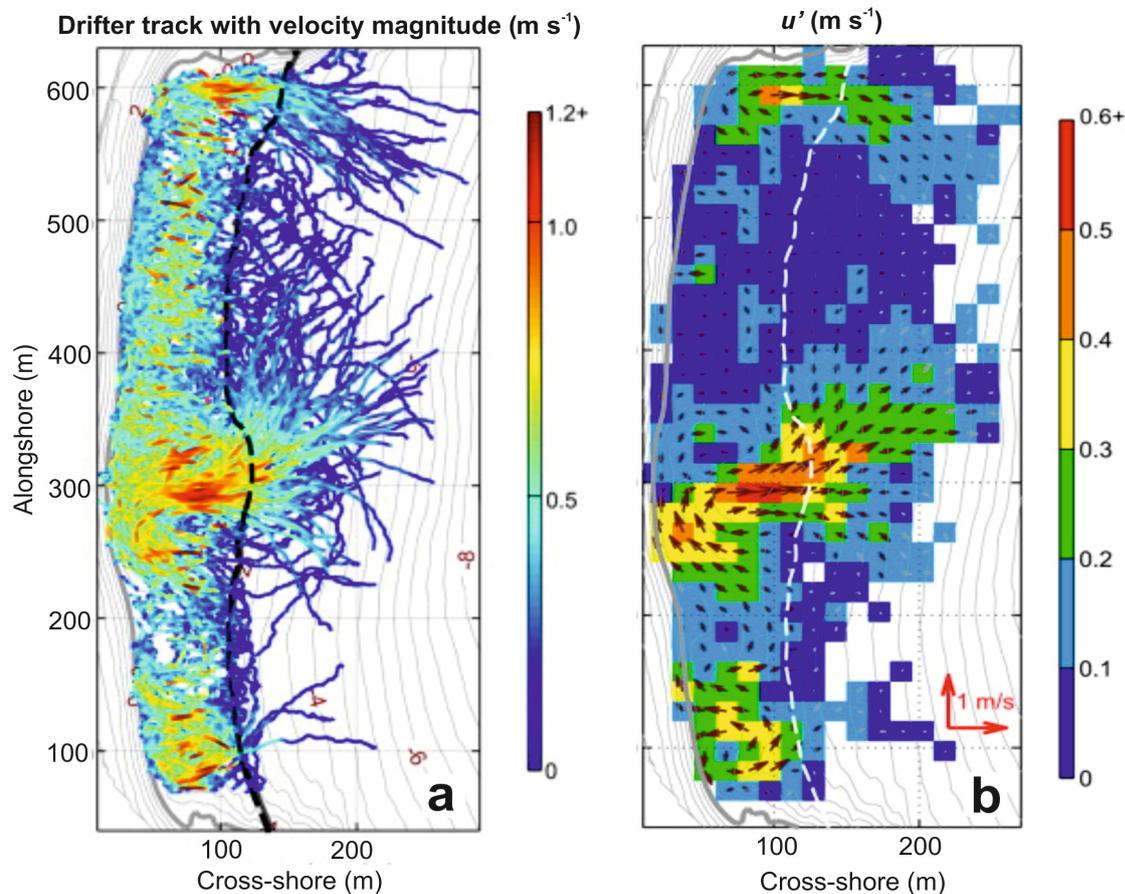


Figure 3: Drifter tracks coloured by velocity magnitude (a), and mean velocity magnitude and direction determined from individual drifter observations (b). Contours are at 0.5 m spacing and the edge of the surf zone is indicated by the black dashed line in (a) and white dashed line in (b). (Source: McCarroll *et al.*, 2014)

Wave Data Analysis

The analysis of surf zone data is very complex and it is far beyond the scope of this article to discuss all types of analysis techniques. The most appropriate analysis procedure depends on the aim of the

research and the type of data available. In this section, the analysis of wave and current data to investigate infragravity waves in the surf zone (see *Introduction*) is discussed as an example. Data collected from the inner surf zone of Perranporth Beach, Cornwall, UK, is used to demonstrate the analysis

techniques. Perranporth is a macrotidal, dissipative beach with a mean tidal range of 5.4 m and is composed of medium sand. Observations of pressure and velocity (u, v, w) were logged continuously at 4 Hz using a co-located pressure transducer and ADV during a large storm in October 2013. The instruments were mounted onto a temporary scaffold frame driven into the sand in the intertidal zone (Figure 2b). The height of the instruments above the bed was measured before and after each tide and a linear function was used to correct for changes in bed level. The pressure data were converted to water surface elevation, with a depth correction using linear wave theory and a high frequency cut-off of 1 Hz (see *Surf Zone Measurements*), and detrended prior to analysis.

There are generally two approaches to analysing wave data; analysis in the time domain and in the frequency domain.

Time domain analysis

The irregular nature of natural waves is evident in the example time series of water surface elevation and cross-shore velocity shown in Figure 4a and 4b. The two time series have limited value in their raw state and are best described quantitatively using statistics. The two most common statistical parameters used to describe a wave time series are the significant wave height H_s (also $H_{1/3}$) which is the mean of the highest one third of waves, and the peak wave period T_p which is the wave period associated with the maximum wave energy derived from the wave spectrum (described below). Other commonly used wave parameters include the mean wave height \bar{H} , maximum wave height H_{max} and zero-crossing wave period T_z . A summary of the various wave parameters that can be derived from both time and frequency domain analysis is given by Morang *et al.* (1997).

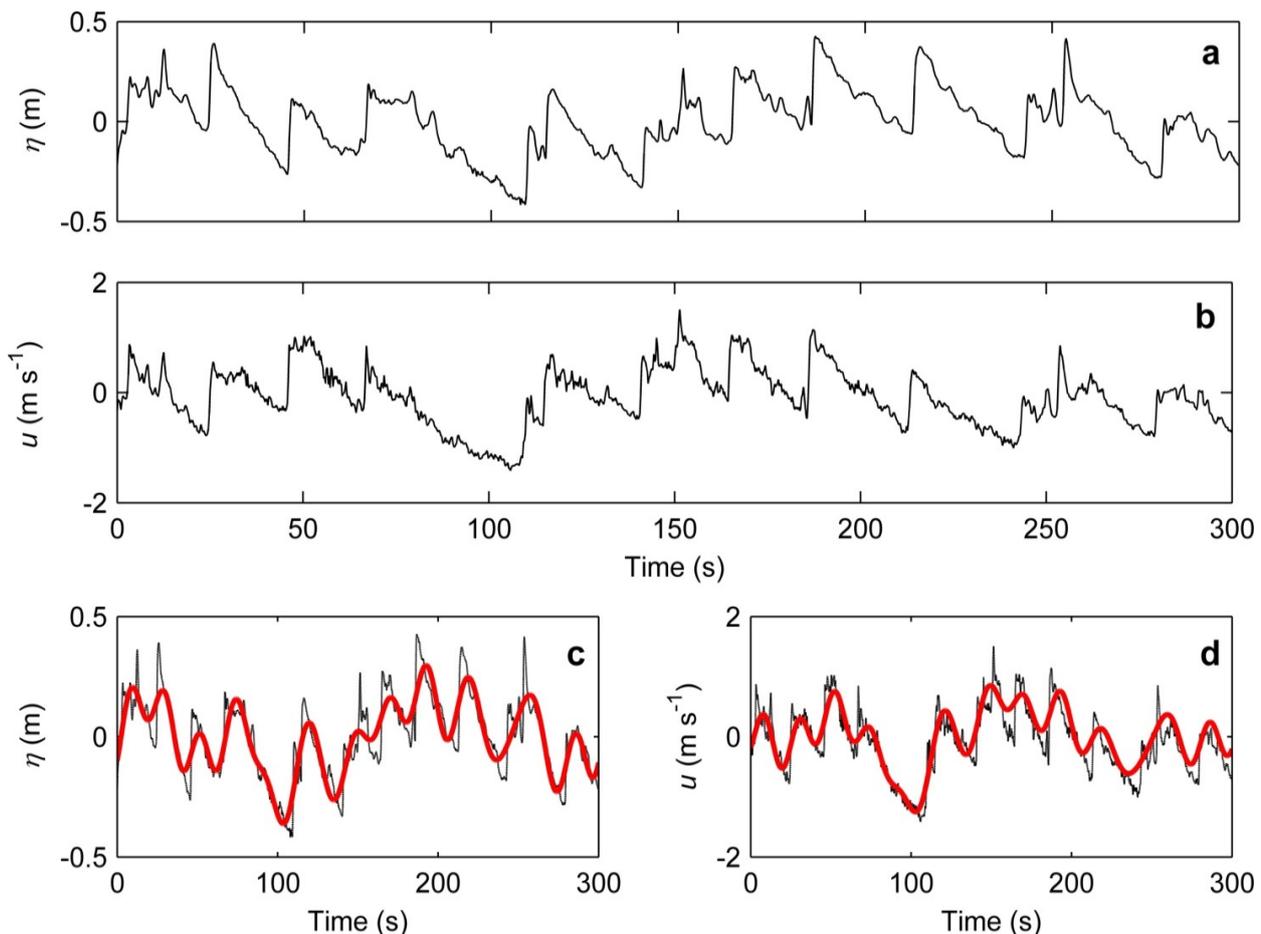


Figure 4: Raw time series (5 mins) of water surface elevation η (a) and cross-shore velocity u (b) collected in the inner surf zone of Perranporth Beach during storm conditions with an offshore significant wave height of 4.31 m. Low pass filtered (0.005-0.05 Hz) water surface elevation (c) and cross-shore velocity (d) time series (red) plotted with the original time series (black).

One method to obtain these wave parameters is through wave-by-wave analysis using the zero-downcrossing method. The zero-downcrossing method defines an individual wave by two successive downward crossings of the mean water level by the water surface elevation. Alternatively, Longuet-Higgins (1952) proposed a method based on the Rayleigh distribution, which calculates the various wave height parameters using the standard deviation of the water surface elevation time series. Significant wave height, for example, can be estimated from

$$H_s = 4\sigma_\eta$$

where σ_η is the standard deviation of the water surface elevation. To calculate wave statistics over particular frequency ranges (e.g. infragravity and incident), high- and low-pass filters can be applied to the original time series. Figure 4c and 4d shows the filtered infragravity components plotted with the original time series from Figure 4a and 4b. The infragravity time series follow closely the original time series for both water surface elevation and cross-shore velocity. This is a clear indication that the dominant water motion is at infragravity frequencies and the calculated significant wave heights reflect this; 0.58 m and 0.30 m for the infragravity and incident bands respectively. This is typical of water motion in the inner surf zone of dissipative beaches.

As mentioned in the introduction, the long wavelengths of infragravity waves inhibit wave breaking allowing some energy to reflect from the shoreline. Therefore, the low frequency signal of a wave time series will comprise both incoming (shoreward) and outgoing (seaward) components. Guza *et al.* (1984) proposed a method to decompose the shoreward and seaward propagating wave signals in the time domain using the water surface elevation and cross-shore velocity by:

$$\eta_{in} = \frac{\eta + \sqrt{h/g} u}{2}$$

$$\eta_{out} = \frac{\eta - \sqrt{h/g} u}{2}$$

where η_{in} and η_{out} are the incoming and outgoing wave components, η is water surface elevation and u is cross-shore velocity. This method was used to separate the infragravity wave signal in Figure 4c into

its incoming and outgoing components and corresponding significant wave heights were calculated. The significant wave height of the outgoing infragravity time series was 0.23 m, indicating that some wave energy was indeed reflected. It is also possible to separate the incoming and outgoing wave components in the frequency domain (discussed below).

Frequency domain analysis

Wave data analysis in the frequency domain is achieved through spectral analysis. Spectral analysis is based on the fast Fourier transformation, which assumes that a time series is composed of a finite number of sinusoids at discrete frequencies. Spectral analysis partitions a time series into its constituent parts and produces an auto-spectrum (sometimes referred to as a wave spectrum); a plot of wave variance (proportional to wave energy) as a function of frequency. An outline and worked example of the various steps involved to produce a wave spectrum is given by Hegge and Masselink (1996). Figure 5a-c shows the water surface elevation, cross-shore and alongshore velocity auto-spectra of an extended version of the time series in Figure 4. The auto-spectra reveal that most of the variance in water surface elevation and cross-shore velocity is at infragravity frequencies; 87.7% and 88.4% respectively. The alongshore velocity auto-spectrum, however, shows less variance over a broader range of frequencies.

An extension of spectral analysis that is useful in the study of infragravity waves is cross-spectral analysis. Cross-spectral analysis is used to determine the level of covariance and phase lag between two time series (Jenkins and Watts, 1968). If the cross-shore structure of infragravity waves is standing, cross-spectral analysis will reveal a phase difference of 90° ($\pi/2$) between the water surface elevation and cross-shore velocity (Suhayda, 1974). Cross-spectral analysis has also been used to investigate the relationship between infragravity waves and wave groups in the surf zone (e.g. Masselink, 1995).

The decomposition of wave energy into incoming and outgoing components in the frequency domain can be achieved in two ways; (1) using pressure and cross-shore velocity measurements from co-located

sensors, or (2) using pressure measurements only from a cross-shore array of pressure transducers. The method of Sheremet *et al.* (2002) is an example of the first technique where the incoming and outgoing energy E at each discrete frequency is calculated as:

$$E^{\pm}(f) = \frac{1}{4} \left[S_p(f) + \frac{h}{g} S_u(f) \pm \left(2 \sqrt{\frac{h}{g}} \right) C_{pu}(f) \right]$$

where S_p and S_u are the pressure and cross-shore velocity auto-spectra respectively and C_{pu} is the $p-u$ cospectrum. The second technique uses calculations of wave celerity and phase lag as waves travel through an array of normally three or more pressures transducers. A number of solutions exist that employ this method such as those of Gaillard *et al.* (1980), Battjes *et al.* (2004) and Van Dongeren *et al.* (2007). The type of

decomposition method used will largely depend on instrument availability and the overall aim of the study. De Bakker *et al.* (2014) performed a comparison between the $p-u$ method of Sheremet *et al.* (2002) and the array method of Van Dongeren *et al.* (2007) using data collected in the surf zone of a dissipative beach in the Netherlands and found relatively good agreement between the two methods.

Following the decomposition of wave energy by one of the above methods, frequency dependant (and bulk) reflection coefficients $R^2(f)$ can be calculated which are simply the ratio of seaward to shoreward propagating wave energy flux $F^{\pm}(f)$ (calculated as $E^{\pm}(f)\sqrt{gh}$).

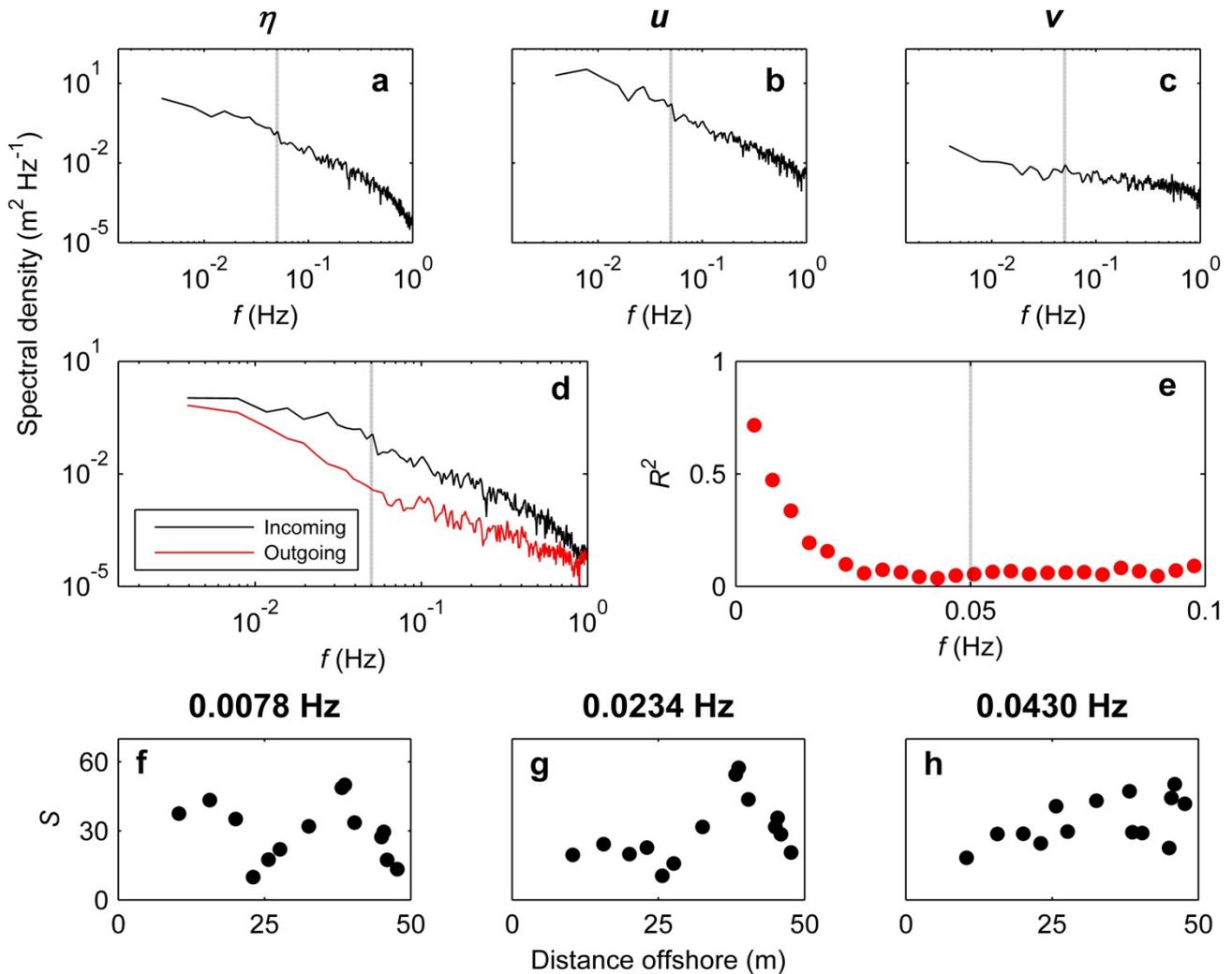


Figure 5: Auto-spectra of water surface elevation η (a), cross-shore velocity u (b) and alongshore velocity v (c). Incoming and outgoing water surface elevation auto-spectra (d) using the separation method of Sheremet *et al.* (2002) and corresponding reflection coefficients R^2 (e). Normalised spectral density S versus distance offshore for three infragravity frequencies (f-h). Vertical dotted lines in a-e indicate the infragravity-incident frequency transition of 0.05 Hz.

Figure 5e shows the frequency dependant reflection coefficients corresponding to the incoming and outgoing spectra in Figure 5d (estimated using the method of Sheremet *et al.* (2002) outlined above). The reflection coefficients show that only frequencies at the lower end of the infragravity band reflect a significant amount of energy; R^2 is less than 0.1 for all frequencies higher than 0.0273 Hz indicative of > 90% energy dissipation.

To gain a detailed insight into the cross-shore structure and transformation of infragravity (and incident) waves, the techniques described above can be applied to data collected at a number of cross-shore locations. This can be achieved through the deployment of an instrument array, or deploying a single instrument rig and using the tide as a surrogate for changing cross-shore position. The latter option requires fewer instruments, but is limited to sites with a large tidal range, linear beach profile and no change in forcing conditions during the study. The cross-shore structure of energy at three discrete frequencies in the infragravity band is demonstrated in Figure 5f-h by plotting the spectral density at these frequencies versus distance offshore. In this example the tide was used as a proxy for changing cross-shore position, hence the spectral density is normalised by the offshore wave height. It can be seen that at 0.0078 Hz there is a clear (anti)nodal structure symbolic of standing waves with antinodes at 15 m and 40 m and a node at 24 m. This is partially lost at 0.0234 Hz and completely absent at 0.0430 Hz which displays a progressive wave pattern. This agrees well with the reflection coefficients in Figure 5e.

Concluding Remarks

The surf zone is highly energetic and can be a challenging environment in which to collect wave and current measurements. The data, however, are invaluable and there have been many advances in instrument technology over the last few decades, which have helped to further our knowledge of surf zone hydrodynamics. Pressure transducers are undeniably the most commonly used and arguably the most valuable device for measuring waves. There has been a shift towards the use of acoustic sensors for measuring surf zone currents, yet EMCMs

remain a valuable tool, especially during turbulent conditions.

Detailed planning, including a thorough understanding of the various sensors and analysis procedures, is crucial to ensure the collection of a high quality, useful dataset. There are several important considerations that need to be included such as the local wave climate, sampling strategy, data retrieval and analysis techniques, and the overall aim of the research. Researchers are advised to review the relevant literature before undertaking a field experiment in order to achieve the optimal set-up for their project.

Acknowledgements

The comments and suggestions of two anonymous reviewers helped to improve the clarity of this article. The author also thanks Dr Mark Davidson and Prof. Paul Russell for their constructive feedback on an early draft. Fieldwork assistance at Perranporth was provided by Dr Mark Davidson, Prof. Paul Russell, Dr Tim Scott and Mr Luis Melo de Almeida.

References

- Aagaard T, Masselink G. 1999. The Surf Zone. In *Handbook of Beach and Shoreface Morphodynamics*, Short AD (ed.). Wiley: Chichester; 72-118.
- Aagaard T, Black KP, Greenwood B. 2002. Cross-shore suspended sediment transport in the surf zone: a field-based parameterization. *Marine Geology* **185**: 283-302.
- Arnaud G, Mory M, Abadie S, Cassen M. 2009. Use of a resistive rods network to monitor bathymetric evolution in the surf/swash zone. *Journal of Coastal Research* **SI 56**: 1781-1785.
- Austin M, Scott T, Brown J, Brown J, MacMahan J, Masselink G, Russell P. 2010. Temporal observations of rip current circulation on a macro-tidal beach. *Continental Shelf Research* **30**: 1149-1165.
- Austin MJ, Masselink G, Scott TM, Russell PE. 2014. Water-level controls on macro-tidal rip currents. *Continental Shelf Research* **75**: 28-40.

- Battjes JA. 1974. Surf similarity. *Proceedings of the 14th Coastal Engineering Conference*, ASCE, 466-480.
- Battjes JA, Bakkenes HJ, Janssen TT, Van Dongeren AR. 2004. Shoaling of subharmonic gravity waves. *Journal of Geophysical Research* **109**: C02009.
- Blenkinsopp CE, Turner IL, Allis MJ, Peirson WL, Garden LE. 2012. Application of LIDAR technology for measurement of time-varying free-surface profiles in a laboratory wave flume. *Coastal Engineering* **68**: 1-5.
- Brander RW. 1999. Field observations on the morphodynamic evolution of a low-energy rip current system. *Marine Geology* **157**: 199-217.
- Brander RW, Short AD. 2001. Flow Kinematics of Low-energy Rip Current Systems. *Journal of Coastal Research* **17**: 468-481.
- Catalan PA, Haller MC, Plant WJ. 2014. Microwave backscattering from surf zone waves. *Journal of Geophysical Research* **119**: 3098-3120.
- Davidson-Arnott RGD, McDonald RA. 1989. Nearshore water motion and mean flows in a multiple parallel bar system. *Marine Geology* **86**: 321-338.
- Davidson-Arnott R. 2010. Introduction to Coastal Processes and Geomorphology. Cambridge University Press: New York.
- De Bakker ATM, Tissier MFS, Ruessink BG. 2014. Shoreline dissipation of infragravity waves. *Continental Shelf Research* **72**: 73-82.
- De Vries S, Hill DF, De Schipper MA, Stive MJF. 2011. Remote sensing of surf zone waves using stereo imaging. *Coastal Engineering* **58**: 239-250.
- Elgar S, Raubenheimer B, Guza RT. 2001. Current Meter Performance in the Surf Zone. *Journal of Atmospheric and Oceanic Technology* **18**: 1735-1746.
- Elgar S, Raubenheimer B, Guza RT. 2005. Quality control of acoustic Doppler velocimeter data in the surfzone. *Measurement Science and Technology* **16**: 1889-1893.
- Feddersen F. 2010. Quality Controlling Surf Zone Acoustic Doppler Velocimeter Observations to Estimate the Turbulent Dissipation Rate. *Journal of Atmospheric and Oceanic Technology* **27**: 2039-2055.
- Gaillard P, Gauthier M, Holly F. 1980. Method of analysis of random wave experiments with reflecting coastal structures. *Proceedings of the 17th International Conference on Coastal Engineering*, ASCE, 204-220.
- Galvin CJ. 1968. Breaker type classification on three laboratory beaches. *Journal of Geophysical Research* **73**: 3651-3659.
- Green MO. 1999. Test of sediment initial-motion theories using irregular-wave field data. *Sedimentology* **46**: 427-441.
- Guedes RMC, Bryan KR, Coco G. 2013. Observations of wave energy fluxes and swash motions on a low-sloping, dissipative beach. *Journal of Geophysical Research* **118**: 3651-3669.
- Guza RT, Thornton EB, Holman RA. 1984. Swash on steep and shallow beaches. *Proceedings of the 19th International Conference on Coastal Engineering*, ASCE, 708-723.
- Guza RT, Thornton EB. 1985. Observations of Surf Beat. *Journal of Geophysical Research* **90**: 3161-3172.
- Guza RT, Clifton MC, Rezvani F. 1988. Field Intercomparisons of Electromagnetic Current Meters. *Journal of Geophysical Research* **93**: 9302-9314.
- Haller MC, Honegger D, Catalan PA. 2014. Rip Current Observations via Marine Radar. *Journal of Waterway, Port, Coastal and Ocean Engineering* **140**: 115-124.
- Hegge BJ, Masselink G. 1996. Spectral Analysis of Geomorphic Time series: Auto-Spectrum. *Earth Surface Processes and Landforms* **21**: 1021-1040.
- Holman RA, Symonds G, Thornton EB, Ranasinghe R. 2006. Rip spacing and persistence on an embayed beach. *Journal of Geophysical Research* **111**: C01006.
- Holman RA, Stanley L. 2007. The history and technical capabilities of Argus. *Coastal Engineering* **54**: 477-491.
- Holman RA, Haller MC. 2013. Remote Sensing of the Nearshore. *Annual Review of Marine Science* **5**: 95-113.
- Huntley DA, Hendry MD, Haines J, Greenidge B. 1988. Waves and Rip Currents

- on a Caribbean Pocket Beach, Jamaica. *Journal of Coastal Research* **4**: 69-79.
- Jenkins GM, Watts DG. 1968. Spectral Analysis and its Applications. Holden-Day: San Francisco.
- Johnson D, Pattiaratchi C. 2004. Application, modelling and validation of surfzone drifters. *Coastal Engineering* **51**: 455-471.
- Komar PD. 1998. Beach Processes and Sedimentation: Second Edition. Prentice Hall: New Jersey.
- Longuet-Higgins MS. 1952. On the statistical distribution of the height of sea waves. *Journal of Marine Research* **11**: 245-266.
- MacMahan J, Brown J, Thornton E. 2009. Low-Cost Handheld Global Positioning System for Measuring Surf-Zone Currents. *Journal of Coastal Research* **25**: 744-754.
- MacVicar BJ, Beaulieu E, Champagne V, Roy AG. 2007. Measuring water velocity in highly turbulent flows: field tests of an electromagnetic current meter (ECM) and an acoustic Doppler velocimeter (ADV). *Earth Surface Processes and Landforms* **32**: 1412-1432.
- Masselink G. 1995. Group bound long waves as a source of infragravity energy in the surf zone. *Continental Shelf Research* **15**: 1525-1547.
- Masselink G, Puleo JA. 2006. Swash-zone morphodynamics. *Continental Shelf Research* **26**: 661-680.
- Masselink G, Hughes MG, Knight J. 2011. Coastal Processes and Geomorphology: Second Edition. Hodder Education: London.
- McCarroll RJ, Brander RW, Turner IL, Power HE, Mortlock TR. 2014. Lagrangian observations of circulation on an embayed beach with headland rip currents. *Marine Geology* **355**: 173-188.
- Morang A, Larson R, Gorman L. 1997. Monitoring the Coastal Environment; Part 1: Waves and Currents. *Journal of Coastal Research* **13**: 111-133.
- Puleo JA, Faries JWC, Davidson M, Hicks B. 2010. A Conductivity Sensor for Nearbed Sediment Concentration Profiling. *Journal of Oceanic and Atmospheric Technology* **27**: 397-408.
- Puleo JA, Lanckriet T, Wang P. 2012. Near bed cross-shore velocity profiles, bed shear stress and friction on the foreshore of a microtidal beach. *Coastal Engineering* **68**: 6-16.
- Puleo JA, Lanckriet T, Blenkinsopp C. 2014. Bed level fluctuations in the inner surf and swash zone of a dissipative beach. *Marine Geology* **349**: 99-112.
- Ridd PV. 1992. A sediment level sensor for erosion and siltation detection. *Estuarine, Coastal and Shelf Science* **35**: 353-362.
- Rodriguez A, Sanchez-Arcilla A, Redondo JM, Mosso C. 1999. Macroturbulence measurements with electromagnetic and ultrasonic sensors: a comparison under high-turbulent flows. *Experiments in Fluids* **27**: 31-42.
- Ruessink BG. 1999. Data report 2.5D experiment Egmond aan Zee. Department of Physical Geography, University of Utrecht, The Netherlands.
- Ruessink BG, Bell PS, Van Enckevort IMJ, Aarninkhor SGJ. 2002. Nearshore bar crest location quantified from time-averaged X-band radar images. *Coastal Engineering* **45**: 19-32.
- Russell PE. 1993. Mechanisms for beach erosion during storms. *Continental Shelf Research* **13**: 1243-1265.
- Saulter AN, Russell PE, Gallagher EL, Miles, JR. 2003. Observations of bed level change in a saturated surf zone. *Journal of Geophysical Research* **108**: 3112.
- Schmidt WE, Woodward BT, Millikan KS, Guza RT, Raubenheimer B, Elgar S. 2003. A GPS-Tracked Surf Zone Drifter. *Journal of Atmospheric and Oceanic Technology* **20**: 1069-1075.
- Scott TM, Russell PE, Masselink G, Wooler A, Short AD. 2007. Beach Rescue Statistics and their Relation to Nearshore Morphology and Hazard: A Case Study for Southwest England. *Journal of Coastal Research* **SI 50**: 1-6.
- Scott TM, Russell PE, Masselink G, Wooler A. 2008. High volume sediment transport and its implications for recreational beach risk. *Proceedings of the 31st International Conference on Coastal Engineering*, ASCE, 4250-4262.
- Senechal N, Coco G, Bryan KR, Holman RA. 2011a. Wave runup during extreme storm

conditions. *Journal of Geophysical Research* **116**: C07032.

Senechal N, Abadie S, Gallagher E, MacMahan J, Masselink G, Michallet H, Reniers A, Ruessink G, Russell P, Sous D, Turner I, Arduin F, Bonneton P, Bujan S, Capo S, Certain R, Pedreros R, Garlan T. 2011b. The ECORS-Truc Vert'08 nearshore field experiment: presentation of a three-dimensional morphologic system in a macro-tidal environment during consecutive extreme storm conditions. *Ocean Dynamics* **61**: 2073-2098.

Sheremet A, Guza RT, Elgar S, Herbers THC. 2002. Observations of nearshore infragravity waves: Seaward and shoreward propagating components. *Journal of Geophysical Research* **107**: C8.

Short AD, Brander RW. 1999. Regional Variations in Rip Density. *Journal of Coastal Research* **15**: 813-822.

Suhayda JN. 1974. Standing waves on beaches. *Journal of Geophysical Research* **79**: 3065-3071.

Thornton EB, Guza RT. 1982. Energy saturation and phase speeds measured on a natural beach. *Journal of Geophysical Research* **87**: 9499-9508.

Thornton EB, Guza RT. 1983. Transformation of Wave Height Distribution. *Journal of Geophysical Research* **88**: 5925-5938.

Van Dongeren A, Battjes J, Janssen T, Van Noorloos J, Steenhauer K, Steenbergen G, Reniers A. 2007. Shoaling and shoreline dissipation of low-frequency waves. *Journal of Geophysical Research* **112**: C02011.

Woodroffe CD. 2002. Coasts: form, process and evolution. Cambridge University Press: Cambridge.

Methods for field measurement and remote sensing of the swash zone

Sebastian J. Pitman¹

¹Ocean and Earth Sciences, National Oceanography Centre, University of Southampton (sjp1e13@soton.ac.uk)



ABSTRACT: Swash action is the dominant process responsible for the cross-shore exchange of sediment between the subaerial and subaqueous zones, with a significant part of the littoral drift also taking place as a result of swash motions. The swash zone is the area of the beach between the inner surfzone and backbeach that is intermittently submerged and exposed by the processes of wave uprush and backwash. Given the dominant role that swash plays in the morphological evolution of a beach, it is important to understand and quantify the main processes. The extent of swash (horizontally and vertically), current velocities and suspended sediment concentrations are all parameters of interest in the study of swash processes. *In situ* methods of measurements in this energetic zone were instrumental in developing early understanding of swash processes, however, the field has experienced a shift towards remote sensing methods. This article outlines the emergence of high precision technologies such as video imaging and LIDAR (light detection and ranging) for the study of swash processes. Furthermore, the applicability of these methods to large-scale datasets for quantitative analysis is demonstrated.

KEYWORDS: run-up, morphodynamics, coastal imaging, video, LIDAR.

Introduction

The beachface is a highly spatially and temporally dynamic zone, predominantly due to swash processes such as wave run up. The foreshore (the intermittently wetted, intertidal area) is the interface between land and sea and is characterised by highly variable hydro- and morpho-dynamic processes. Understanding the evolution of the foreshore is of critical importance to coastal oceanographers, planners and engineers because energy delivered to this region drives the erosive or accretive response of the beach (Stockdon *et al.*, 2006). Swash action is the dominant process responsible for moving sediment cross-shore between the subaerial and subaqueous zones, with a significant part of the littoral drift also taking place in this zone (Masselink and Puleo, 2006). The swash zone (Figure 1) is defined as the boundary between the inner surf zone and the back beach (Ruggiero *et*

al., 2004) and its dominant responses are largely well understood. It is the most energetic zone in terms of bed sediment movement and is characterised by strong and unsteady flows as a result of run up and backwash, within which single events can cause changes of up to 43 mm in bed level (Blenkinsopp *et al.*, 2011). It is important to recognise that this swash zone is part of an integrated system comprising local groundwater dynamics, the beachface and the surf zone, with the feedback from surf to swash of critical importance when considering hydrodynamics (Masselink and Puleo, 2006). It has also recently been shown that swash zone flows exert influence not just locally (overtopping, littoral drift, *etc.*), but they also affect the dynamics of the surf zone itself (Brocchini, 2006).

Run up is described here as a set of discrete water level maxima measured on the foreshore, with respect to the still water level;

that which would occur in the absence of forcing by the incident wave field (Grant, 1948; Guza and Thornton, 1982). The two components of run up; wave swash and wave set-up, operate on very different temporal scales, as a result of the different forcing factors (Senechal *et al.*, 2011).

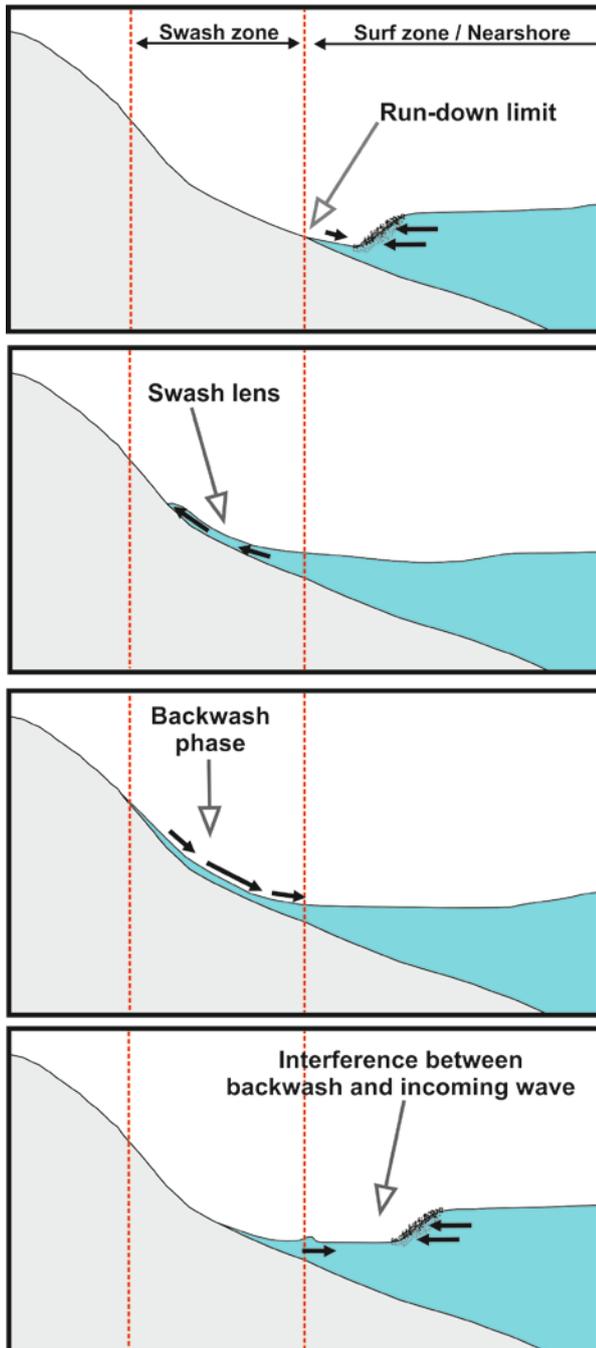


Figure 1: Representation of a typical swash cycle. Arrows represent flow direction with size indicative of relative magnitude. Red lines delineate the swash zone based on limit of backwash (a) and maximum limit of swash excursion (c). Modified from Masselink and Puleo (2006).

Swash, the time-varying, fluctuating component, operates on frequencies comparable to the incident wave field from which it stems, whereas set-up refers to the mean water level as a result of wave breaking (Guza and Thornton, 1982; Nielsen and Hanslow, 1991). The wave run up height is generally normalised by the incident wave height, as they are of the same order of magnitude (Kobayashi, 1997). Set-up is relatively small when compared to swash action on a steeply sloping beach, where there is an appreciable degree of wave reflection (Battjes, 1974). This wave driven run up converts kinetic energy into potential energy as it traverses up the beachface (uprush phase), before gravity-driven flows act to return the flow down the slope of the foreshore (backwash phase). There is typically interference between subsequent waves, with the backwash of preceding waves colliding with the uprush of the next wave, meaning individual waves do not often complete a full and balanced cycle of uprush and backwash (Erikson *et al.*, 2005). The motion of fluid traversing over the beachface may be affected somewhat by infiltration and/or exfiltration from the beachface. Infiltration will remove water from the swash and generally aids with progradation of the beachface, whereas exfiltration adds water to the backwash and generally aids erosion or scarp of the beachface (Masselink and Puleo, 2006).

When determining various morpho- and hydro-dynamic properties and states of the beach, the non-dimensional Iribarren number or surf similarity parameter is commonly used (Battjes, 1974);

$$\xi = \frac{\tan\beta}{(H_o/L_o)^{1/2}} \quad (1)$$

where β is beach slope, L_o is the deepwater wavelength given by linear wave theory and H_o is the offshore wave height. This is often referred to as the dynamic beach steepness parameter (Stockdon *et al.*, 2006), accounting for the antecedent beach slope as well as the incident wave conditions. This property has proved useful in empirically determining run up (Holman and Sallenger, 1985; Holman, 1986; Ruggiero *et al.*, 2004; Stockdon *et al.*, 2006) and illustrates well the dependence of run up on beach slope and wave conditions.

Miche (1951) also provided a formula useful in estimating the maximum swash amplitude;

$$\varepsilon_s = \frac{a_s \omega^2}{g \beta^2} \quad (2)$$

where a_s is the vertical amplitude of the shoreline motion, ω is the angular wave frequency, g is the gravitational component and β is the beach slope. In the Miche equation, maximum swash amplitude is assumed to be related to the limiting amplitude of a non-breaking standard wave on a planar beach slope (Baldock and Holmes, 1997). Miche postulated that an increase in incident wave height would eventually lead to wave breaking, and therefore the amplitude would be saturated as a further increase in wave height simply only increases the amplitude of the progressive component, which would then dissipate through wave breaking with no shoreline amplitude (Ruggiero *et al.*, 2004). This hypothesis therefore proposes that swash heights do not increase with increasing offshore wave height.

When considering swash processes, there are three main hydrodynamic parameters of interest; 1) swash excursion, 2) swash depth, and 3) swash velocity.

Swash excursion refers to the spatial coverage of the varying water surface. This excursion up the beach is typically defined in terms of its vertical elevation, rather than the horizontal extent of run up (Holland *et al.*, 1995). This vertical excursion is usually against the Ordnance Datum (in the UK), which provides a static reference point, especially useful in macrotidal environments. This parameter is of interest because the excursion of swash delineates the area of the beach subject to bed shear stresses, and ultimately sediment transport.

The depth of swash is important because it directly influences how much sediment can be transported in the water column. Thin layers of swash will be confined to transporting sediment as bed load, whereas deeper swash lenses will be capable of transporting suspended sediment as well as transporting bed load. The shape of the swash lens (*i.e.* the swash surface profile) has been postulated as indicative of flow velocities. Early observation of the swash

lens during backwash showed that at flow reversal, the leading edge of the swash remains static until a large amount of water mass downslope has moved seaward (Emery and Gale, 1951). Hughes (1992, 1995) identified differing hydrodynamic characteristics within the backwash based upon the free surface profile of the swash lens. He was able to relate these different swash lenses to different methods of sediment transport within the backwash. Therefore, measurement of swash depth across the foreshore may provide clues as to the complicated hydrodynamics at work.

Bed shear stress is the mechanism driving initial sediment motion and is perhaps the most important hydrodynamic parameter of swash motion. It is, however, often replaced with a surrogate parameter in that of flow velocity (Hughes *et al.*, 1997). A degree of asymmetry in swash duration and velocities has often been observed both in laboratory and field investigations (Kemp, 1975; Raubenheimer *et al.*, 1995; Baldock and Holmes, 1997; Power *et al.*, 2011). Of particular interest is the vertical velocity structure (Blenkinsopp *et al.*, 2010), however, the depth averaged velocity is also a useful parameter in estimation of sediment transport.

This paper presents an overview of swash measurement through *in-situ* measurement and the growing shift towards remotely sensed measurements. Measurement of swash processes is often coincident with measurement of sediment movement within the swash zone; sediment transport is covered in Section 3.2.4 and thus not included here. The methods and data presented here, although applicable across the whole spectrum of swash environments, are predominantly collated from the work of Plymouth University (Devon, UK) on Slapton Sands; a coarse grained gravel barrier in SW England.

***In-situ* measurement**

Field measurement of the spatial extent, free surface profile and velocity profiles within the swash zone has incorporated a number of *in-situ* techniques, outlined below.

Swash excursion

Some of the first quantitative measurements of swash excursion were undertaken using resistance wires (Guza and Thornton, 1982). This involved two electronic resistance wires, elevated to a nominal distance above the bed. The wires ran cross shore and the swash action running up that profile submerged the wires, shorting out the current path, resulting in different levels of resistance (Holman and Guza, 1984).

The wires therefore recorded the most seaward location at which water depth exceeded the vertical elevation of the wires. This method typically yielded accuracies of O(cm) in terms of the horizontal (cross-shore) extent of run up. The wires were calibrated by submerging known lengths in water prior to field deployment (Raubenheimer *et al.*, 1995). They were further calibrated after the measurements, as Guza and Thornton (1982) noted resistance gains of 4.5% occurred between pre- and post-experiment calibration. The thin run up lense means this method is extremely sensitive to wire elevation, such that lower wire elevations will indicate a larger total extent of run up (Holman and Guza, 1984; Holland *et al.*, 1995). Raubenheimer and Guza (1996) found a 20 cm vertical wire separation accounted for a cross-shore difference in run up excursion of 20 m.

Swash depth

In the same way that horizontal resistance wires have been used to measure the spatial extent of swash, vertical capacitance probes have previously been used to measure swash depth (Waddell, 1976; Hughes, 1992). The probes consist of vertically suspended wires, which when submerged, register an increased capacitance to that recorded in air alone. The change in current is received as an amplitude variation, and these signals can then be calibrated against signals received in known water depths, providing an indication of swash depth in the field (Foote and Horn, 2002).

Instruments deployed to measure bed levels are typically concerned with the change in morphology on the beachface (e.g. Waddell, 1980), however, the same instruments, when set to log at rates of > 1 Hz also give good

indication of swash motion over the beachface. The sensors used in this type of instrumentation typically transmit ultrasonic sound toward a target (the beachface) and measure the time before an echo is received back, resulting in high accuracy, high precision measurements (Massa, 1999). The time lag before the echo correlates to distance and is calibrated on the medium through which the sound passes (i.e. air).

Turner *et al.* (2008) described a bed level acoustic array mounted on a scaffold frame that collected data of bed level changes at 4 Hz immediately prior to each swash event. The bed level sensors offered a response from the first non-gaseous surface (i.e. water or sediment). Therefore, the method was appropriate and applicable to measurement of the swash lens. Poate *et al.* (2013) demonstrated the ability of the scaffold sensor array to conduct continuous measurement of swash action through non-daylight hours in energetic conditions.

Pressure transducers (PTs), more commonly used in the surf zone or deep ocean, can also be used to measure swash depth. They measure the pressure applied by the water above and convert this force into an equivalent water depth. PTs give discreet observations, and therefore interpolation is required to make inferences about the swash lens. They are often deployed in conjunction with other sensors, such as run up wires for a more continuous dataset.

Masselink and Russell (2006) deployed 3 rigs, each with 3 PTs attached, in a field experiment investigating flow velocities and sediment transport in the swash zone. PTs are ideally deployed flush with the beachface to give an idea of the total amount of water above the bed. However, this is problematic because saturated sediment washes over and settles upon the PTs, skewing the results. Therefore, Masselink and Russell (2006) set any recorded values of ≤ 0.02 m to zero, and disregarded data in this range.

Swash velocities

Bidirectional current meters in the form of ducted impellers have been used to ascertain swash velocities (Sonu, 1972; Sonu *et al.*, 1974; Teleki *et al.*, 1976). The rotation of the three-bladed impeller in the duct gives output

in the form of pulse density and polarity, which is translated to current speed and direction. Sonu (1972) estimated the threshold sensitivity of this method to be 5 cm/sec, with very little signal degradation when the impellor was orientated less than 30 degrees from the mean flow direction. The purpose of the duct on these meters is to protect them from derogatory effects of flows at large incidence angles (Smith, 1978). The impellor current meters are limited somewhat in that they can only ascertain currents in one direction at a time; however, their application to the swash zone is proportionate as generally the swash flow reverses in a depth-uniform manner, unlike the surf zone. Furthermore, the current meters are only able to ascertain meaningful velocity measurements up to a relatively small maximum velocity (Holland *et al.*, 2001).



Figure 2: An example of a mobile Argus station deployed by Plymouth University at Slapton Sands for a short field experiment. Two cameras are mounted on a scaffold rig. Checkerboards are placed on the beach, within the cameras field of view to act as ground control points (Photo: S. Pitman).

Electromagnetic current meters (EMCMs) have also been used widely to measure swash zone velocities, although they are more commonly used in the surf zone. The

EMCM works by inducing the creation of voltage by movement of a conductor (water) through a magnetic field. Here, the magnetic field is created by a coil inside the head of the instrument (Butt, 1999). The sensors are designed to be used fully submerged, however, with calibration, they can be adapted to suit the continuous emergence and submergence occurring in the swash zone. These meters often have small discus heads, and can be considered to give errors in the region of 8% of actual swash velocities (Butt *et al.*, 2001). These sensors need to be mounted at heights of few cm to avoid interference from the bed (Butt and Russell, 1999; Butt *et al.*, 2004).

Acoustic Doppler Velocimeters (ADV) have been used successfully to quantify velocities in the swash zone. These small and robust sensors are effective at measuring oscillatory flows such as those in the swash, typically capable of measuring flow velocities between 0 and 2.5 m/s (Kraus *et al.*, 1994). The sensors work by emitting short acoustic pulses into the water and receive a signal back after the pulse is scattered back by reflectors in the water (Elgar *et al.*, 2001). The phase shift in these reflectors between several successive returns allows for information on cross- and along-shore velocities, as well as any vertical velocity components to be collected (Lhermitte and Serafin, 1984). The ADV returns can be corrupted by bubbles and sediment in the water column created by breaking waves, but these signals are often distinguishable using a ratio between signal and noise, meaning they can be discarded where appropriate (Elgar *et al.*, 2005).

Remote sensing

Predictive formulas for run up are critical for coastal planners, engineers and researchers, because they provide estimation based on relatively easy to measure variables such as the offshore wave conditions and beach slope. However, the *in situ* measurement of swash processes is inherently difficult and complex (Blenkinsopp *et al.*, 2011), proving challenging for even the most robust and advanced hydrodynamic equipment (Masselink and Puleo, 2006). This problem seems counter-intuitive as the swash zone is readily accessible when compared to the surf zone or offshore regions; however, the

instruments used in the swash zone are often delicate and are exposed to energetic wave breaking and sediment transport regimes. Many coastal processes, especially in the energetic swash zone, are poorly understood because of the difficulty in collecting continuous, long-term and large scale field measurements, especially with high spatial and temporal resolution (Guedes *et al.*, 2011, Holman and Stanley, 2007). Remote sensing systems capable of monitoring coastal processes, such as the Argus video imaging system (Holman and Stanley, 2007), have experienced a period of significant interest and development over the past 30 years (Guedes *et al.*, 2011; Holman and Stanley, 2007). Progress in this area is largely driven by the aforementioned difficulties in obtaining *in situ* measurement. Coastal imaging systems can be permanent fixtures or can be deployed temporarily to compliment a field experiment (Figure 2). The more commonly employed remote sensing techniques are described below.

Video image analysis

Time lapse photography has been used in the study of the nearshore for over 40 years (Sonu, 1972; Sasaki *et al.*, 1976). Quantitative data is obtained by interrogating the image for optical signatures that are either directly or indirectly created by nearshore processes, such as the concentration of breaking waves over a submerged bar showing up as a high intensity band in the image (Kingston *et al.*, 2000; Lippmann and Holman, 1989, 1990; Plant and Holman, 1998). The first to directly apply the technique to swash motion were Holman and Guza (1984). They found that manual digitisation of swash events from photographs, although subjective, gave results comparable to those obtained by resistance wires. This method involved frame-by-frame digitisation of the run-up edge in images, as it moves through the uprush and backwash phases, from cameras positioned to look alongshore. The position of this leading edge was converted to a known cross-shore location to create a time series of run up. The authors report times of 30 minutes to digitise a 2048 point dataset. Holman and Bowen (1984) investigated the subjectivity of this method to measure swash height by making replicate digitisations of the same film using different operators. This resulted in errors of between 15 and 20 %.

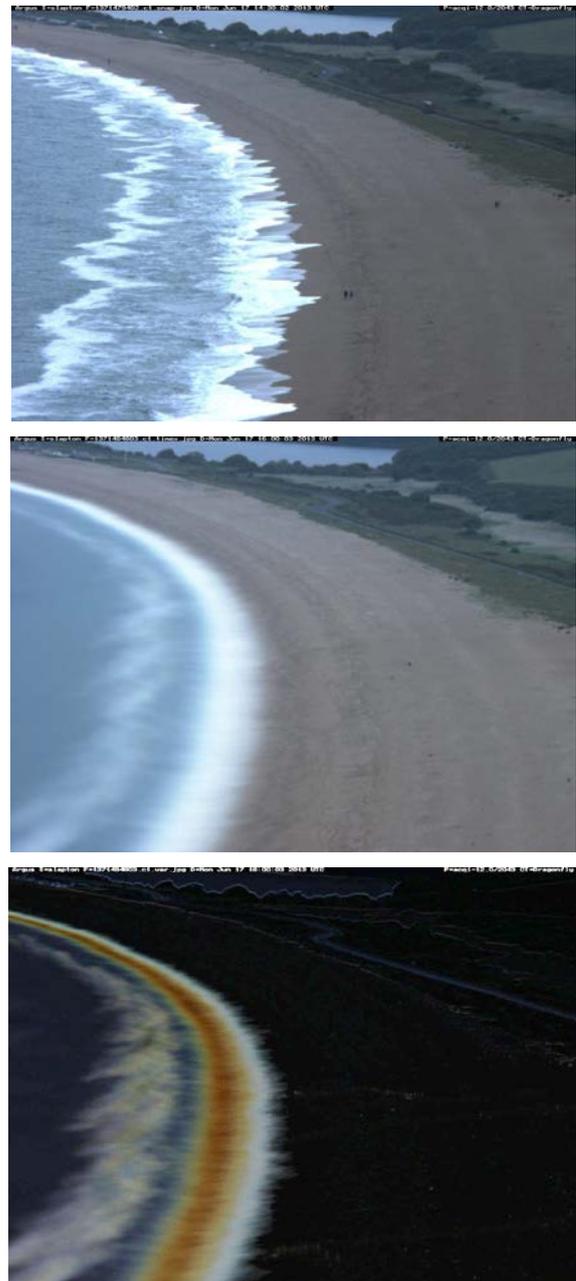


Figure 3: Corresponding examples of (a) snap, (b) time-exposure and (c) time-variance images from an Argus camera system at Slapton Sands, Devon, UK. Images courtesy of Plymouth University.

Aagard and Holm (1989) developed the use of video images further by pre-defining a transect line in the image, and stripping out only the pixels along this transect line in images sampled at 8 Hz. These transects were then used to create a visual timestack which could be digitised and ultimately allowed the backwash phase to be more effectively tracked. An example is given in Figure 4, where 180 s of run up has been recorded at 2 Hz, before being manually digitised. The timestack approach allows

researchers to clearly see the leading edge of the swash lens as it propagates across the beach, but it is sometimes harder to pick out the limit of run down.

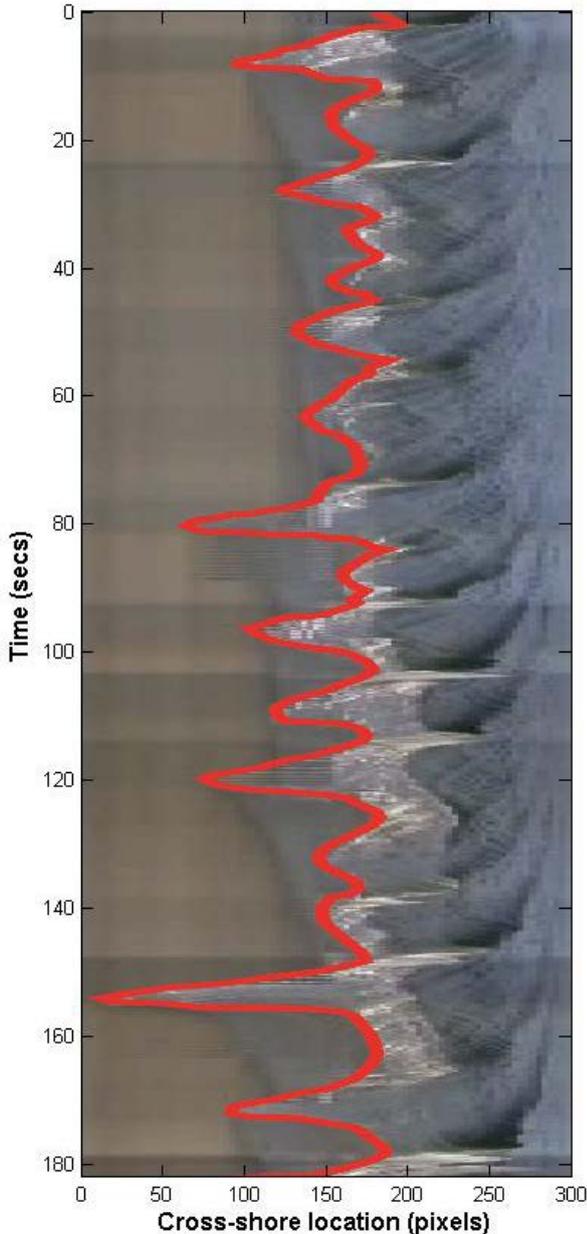


Figure 4: An example of a digitised timestack of run up over 180 s.

Video imaging has also been used as an approach to quantify swash velocities. Particle image velocimetry (PIV) is an approach that uses images sampled at short intervals to track individual particles or clusters of particles, translating the shift in position into a velocity measurement (Adrian, 1991; Grant 1997). Typically, the flow would be artificially 'seeded' with particles (Grant, 1997), however, PIV has been applied to the swash zone where the technique instead tracks the propagation of image 'texture' in the form of naturally occurring foam patterns

(Holland *et al.*, 2001). A synthetic grid is overlaid onto the area of interest, with the nodes representing the start point of estimated velocity vectors. The technique's output is a trade-off between the size of the field of view, computational time and acceptable error. The resulting data from PIV requires some degree of post processing to remove spurious vectors created by incorrect correlation. This occurs when the automation believes two different particles in subsequent images to be the same, thus creating a vector and velocity measurement between them. Holland *et al.* (2001) successfully adapted PIV to work in the field, applying it to a 25 x 40 m region of the swash zone, with grid points spaced at 0.8 m in each direction. They concluded that the method was effective at measuring flow speeds in excess of 4 m/s and could accurately represent flow structures with large spatial variation.

Meteorological factors influence the ability to remotely sense beach processes. Image quality is highly susceptible to fog, high and low light conditions, rain and temperature variations; all potentially creating images of inadequate quality for quantitative analysis. There are cameras capable of dealing with such adverse conditions, but the expense involved in such systems is generally prohibitive (Holman and Stanley, 2007). Such cameras are primarily based on infra-red imaging and thus can work in fog and low-light conditions.

Extreme sunshine, or sunlight at low incidence angles can create too much reflection from both the water and the beachface. This makes scientific interrogation of pixel intensities in the image challenging, as the high reflectivity obscures the processes of interest. This is demonstrated by Figure 5, where a transect line has been taken across a concurrent timex and variance image. High reflectivity of the beach in Figure 5a obscures the expected intensity maxima. This is a datum postulated by Plant and Holman (1997) to be indicative of the shoreline, located in these images at 60 m cross-shore. In Figure 5a, however, the intensity maxima has been displaced onshore to a location where incident sunlight has created high intensity reflection on the beachface. It could further be postulated that an ill-positioned camera may suffer from reflection of sunlight off of the water during

sunrise or sunset, obscuring the trend expected by the shoreline intensity maxima model. This effect appears to be mitigated somewhat by the use of time-variance images as opposed to time-exposure images. This approach has been employed in Figure 5b, where the subsequent pixel intensities (Figure 5c) show much less noise around the maximum intensity values, with a pronounced double peak at 60 m, coincident with the shoreline.

Image collection is also influenced by temperature variations over the diurnal cycle, especially when cameras are mounted on high scaffold frames. Solar heating creates daily shifts in the tilt of the camera by ~ 3 pixels (Holman and Stanley, 2007). At the furthest extent of the field of view, a shift of 3 pixels in the y (*along-shore*) direction could equate to a real-world alongshore shift of $O(10\text{ m})$, potentially large relative to scales of swash motion.

Detection and ranging

An emerging method for remotely sensing swash and the near-shore uses light- and/or radio- detection and ranging, or LIDAR and RADAR, respectively. These systems work by remotely sensing the free-surface elevation, normally of the surf-zone (Bell, 1999; Haller and Lyzenga, 2003), but also that of the swash-zone (Blenkinsopp *et al.*, 2010). For the purposes of this discussion, RADAR and LIDAR are grouped together as the mechanism and outputs are largely the same, with the obvious exception that one uses light (laser) and one uses sound (radio waves). Both systems work by emitting a pulse and measuring the round-trip time and known directionality of that pulse in relation to a sensor (Feagin *et al.*, 2014). Generally, the return of such pulses from the water surface would not be sufficient for quantitative analysis unless the incidence angle was

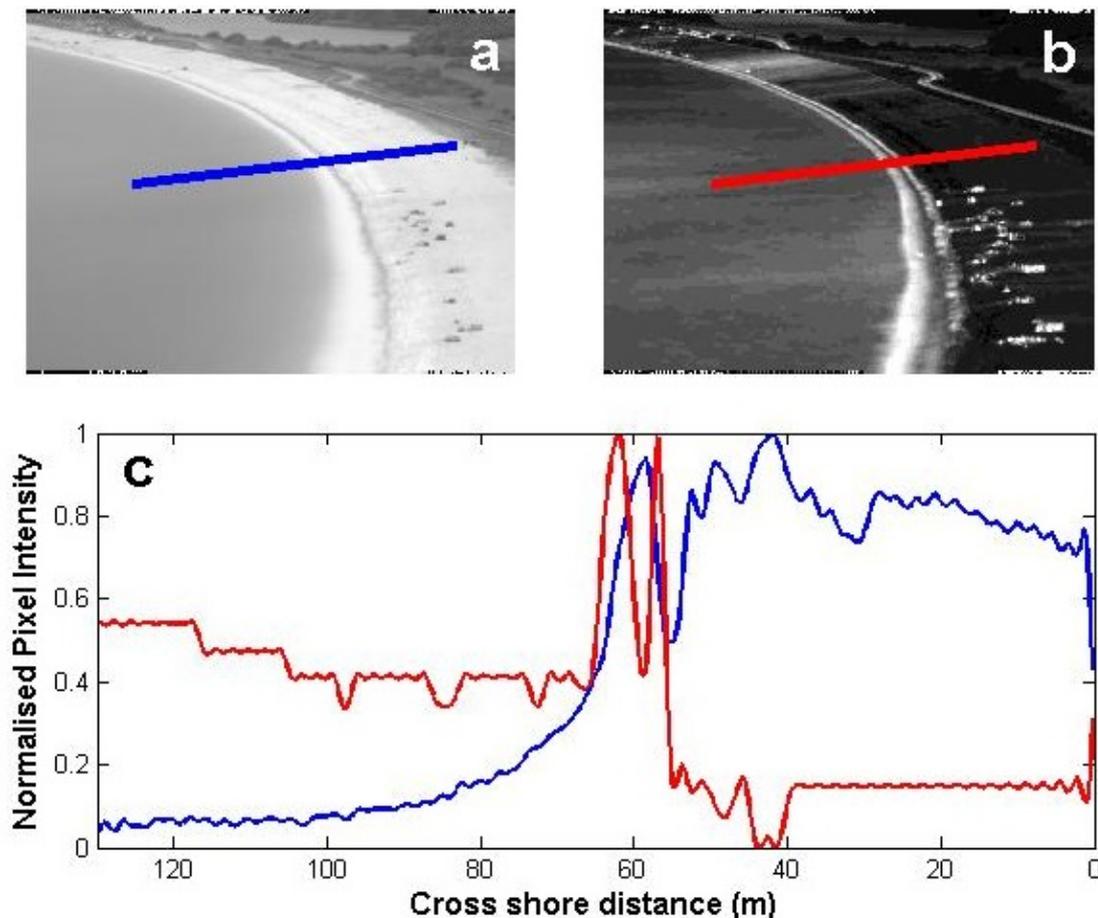


Figure 5: (a) A timex image and (b) corresponding variance image of Slapton Sands, Devon, UK, taken in bright sunshine on 30 Jul 2009. Cross-shore transect lines have been superimposed onto the images. Pixel intensities along this transect were measured, normalised and are presented in (c) with blue and red lines corresponding to the timex and variance images respectively. almost perpendicular. However, the aerated and turbulent nature of the swash means the free-surface roughness is conducive to ample return signals being created (Blenkinsopp *et*

al., 2010). It is sometimes useful to apply a time-averaging approach, much in the same way as described for video images. This reduces the scatter from ripples on the ocean surface and instead gives a solid signal back from breaking waves. This method was successfully used by Ruessink *et al.* (2002) to infer sandbar morphology from patterns of persistent wave breaking. A time-variance approach may be useful for detecting swash action on the beachface; the intermittent covering of the beachface by swash could be assumed to create a high variance feedback signal in either of the detection and ranging methods, if sampled at high enough resolution.

Unlike the video imaging approach, detection and ranging methodologies are not limited by daylight and are limited to a much lesser extent by inclement weather (McNinch, 2007). Detection and ranging systems are advantageous in the study of swash processes as they enable rapid data collection, with high spatial and temporal resolution. They are extremely useful for deployment during field studies, but are unable to be left continuously recording in the same way as video imagery. Laser scanning, in particular, is experiencing great interest from the coastal community and is rapidly becoming the tool of choice for surveying (Vousdoukas *et al.*, 2014), yet the possibilities for high resolution water surface measurement is still largely unexplored.

Application of methodology

The availability of remote sensing methods to measure swash parameters means that vast datasets, over a wide range of conditions, can be compiled and subsequently investigated. Figure 6 demonstrates the ability of video imaging systems to collect quantitative data of good spatial and temporal resolution, without any need to deploy to the field, once validation has been completed.

Figure 6 shows swash action during a storm on Slapton Sands in April 2008. Swash heights have been extracted from timex images sampled twice hourly, during every hour of daylight in an 85 hour period,

spanning the onset of the storm to its decay. Swash heights on the beach were controlled primarily by wave height (Figure 6a), with maximum daytime swash heights of 5.9 m recorded. The pre- and post-event beach profiles (Figure 6b) show that erosion up to a height of 5.8 m was observed, which corresponds with the maximum height of the remotely sensed swash estimates (Figure 6c). Furthermore, the spatial response of the beach can be estimated by investigating other transect lines in the image (Figure 6d), with results showing that maximum swash heights correlate well to maximum heights of morphological change. In Figure 6d, the bars represent measured maximum elevations of profile change, with the triangles representing the video-derived maximum swash heights. The correlation between the two shows that maximum swash height is a good indicator of effective run-up (i.e. that which has an effect on the beach profile).

This method is applicable to any of the useable hourly images at numerous sites worldwide, provided adequate antecedent morphological data is available. However, often useable images are only collected during calm conditions (i.e. not hindered by weather). The exception to this is when looking at large, flat, dissipative beaches where swash motion under calm conditions is often indistinguishable from the beach face. The effect of higher energy conditions is to create a more energetic swash, ultimately creating a better visible signature in the images.

The combination of multiple methods creates an extremely powerful, accurate tool for assessing small-scale processes such as swash, the measurement of which requires mm scale accuracies. Vousdoukas *et al.* (2014) have recently combined the use of laser scanning and video techniques to reduce root mean square errors in swash elevation on the beachface by an order of magnitude from $O(10\text{ cm})$ to $O(\text{cm})$. This process involved various forms of sensor calibration and data processing, but ultimately yields a robust method for investigating swash processes.

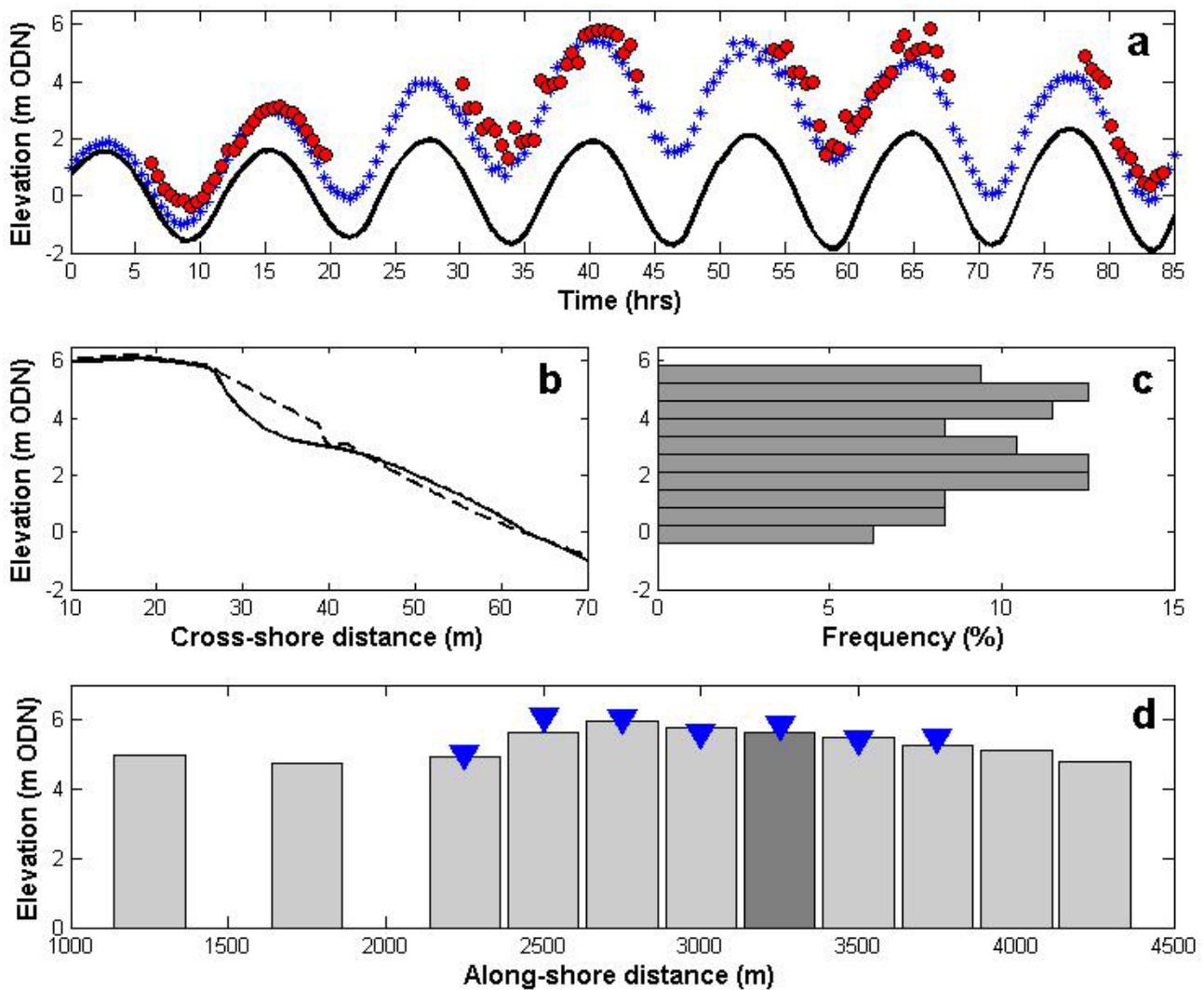


Figure 6: Example of swash extraction from a storm event between 16 and 19 April 2008 at Slapton Sands, Devon, UK. (a) The tidal signal (solid black line) and significant wave height (blue stars) are presented alongside run up derived from video images (red circles) for an 85 hour period. The pre- and post-event beach profile are presented as dashed and solid lines respectively (b) for a position approximately 3200 m alongshore. A histogram of run up heights is presented (c) which shows good correspondence between the maximum run up and the maximum elevation of profile change (effective run up) observed in the post-event profile (b). Run up at various alongshore locations has been presented as bars (d), with the darker bar corresponding to the individual profile presented above. Blue triangles represent video-derived maximum run up at locations that appeared in the field of view of the camera.

Concluding remarks

There is a progressive shift towards remote sensing methods in the study of nearshore processes, however, with the intricate nature of processes such as swash still not fully understood, the collection of *in situ* field data is far from being superfluous. The deployment of instruments into the surf, although sometimes logistically challenging, provides key information on parameters such as the vertical velocity structure of the swash zone. Resistance wires provide a simple and

cheap method for measuring both swash depth and also swash excursion. Their simple construction means that even in the event of energetic events damaging them, the fiscal cost to the researcher is minimal. They are able to provide O(cm) accuracy, depending on how they are deployed, in terms of their elevation above the bed. Recently, the advancing state of the art has welcomed a shift to more technically advanced *in situ* methods, such as PTs, bed level sensors and ADVs. The implementation of these instruments has reduced errors in

measurement down to mm scales, however, the cost and fragility of the equipment has increased exponentially. The sensors now deployed on these units can not only be fouled by seaweed and detritus, but they are susceptible to damage in the energetic swash. Their suitability to be deployed on calm, dissipative sand beaches is clear, however, the same is not always true for the swash zone of a gravel beach. Indeed, some recent studies (e.g. Poate *et al.*, 2013) have said that much future work on gravel beach swash should be confined to remote sensing methods alone.

The growing availability of remote sensing methods has acted to advance both the understanding of the swash zone and also the capabilities of researchers. Coastal imaging systems can remain in place unhindered for years, with minimal maintenance costs. Their continual logging means that when an event of interest takes place, the data for that event can be interrogated and analysed, without the need to pre-plan a large scale field deployment. Although they are currently limited to day light monitoring and somewhat calm conditions, the reducing cost of technology will increase the availability of systems capable of sampling in the infra-red spectrum, ultimately meaning the imaging system can capture events of interest under any prevailing conditions.

It is clear that advances in technology have been instrumental to studies of the coastal zone. The application of remote sensing methods can create vast databases of swash processes, which can ultimately be compared to incident conditions to help parameterise swash. These data can be used to validate existing equations such as that of Stockdon *et al.* (2006), to ultimately enable coastal planners to utilise the most appropriate equation in their estimation of extreme swash events.

In lieu of new technologies becoming available, the combination of existing methodologies, especially with regard to remote sensing, can vastly help to reduce errors in datasets. The continued development of methods such as that of video imaging, and terrestrial detection and ranging, will be instrumental in bettering our

understanding of coastal processes over both spatial and temporal scales.

Acknowledgements

This article's scope, depth and structure was greatly improved following the comments from two anonymous reviewers. The author acknowledges the studentship offered by the University of Southampton and the time allotted within this to complete this article. Thank to Plymouth University for sharing ARGUS imagery of Slapton Sands that was used for quantitative analysis and swash extraction. This article was greatly improved after the constructive comments of Dr. S.L. Gallop.

References

- Aagard T, Holm J. 1989. Digitization of wave run-up using video records. *Journal of Coastal Research* **5**: 547-551.
- Adrian RJ. 1991. Particle-imaging techniques for experimental fluid mechanics. *Annual Review of Fluid Mechanics* **23**: 261-304.
- Baldock TE, Holmes P. 1997. Swash hydrodynamics on a steep beach in: Thornton EB (ed) *Coastal Dynamics '97*. American Society of Civil Engineers, New York.
- Battjes JA. 1974. Surf similarity. *Proceedings of the 14th Coastal Engineering Conference*. Copenhagen, Denmark. ASCE. pp. 466-480.
- Bell PS. 1999. Shallow water bathymetry derived from an analysis of X-band marine RADAR images of waves. *Coastal Engineering* **37**: 513-517.
- Blenkinsopp CE, Mole MA, Turner IL, Peirson, WL. 2010. Measurements of the time-varying free-surface profile across the swash zone obtained using an industrial LIDAR. *Coastal Engineering* **57**: 1059-1065.
- Blenkinsopp CE, Turner IL, Masselink G, Russell PE. 2011. Swash zone sediment fluxes: field observations. *Coastal Engineering* **58**: 28-44.
- Brocchini M. 2006. Integral swash-zone models. *Continental Shelf Research* **26**: 653-660.
- Butt T. 1999. Sediment transport in the swash-zone of natural beaches. *PhD Thesis*,

- Plymouth University, Institute of Marine Studies. Plymouth, Devon, UK.
- Butt T, Russell P. 1999. Suspended sediment transport mechanisms in high-energy swash. *Marine Geology* **161**: 361-375.
- Butt T, Russell P, Turner I. 2001. The influence of swash infiltration-exfiltration on beach face sediment transport: onshore or offshore? *Coastal Engineering* **42**: 35-52.
- Butt T, Russell P, Puleo J, Miles J, Masselink G. 2004. The influence of bore turbulence on sediment transport in the swash and inner surf zones. *Continental Shelf Research* **24**: 757-771.
- Elgar S, Raubenheimer B, Guza RT. 2001. Current meter performance in the surf zone. *Journal of Atmospheric and Oceanic Technology* **18**: 1735-1746.
- Elgar S, Raubenheimer B, Guza RT. 2005. Quality control of acoustic Doppler velocimeter data in the surfzone. *Measurement Science and Technology* **16**: 1889-1893.
- Emery KO, Gale JF. 1951. Swash and swash mark. *Transactions of the American Geophysical Union* **32**: 31-36.
- Erikson L, Larson M, Hanson H. 2005. Prediction of swash motion and run-up including the effects of swash interaction. *Coastal Engineering* **52**: 285-302.
- Feagin RA, Williams AM, Popescu S, Stukey J, Washington-Allen RA. 2014. The use of terrestrial laser scanning (TLS) in dune ecosystems: The lessons learned. *Journal of Coastal Research* **30**: 111-119.
- Foote M, Horn D. 2002. Using video and GIS to measure two-dimensional water surface elevations in the swash zone. *Transactions in GIS* **6**: 43-68.
- Grant US. 1948. Influence of the water table on beach aggradation and degradation. *Journal of Marine Research* **7**: 655-660.
- Grant I. 1997. Particle image velocimetry: a review. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* **211**: 55-76.
- Guedes RMC, Calliari LJ, Holland KT, Plant NG, Pereira PS, Alves FNA. 2011. Short-term sandbar variability based on video imagery: comparison between time-average and time-variance techniques. *Marine Geology* **289**: 122-134.
- Guza RT, Thornton EB. 1982. Swash oscillations on a natural beach. *Journal of Geophysical Research* **87**: 483-491.
- Haller MC, Lyzenga DR. 2003. Comparison of RADAR and video observations of shallow water breaking waves. *IEEE Transactions of in Geosciences and Remote Sensing* **41**: 832-844.
- Holland KT, Raubenheimer B, Guza RT, Holman RA. 1995. Run up kinematics on a natural beach. *Journal of Geophysical Research* **100**: 4985-4993.
- Holland KT, Puleo JA, Kooney TN. 2001. Quantification of swash flows using video-based particle image velocimetry. *Coastal Engineering* **44**: 65-77.
- Holman RA. 1986. Extreme value statistics for wave run-up on a natural beach. *Coastal Engineering* **9**: 527-544.
- Holman RA, Bowen AJ. 1984. Longshore structure of infragravity wave motions. *Journal of Geophysical Research* **89**: 6446-6452.
- Holman RA, Guza RT. 1984. Measuring run up on a natural beach. *Coastal Engineering* **8**: 129-140.
- Holman RA, Sallenger AH. 1985. Setup and swash on a natural beach. *Journal of Geophysical Research* **90**: 945-953.
- Holman RA, Stanley L. 2007. The history and technical capabilities of Argus. *Coastal Engineering* **54**: 477-491.
- Hughes MG. 1992. Application of a non-linear shallow water theory to swash following bore collapse on a sandy beach. *Journal of Coastal Research* **8**: 562-578.
- Hughes MG. 1995. Friction factors for wave uprush. *Journal of Coastal Research* **11**: 1089-1098.
- Hughes MG, Masselink G, Brander RW. 1997. Flow velocity and sediment transport in the swash zone of a steep beach. *Marine Geology* **138**: 91-103.
- Kemp PH. 1975. Wave asymmetry in the nearshore zone and breaker area *in*: Hail J, Carr A. (eds). *Nearshore Sediment Dynamics and Sedimentation*. John Wiley, Hoboken, New Jersey.

- Kingston KS, Ruessink BG, Van Eckenwort IMJ, Davidson MA. 2000. Artificial neural network correction of remotely sensed sandbar location. *Marine Geology* **169**: 137-160.
- Kobayashi N. 1997. Wave run up and overtopping on beaches and coastal structures. Centre for applied coastal research. Research Report No CACR-97-09. Delaware, USA.
- Kraus NC, Lohrmann A, Cabrera R. 1994. New acoustic meter for measuring 3D laboratory flows. *Journal of Hydraulic Engineering* **120**: 406-412.
- Lhermitte R, Serafin R. 1984. Pulse-to-pulse coherent Doppler signal processing techniques. *Journal of Atmospheric and Oceanic Technology* **1**: 293-308.
- Lippmann TC, Holman RA. 1989. Quantification of sand bar morphology: a video technique based on wave dissipation. *Journal of Geophysical Research* **94**: 995-1011.
- Lippmann TC, Holman RA. 1990. The spatial and temporal variability of sandbar morphology. *Journal of Geophysical Research* **95**: 11575-11590.
- Massa DP. 1999. Choosing an ultrasonic sensor for proximity or distance measurement, part 1: acoustic considerations. *Sensors* **16**. (www.sensorsmag.com).
- Masselink G, Puleo JA. 2006 Swash-zone morphodynamics. *Continental Shelf Research* **26**: 661-680.
- Masselink G, Russell P. 2006. Flow velocities, sediment transport and morphological change in the swash zone of two contrasting beaches. *Marine Geology* **227**: 227-240.
- McNinch JE. 2007. Bar and swash imaging RADAR (BASIR): A mobile X-band RADAR designed for mapping nearshore sand bars and swash-defined shorelines over large distances. *Journal of Coastal Research* **23**: 59-74.
- Miche R. 1951. Exposes a l'action de la houle. *Annales Pont et Chaussees* **121**: 285-319.
- Nielsen P, Hanslow DJ. 1991. Wave runup distributions on natural beaches. *Journal of Coastal Research* **7**: 1139-1152.
- Plant NG, Holman RA. 1997. Intertidal beach profile estimation using video images. *Marine Geology* **140**: 1-24.
- Plant NG, Holman RA. 1998. Extracting morphologic information from field data. *Proceedings of the 26th International Coastal Engineering Conference*, ASCE, New York, USA, pp. 2773-2784.
- Poate T, Masselink G, Davidson M, McCall R, Russell P, Turner I. 2013. High frequency in-situ field measurements of morphological response on a fine gravel beach during energetic wave conditions. *Marine Geology* **342**: 1-13.
- Power HE, Holman RA, Baldock TE. 2011. Swash zone boundary conditions derived from optical remote sensing of swash zone flow patterns. *Journal of Geophysical Research*, **116**: doi:10.1029/2010JC006724.
- Raubenheimer B, Guza RT. 1996. Observations and predictions of run-up. *Journal of Geophysical Research* **101**: 25575-25587.
- Raubenheimer B, Guza RT, Elgar S, Kobayashi N. 1995. Swash on a gently sloping beach. *Journal of Geophysical Research* **100**: 8751-8760.
- Ruessink BG, Bell PS, van Enckevort IMJ, Aarninkhof SGJ. 2002. Nearshore bar crest location quantified from time-averaged X-band RADAR images. *Coastal Engineering* **45**: 19-32.
- Ruggiero P, Holman RA, Beach RA. 2004. Wave run-up on a high-energy dissipative beach. *Journal of Geophysical Research* **109**: doi: 10.1029/2003JC002160.
- Sasaki T, Horikawa K, Hotta S. 1976. Nearshore currents on a gently sloping beach. *Proceedings of the 15th International Conference on Coastal Engineering*, ASCE, Hawaii, USA, pp. 626-644.
- Senechal N, Coco G, Bryan KR, Holman RA. 2011. Wave run up during extreme storm conditions. *Journal of Geophysical Research* **116**: doi: 10.1029/2010JC006819.
- Smith JD. 1978. Measurement of turbulence on ocean boundary layers. *Working Conference on current measurements*. Technical Report, University of Delaware, USA.
- Sonu CJ. 1972. Field observations of nearshore circulation and meandering

currents. *Journal of Geophysical Research* **77**: 3232-3247.

Sonu CJ, Pettigrew N, Fredericks RG. 1974. Measurements of swash profile and orbital motions on the beach **in:** *Ocean Wave Measurement and Analysis*. American Society of Civil Engineers **1**: 621-638.

Stockdon HF, Holman RA, Howd PA, Sallenger Jr AH. 2006. Empirical parameterisation of setup, swash, and run up. *Coastal Engineering* **53**: 573-588.

Teleki PG, Musialouski FR, Prins DA. 1976. Measurement techniques of coastal waves and currents. US Army CERC Miscellaneous Report No 76.

Turner IL, Russell PE, Butt T. 2008. Measurement of wave-by-wave bed-levels in the swash zone. *Coastal Engineering* **55**: 1237-1242.

Vousdoukas MI, Kirupakaramoorthy T, Oumeraci H, de la Torre M, Wübbold F, Wagner B, Schimmels S. 2014. The role of combined laser scanning and video techniques in monitoring wave-by-wave swash zone processes. *Coastal Engineering* **83**: 150-165.

Waddell E. 1976. Swash-groundwater-beach profile interactions **in:** Davis RA, Etherington RL (eds.) *Beach and Nearshore Sedimentation*. Society of Economic and Paleontological Mineralogists Special Publication **24**: 115-125.

Waddell E. 1980. Wave forcing of beach groundwater. *Proceedings of the 17th International Conference on Coastal Engineering*, ASCE, Sydney, Australia, pp. 1436-1452.

Particle Imaging Velocimetry

Christopher Adam Unsworth ¹

¹Department of Geography Environment and Earth Sciences, University of Hull, UK
(C.A.Unsworth@2011.hull.ac.uk)



Particle Imaging Velocimetry (PIV) is one of the least intrusive methods for measuring flow fields. The key advantage of PIV is the ability to reveal temporal linkages in fluid motion on a scale from micrometres to metres, nanoseconds to minutes with measurement accuracy high enough to permit high order statistical analysis of flow turbulence. It is an ideal method for measuring flow conditions, turbulence and Coherent Flow Structures in a range of geophysical flows.

KEYWORDS: Particle Imaging Velocimetry, Flow Turbulence, Coherent Flow Structures, Velocity, Measurement

Introduction

Environmental Fluid Flow Measurement

For an array of environmental processes, fluid flow provides the medium for landscape change. Yet this fluid medium is not an inert passenger, flows also act and produce structures that play a key role in governing sediment movement. The measurement of these turbulent fluid flows is therefore critically important in understating many of the processes in geomorphology.

Particle Imaging Velocimetry (PIV) is a measurement technique that was developed to measure wall bounded turbulent flows (Meinart 1983; Yao and Adrian 1984; Adrian and Yao 1985; Adrian 1991; Adrian 2005). The advantage of PIV is whole flow field measurement with minimal intrusion from instrumentation. A typical 2D PIV system is illustrated in Figure 1. From full flow field visualisation, the structure and spatial linkages in turbulence can be observed, measured and proved, rather than inferred, from single spatial point measurements (e.g. Bennett and Best 1995). This allows for spatial and temporal analysis techniques to be applied, from vorticity measurement (Lourenco and Krothapalli 1995; Adrian *et al.* 2000a) to eddy detection (Chakraborty *et al.* 2005) and identification of Lagrangian Coherent Structures (LCS) (Haller and Yuan

2000; Hardy *et al.* 2010a, 2010b; Haller and Sapsis 2011).

PIV systems are deployed in laboratories to provide 2D, 2.5D and full 3D flow measurements. In-field PIV can provide 2.5D and large scale water surface velocity measurements from UAV, Helium Balloon or a fixed position (e.g. Coz *et al.* 2014). PIV provides instantaneous measurement of velocities across μm -mm-m sized areas have been used in geomorphological research areas such as; around vegetation (Cameron *et al.* 2013; Okamoto and Nezu 2013; Ricardo *et al.* 2014), turbulence and sediment transport (Coleman and Nikora 2009, Ferreira *et al.* 2010), flow in local scour cavities (Unger and Hager 2007; Ferreira 2011), Coherent Flow Structures (CFS) in rivers (Best 2005; Hardy *et al.* 2009, Cooper and Tait 2010), secondary circulation in meander bends (Blanckaert *et al.* 2013) and surface velocity measurement through Large Scale PIV (LS-PIV) (Jodeau *et al.* 2008; Coz *et al.* 2014). Figure 2 displays an example of results from 2D PIV, showing an instantaneous vector field over a dune revealing flow separation in the dune lee. Figure 3 shows the same data decomposed into downstream and vertical velocities, Reynolds stress and Vorticity, revealing the larger turbulent structures in the flow.

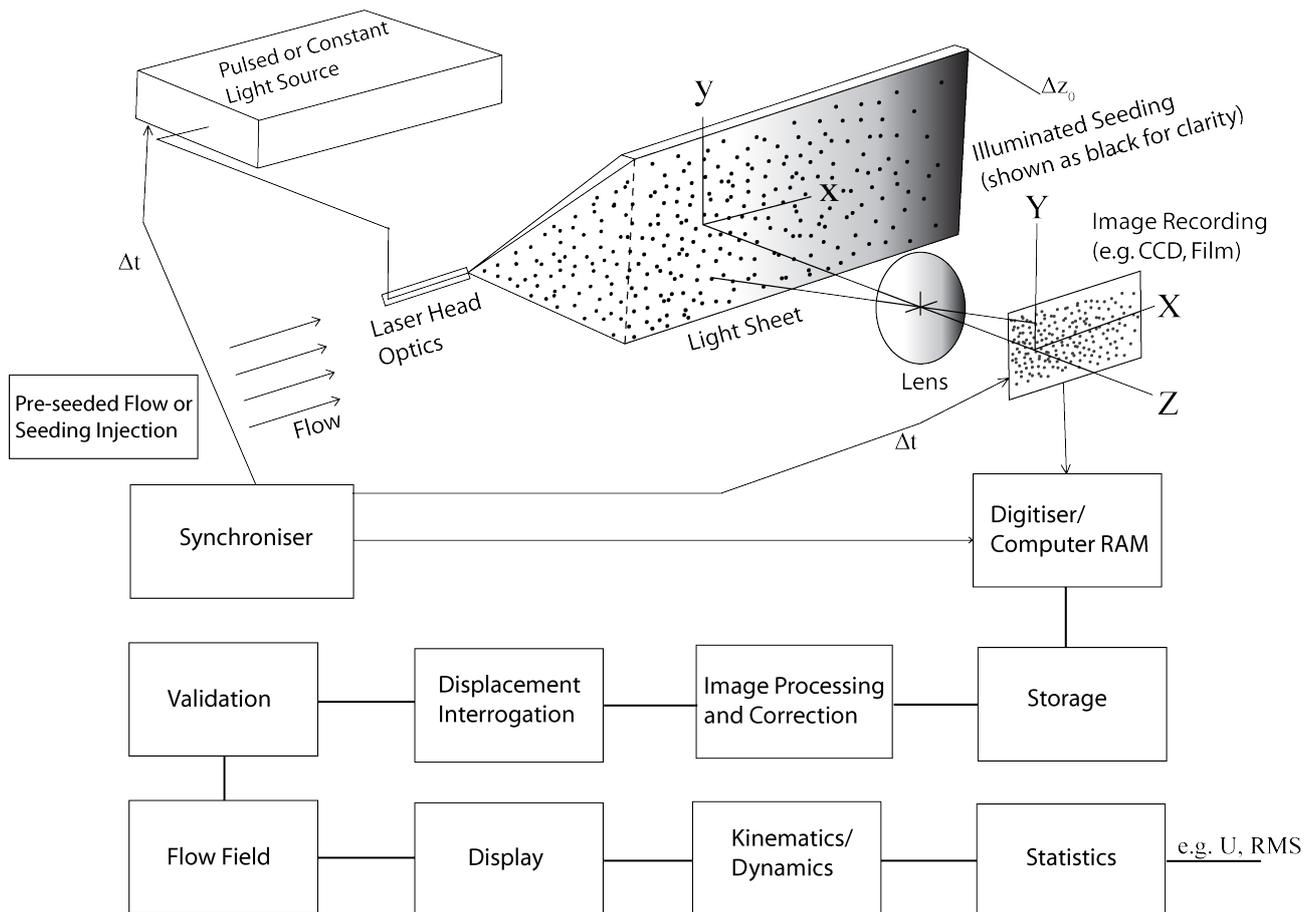


Figure 1. A typical monoscopic (2D-PIV) system. Many aspects of this setup can be changed, such as light sheet illumination source, angle into the flow (e.g. through a transparent wall to further reduce flow impact), light sheet size, intensity and shape. Seeding can be injected into the flow at specific points or fully mixed into the flow. Multiple cameras can be used to increase spatial area and/or resolution. An additional camera mirroring the above light sheet-camera arrangement (i.e. both cameras facing each other) can be used to produce 3D velocity measurements in the light sheet plane (Stereoscopic PIV). Multiple Light sheets and cameras can be used to measure several planes of the flow simultaneously (Tomographic or Holographic PIV). Figure adapted from Adrian (2005).

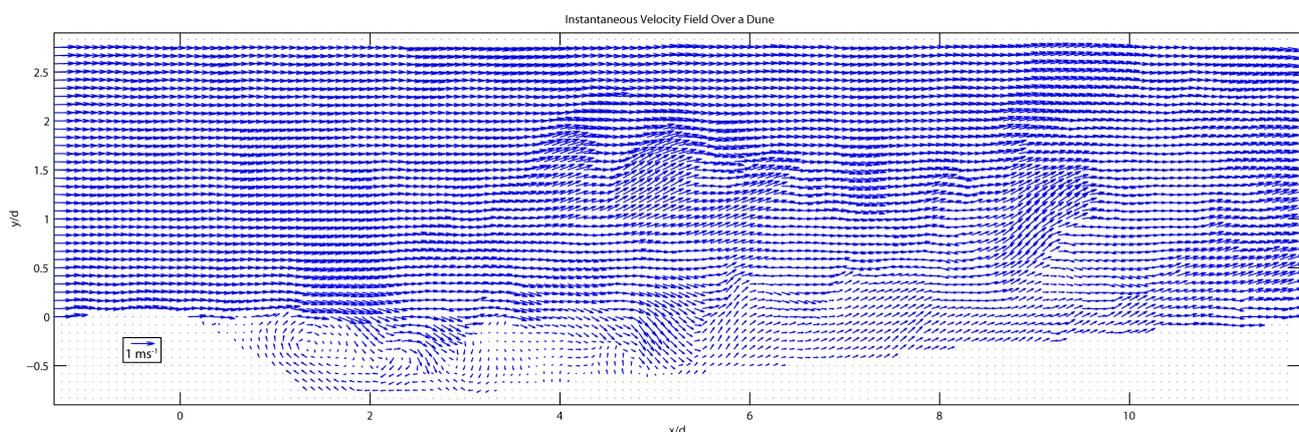


Figure 2. Instantaneous 2D flow field over the centreline of a 2D dune shape. Axis are normalised by the dune height d . Vector arrows have been doubled in size, and the downstream (u) component suppressed by 68% (Tomkins and Adrian 1999; Adrian et al. 2000b) to highlight the rotational flow over the bedform. Flow direction is left to right. Flow separation downstream of the crest can be seen in the shorter vector magnitudes and reversed direction.

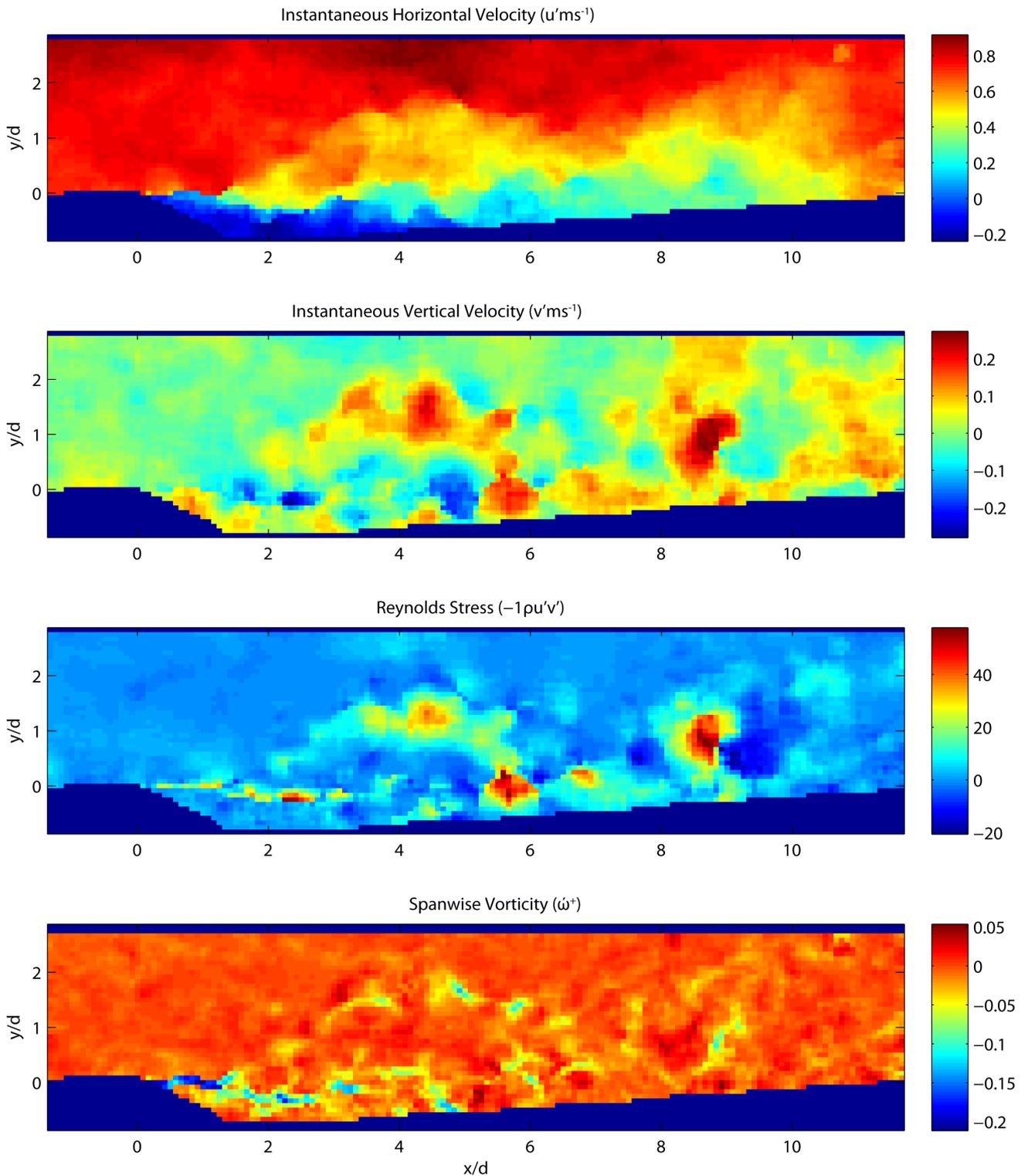


Figure 3. Instantaneous 2D PIV data. The flow field is revealed through downstream, vertical velocities, Reynolds stress and Vorticity, illustrating how PIV can display a host of flow field information including detailed structure of the turbulence. The field of view is over the centreline of a 2D dune shape fixed to the base of a flume. Axis are normalised by the dune height d . Taken at the same instantaneous point in time as Figure 2.

Methodological Overview

Basic Principles

For comprehensive reviews of standard PIV methodology and theory the books Raffell *et al.* (2007) and Adrian and Westerweel (2011) are applicable for beginners and advanced users.

The fundamental measurement in PIV is displacement and time, achieved through a now well-defined basic setup (Figure 1). Firstly, the fluid (water/air) of known density, viscosity and temperature is seeded with an even distribution of μm scale spherical particles/droplets. Particle seeding is of equal density to the fluid, small enough to respond to flow movements with minimal slip and faster than the measurement frequency whilst scattering enough light to be seen (see Hjelmfelt and Mockros 1966; Adrian 1991; Melling 1997, among other solutions). Scattering particle refractive index is different from the fluid which they are seeded into, so light is scattered effectively (Adrian and Westerweel 2011). Secondly, a light sheet produced from laser or halogen bulb is focused into a thin plane with a series of optics. A thin plane of light is required so that a key assumption in the vector calculation is achieved; a spatially consistent grid of particles. Thirdly, a camera is placed perpendicular to the light sheet. Orthogonal camera to light sheet angles is possible but requires orthorectification (see Jodeau *et al.* 2008). The camera records light intensity with a high contrast ratio. Minimum measured pixel size, cost and availability tend to define camera resolution, which typically ranges from 1MP to 20MP. The camera sensor pixel size is an important consideration as smaller camera pixels improves the accuracy of each picture at a given resolution and reduces the error in vector calculation via a higher signal to noise ratio. Images produced look like a starry night sky (Figure 4). Finally, a synchroniser is required to keep the laser and image capture synchronous.

High density seeding (measurement via group displacement) is the typical method used in PIV (Adrian 1997; Westerweel 1997). A grid of interrogation regions is overlain over each image. The size of the grid can be varied to match particle seeding density, usually in powers of 2 starting at 8 (i.e. a grid of squares each 4x4 pixels).

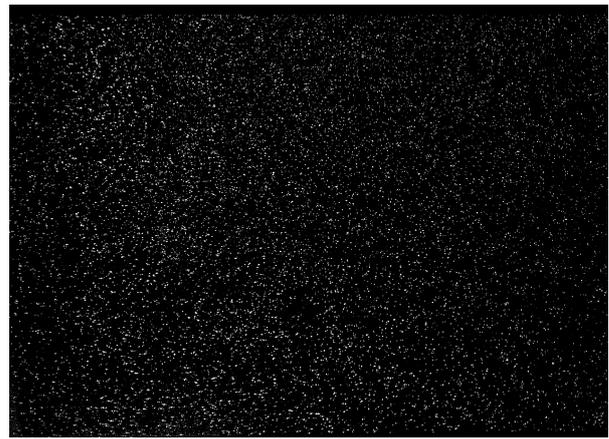


Figure 4. An example of a raw image of lab 2D-PIV, illuminated from below. Field of view is downstream of a large bar fixed to the bed, flow is left to right.

The technique of measuring particle displacement in PIV is termed cross-correlation (Keane and Adrian 1992). Particle displacement is measured by finding the maximum spatial cross correlation between two interrogation region images separated in time. This provides a known direction and magnitude shift for each interrogation region in the field of view. Known time between images then produces the velocity vector. This process is performed repeatedly over the entire array of images via a 2D Fast-Fourier Transform (Adrian 1991; Keane and Adrian 1992; Westerweel 1993). If the particles in the interrogation window do not move with equal direction and magnitude then the calculated vector correlation coefficient will be low and is therefore easily removed in post-processing using a user defined threshold of correlation. Consequently, interrogation region size and measurement frequency is crucial for realistic measurements of highly turbulent flow, e.g. shear flow.

Interrogation regions are traditionally square and regular. However, as particles can sometimes move in and out of interrogation regions and thus reduce image correlation, an adaptive grid size and shape can be used to reduce this error. This also aids in measurement of poorly seeded areas, or areas with high shear (Gruen 1985; Lourenco and Krothapalli 1995; Theunissen *et al.* 2010). Recent algorithms can dynamically identify ideal window size and shape so that the number of particles in each window is above a user-defined threshold (Theunissen *et al.* 2007, 2010).

Illumination of seeding particles for photographic capture typically falls into two methods. Dual pulse pairs, which are two very closely spaced pulses of light that are recorded and used to produce a displacement map for a single time step.

The time displacement of the pulse pairs can be tuned to an expected amount of particle displacement and ideal interrogation window size. This allows for greater displacement measurement accuracy, but requires twice the data than single pulsed measurement. Single pulsed measurement uses the displacement between two single pulses at the output measurement frequency to derive displacement- therefore any tuning to the flow character changes the output measurement frequency. Single pulse does however allow for increased maximum length of measurement from halving the data required for each time step. Equally accurate measurements compared to dual-pulse are possible but measurement frequency needs to be carefully tuned to the range of probable instantaneous velocities.

General Advantages and Limitations

As stated, the principal advantage of PIV is whole flow field velocity measurement. This can be obtained across a wide range of measurement frequencies (e.g. 1-10-100-1000 Hz available) with a dynamic velocity range typically of 1:200 (Adrian 1997). PIV velocity accuracy ranges from 0.2%-2% of flow field of view size (e.g. +/- 0.08 mm s⁻¹, Hardy *et al.* 2011). Although this is less good than LDA and LDV velocity measurements (e.g. ~0.002% of velocity range, Dantec 2014) it is still adequate for higher order statistical analysis such as vorticity (e.g. Chakraborty *et al.* 2005). Direct comparison between PIV and single point laser measurement's (LDA/LDV) is cumbersome as PIV accuracy scales with field of view size and LDA/LDV's accuracy scales with velocity range. It is certainly possible to obtain LDA/LDV accuracy measurements with PIV, but this may require small fields of view and/or very high resolution and low noise cameras.

Measurements can be made at high spatial resolution up to a μm scale, however there is usually a trade-off between spatial resolution and the size of the field of view. This can be

somewhat mediated if only mean flow measurements are needed- thus allowing for sections of a larger field of view to be measured individually then stitched together.

Flumes with transparent walls can be used to provide near-intrusion free measurement of the whole flow field. However, the addition of particle seeding, cameras, laser optics and flow stabilises to the flow does reduce how unaffected the flow measurements are.

PIV can be used at microscopic scales (endoscopic PIV, e.g. Blois *et al.* 2014) up to 10's of meters (Large Scale PIV, e.g. Muste *et al.* 2008). Maximum measurement period is a combination of the amount of computer RAM, size of each digital image, number of images captured per second and the available on board camera memory. Typical measurement times are less than 10 minutes, occupying the Macro and Meso scales of velocity spectrum (Buffin-Bélanger and Roy 2005; Marquis and Roy 2011). The length of time that the laser can be operating may also be a factor as well as any potential experimental sensitivity to temperature change produced from the laser light.

Although lower-grade and cheaper systems are becoming available, PIV is usually very expensive. The high power class 4 Nd:YAG lasers often require extensive safety protocols and user training. The construction and setup of a PIV system is often very time-consuming with low tolerance for error. If flow around objects needs to be measured it should be noted that reflections can reduce data quality. Flow around transparent objects can be measured, however if the object's refractive index is not matched to the fluids' then the laser light will refract and illuminate out-of-plane areas. If the object is not transparent, flow down-view of the laser will not be illuminated.

Types of PIV Data Collection

Standard 2D PIV produces vectors in two dimensions co-planer to the light sheet (e.g. Figures 2, 3 and 5; Willert and Gharib 1991; Adrian 1997; Westerweel 1997). Stereoscopic PIV utilises two cameras positioned off-axis to a single light sheet whilst viewing the same area to produce depth of field (Wheatstone 1838; Arroyo and Greated 1991; Prasad and Adrian 1993;

Soloff *et al.* 1997; Prasad 2000). This overcomes one of the main shortfalls of 2D PIV- the inability to simultaneously resolve all three components of velocity in a field of view. An especially significant drawback as turbulence is inherently three-dimensional (Gioia *et al.* 2010) and that geomorphological flows can produce strong flow three-dimensionality, e.g. river meanders (Bagnold 1960; Dietrich *et al.* 1979; Ferguson *et al.* 2003; Blanckaert 2011), bedforms (Maddux *et al.* 2003; Parsons *et al.* 2005; Venditti 2007; Omid and Piomelli 2013) and at confluences and around bars (Best and Roy 1991; Biron *et al.* 1996; Fujita *et al.* 1998; Lane *et al.* 1998).

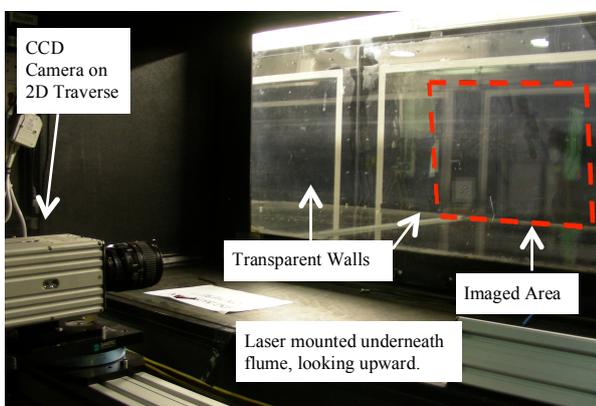


Figure 5. Example of 2D-PIV in a recirculating flume. Here the laser case and optics are outside the flume. A transparent flume floor and walls facilitate illumination and image capture.

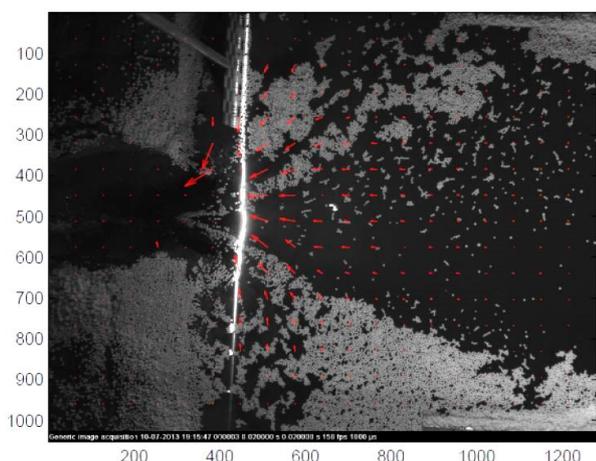


Figure 6. Example of Large Scale Surface PIV, with calculated vectors superimposed in red. (Taken from Amaral *et al.* 2014).

Large Scale PIV (LS-PIV, Figure 6) differs from traditional PIV namely in scale and the cameras oblique angle to the measurement surface. Therefore images need to be orthorectified before image interrogation (Jodeau *et al.* 2008). Flow tracers also differ as they are required to float and have higher reflectance than laboratory scale PIV as the tracers are commonly illuminated by the sun or overhead lamps. Care is also needed to reduce glare from the water surface. LS-PIV has been used to measure flow discharge (Creutin *et al.* 2003; Jodeau *et al.* 2008; Muste *et al.* 2008) water movement in floodplains and around river engineering structures (Fujita *et al.* 1998).

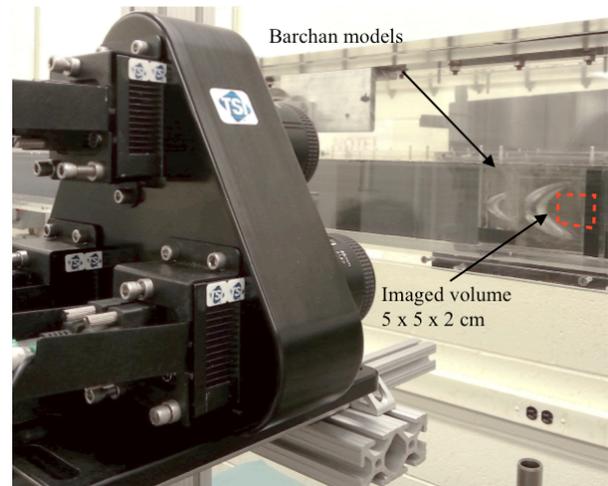


Figure 7. Example of a Volumetric PIV setup from TSI ©. This is a 3 camera system. Image credit: Gianluca Blois.

Volumetric (4D-PIV, Figure 7) velocity fields can be collected using Tomographic PIV (multiple camera) or holographic (infinite focal length camera) PIV methods. These techniques have been successfully used to measure 4D velocity fields in small volumes where distances between interrogation window measurement volumes are small enough that 3D cross correlation is possible (Barnhart *et al.* 1994; Elsinga *et al.* 2006; Kitzhofer *et al.* 2011; Schröder *et al.* 2011). Therefore the technique has generally been applied in micro-PIV studies (Malkiel 2003; Sheng *et al.* 2008). Recent advances in plenoptic (light field) cameras are allowing volumetric measurement in larger volumes- but with exponentially higher processing times (Fahringer and Thurow 2012; Lynch *et al.* 2012). In geomorphology, volumetric flow measurement has recently been used to

measure CFS over Barhcan dunes (Blois *et al.* 2012).

Assessment of Data Collection Parameters

Prior to collection of the desired flow data, an assessment of the capability of the PIV system and the data required to meet the needs of the researchers and their aims should to be performed. There are some practical trade-offs that need to be accounted for (See Chapter 10 in Adrian and Westerweel, 2011 for an extensive review of PIV set-up guidelines):

- Field of view size: In general, a larger field of view will reduce the spatial resolution of vectors by increasing the area that each camera pixel is measuring. This can be somewhat mediated by camera resolution and lens type but this may be costly if new equipment needs to be purchased.
- Seeding: Seeding needs to follow the flow with minimal lag and slip whilst reflecting enough light to be visible (see chapter 2 in Adrian and Westerweel 2011). There is a need for adequate seeding of the flow for the spatial resolution required. A minimum of 7 particles per interrogation area is recommended when using cross-correlation (Keane and Adrian 1992). Conversely, particularly high quantities of seeding may affect light intensity further away from the light source due to obscuration.
- Obstructions to the field of view: Unless the object under interrogation is transparent and the same reflective index as the fluid medium then flow “downstream” of the light source will be darkened/obscured, curtailing measurements.
- Spatial resolution of vectors: This should be known prior to entering the lab. If the data is required to validate a model then vector grid size should ideally smaller than then numerical model grid (Hardy *et al.* 2005). High spatial resolution may require a reduction in field of view and scale of experiment, which may alter the scaling that can influence the

experimental results (Peakall *et al.* 1996).

- Range of probable particle displacements: A known flow velocity is recommended so as to define sample rate, interrogation window size and overlap. The particle displacement between images-pulse pairs needs to be both large enough to be detectable but smaller than the interrogation region size possible. This can be mediated slightly with variable interrogation sizes and overlaps. This should also guide whether single or double pulse measured is used.

Data Quality and Post processing

There are several common steps that are taken to improve data quality and are commonly available in PIV capture and processing software packages. The primary objective of pre-processing the images is to increase the signal to noise ratio in the FFT calculation. Some suggestions are: removing the mean intensity from each frame; Applying a high-pass filter to the images. Manipulate the image interrogation regions by applying a Gaussian filter. Use an interrogation area overlap, dependent on particle displacement length and interrogation grid size. A signal to noise ratio floor of 1.2 is very common and can be used to prevent vector calculation in poorly seeded/illuminated regions (Keane and Adrian 1992).

Post-processing of the calculated velocity vectors is essential as incorrect vector calculation from loss of pairs, phantom particles; poor seeding concentration will all produce spurious vectors. A known upper velocity magnitude limit based upon mean flow conditions can be used to remove obviously incorrect vectors. Several vector outlier detection methods have been produced, of which the 2D median filter (Westerweel and Scarano 2005) is the most highly cited.

After the removal of poor vectors (quantity of which should be recorded as a gauge of data quality) there is the option of interpolating replacement vectors, of which there are many methods. These methods generally use the surrounding vectors to linearly interpolate replacement vectors (e.g. Garcia 2011).

Some authors prefer not to interpolate vectors as the values of the interpolated values may not reflect the real fluid flow, particularly if there is strong rotation (e.g. Hardy *et al.* 2005).

High Cost Systems

High cost systems are usually based in well-funded laboratories and offer particularly high accuracy, precision and resolution as well as setup and collection experience. High cost PIV setups are generally purchased as an entire system (lasers and optics, camera(s), computer(s) and processing software) and often cost in the order of £100,000's. PIV systems can be purchased from companies such as Dantec®, TSI® and LaVision®, among others. Usually, additional equipment is required, such as a traverse and controller system and laser light shielding for safety. Software packages tend to provide support for acquisition and processing, with a large range of processing options available that can; filter/improve image quality, perform a wide range of PIV interrogation techniques, vector filtering and interpolation and even higher order statistical analysis like Proper Orthogonal Decomposition (Berkooz *et al.* 1993).

Low cost Systems

There is an established and growing low cost option than can offer very reasonable capabilities. These systems tend to be more "home-made" and use strobe lighting instead of lasers and general SLR-type cameras to reduce costs. Continuous lasers are also used, reducing the complexity of equipment (e.g. Aleixo *et al.* 2012). Moreover, zero cost open access software is available, e.g. the free Matlab PIV toolbox (Thielicke 2014) for data collection and processing. Depending on the requirements of the research, a low cost option may have the capability to produce the desired dataset (e.g. Ryerson and Schwenk 2012) - but careful planning is certainly required. One advantage is that test runs can be performed at very low cost to establish the feasibility of a project.

Examples of PIV in Geomorphological Research

PIV tends to be used in geomorphological research to measure CFS as well as general

boundary conditions (mean velocities). In the section below, several papers that have used PIV and PIV based image interrogation are quickly summarized, and highlight how the application of PIV facilitated the authors' efforts to advance their subject.

Gravel beds

The particularly high bed roughness and irregularity of gravel beds, alongside their relatively shallow submergence in many environs exerts a significant control on the structure of fluid flowing over such beds (Wiberg and Smith 1991; Dinehart 1992; Cooper and Tait 2010). Yet the understanding of the origin of macro turbulent CFS in gravel bed rivers had not been quantitatively measured until Hardy *et al.* (2009, 2010a) used laboratory 2D PIV to capture the origin, scale and development of gravel clast produced CFS. The PIV data was used to produce mean flow, turbulence production, Reynolds Stress and Quadrant event statistics (Rao *et al.* 1971; Lu and Willmarth 1973; Bogard and Tiederman 1986; Nelson *et al.* 1993), with primary frequencies of eddy motion calculated via wavelet analysis (e.g. Torrence and Compo 1998). Hardy *et al.* (2009) found that at a low Reynolds number near bed exerted a strong control on the outer flow via these near bed generated flow structures. With higher Reynolds numbers the length flow separation behind gravel clasts increased and the separated flow regions actually merged to form a layer of skimming flow over the bed (Hardy *et al.* 2009).

It has been observed that sediment entrainment in gravel beds occurs in patches rather than a continuum (Drake *et al.* 1988). Hardy *et al.* (2010a) showed how the scale of bed roughness controlled the formation and character of macro-turbulent CFS. Indicating that large-scale CFS controlling these entrainment patches originate from bed generated turbulence and that outer flow layer flow structures are the result of flow topography interactions in the near bed region (Hardy *et al.* 2010a).

Channel shape

River meanders produce complex flow features, such as strong secondary

circulation and flow separation (Hooke 1975; Jackson 1975; Bridge and Jarvis 1977; Dietrich *et al.* 1979; Ferguson *et al.* 2003; Seminara 2010; Blanckaert 2011; Zinger *et al.* 2013). Whilst the significance of flow separation in meanders had been noted for some time (Ferguson *et al.* 2003), the flow and channel shape controls on the occurrence of flow separation were ambiguous. Blanckaert *et al.* (2013) define some the formative conditions for flow separation in meanders via the use of 3D PIV and LS-PIV to measure surface velocity and thus identify locations of flow separation in a flume facility. The LS-PIV velocity measurements were calculated using a version of “SuperPIV” whereby PIV cross-correlation calculated velocity and direction in a grid- aiding a particle tracking algorithm which produced velocity vectors per particle. The result was then interpolated onto a grid producing average flow velocity across the meander bends (Blanckaert *et al.* 2013).

The stereoscopic PIV system used facilitated non-intrusive measurement of the three components of velocity in the highly three-dimensional flow structure of an idealised sharp meander bend. This was particularly useful in revealing locations of flow stagnation, local (shear) and global (secondary flow) rotation, and the location of impinging flow both in a flat bed and with an immobile bed.

In combining these PIV techniques the authors were able to reveal some of the formative conditions for flow separation in meander bends that had previously been inferred.

Surface manifestation of coherent flow structures

Water flows over rough beds are often perturbed by upwelling fluid which often contains more suspended sediment than the surrounding flow (Matthes 1947; Coleman 1969; Jackson 1976; Lapointe 1992; Kostaschuk and Church 1993; Best 2005). These structures are generated at the rough bed boundary and mix fluid across the entire water depth (Muller and Gyr 1986; Kadota and Nezu 1999; Best 2005). Whilst qualitative description of boil types has been made (Babakaiff and Hickin 1996) it has been very difficult to quantitatively measure their size,

shape and velocity. Talke *et al.* (2013) used PIV algorithms on Infer-Red (IR) camera data of these upwelling boils in a tidal river. Temperature difference between a cooler surface and warmer sub-surface waters from the ebb tide that were upwelling were used to differentiate between normal flow and boils and facilitated a contrast that could be measured with PIV image interrogation (see Chickadel *et al.* 2011). Talke *et al.* (2013) were able to successfully measure the rate of Turbulent Kinetic Energy production and dissipation at the water surface and compare it to the rates of TKE production beneath and link this to the existence, size and intensity of boils to the unbalanced TKE budget. Best *et al.* (2010) used a similar technique on the acoustic backscatter from a multibeam sonar to identify and measure CFS under water in the Mississippi River (Best *et al.* 2010).

Other geomorphological areas that have utilised PIV include: Tracking sediment movement (Coleman and Nikora 2009) and Numerical Model Validation (Hardy *et al.* 2005). Recent advances in PIV have involved measurement of flow around vegetation (Cameron *et al.* 2013; Okamoto and Nezu 2013; Yager and Schmeckle, 2013), volumetric measurements of turbulence over river bedforms (Blois *et al.* 2012), endoscopic measurement of hyporheic flow (Blois *et al.* 2014), and coupled measurements of flow and sediment transport (e.g. Amir *et al.* 2014; Schmeckle 2014). Additionally, LS-PIV from fixed or mobile (UAV) positions has been developed in the past 5 years and facilitates measurement of velocities over very large areas (Coz *et al.* 2014), around engineered structures (Jodeau *et al.* 2008) and surface velocities (Chickadel *et al.* 2011). PIV is providing researchers with unparalleled ability to measure the processes in environmental fluid flows.

Summary

PIV is one of the least intrusive methods of measuring flow fields. Using PIV can be expensive, time consuming and complicated but it has the ability to provide velocity measurement across a range of geomorphological scales, even volumetrically, with accuracy sufficient to determine higher order statistical analysis. It is able to reveal temporal linkages in fluid motion on a scale from micrometres to

meters, nanoseconds to minutes. It is an incredibly powerful research tool that has been- and continues to be- revolutionary in nearly every subject it used, including geomorphology.

Acknowledgements

The Author would like to thank the helpful reviews from Rich Hardy and Rui Ferreira which greatly aided the clarity of the manuscript, Daniel Parsons who provided comments on a draft version of the paper, and Stuart McLelland for helpful discussions on PIV. The article was written whilst the author was a recipient of a University of Hull PhD scholarship.

References

- Adrian RJ. 2005. Twenty years of particle image velocimetry. *Experiments in Fluids*, **39**(2): 159 – 169.
- Adrian R. 1997. Dynamic ranges of velocity and spatial resolution of particle image velocimetry. *Measurement Science and Technology*, **8**: 1393.
- Adrian RJ. 1991. Particle-Imaging Techniques for Experimental Fluid-Mechanics. *Annual Review of Fluid Mechanics*, **23** (1): 261 – 304.
- Adrian RJ, Westerweel J. 2011. *Particle Imaging Velocimetry*, Cambridge: Cambridge Aerospace Series.
- Adrian RJ, Yao CS. 1985. Pulsed laser technique application to liquid and gaseous flows and the scattering power of seed materials. *Applied Optics*, **24** (1): 44 – 52.
- Adrian RJ, Christensen KT, Liu ZC. 2000a. Analysis and interpretation of instantaneous turbulent velocity fields. *Experiments in Fluids*, **29**(3): 275 – 290.
- Adrian RJ, Meinhart CD, Tomkins CD. 2000b. Vortex organization in the outer region of the turbulent boundary layer. *Journal of Fluid Mechanics*, **422**: 1 – 54.
- Aleixo R, Zech Y, Soares-Frazão S. 2012. Turbulence Measurements. In: Murillo Munoz RE (Ed.) *Dam-Break Flows: Proceedings Of The River Flow 2012 Conference*. San Jose, Costa Rica, Vol. 1, 311 – 318.
- Amaral S, Jónatas R, Bento AM, Palma J, Viseu T, Cardoso R, Ferreira RML. 2014. Failure by overtopping of earth dams. Quantification of the discharge hydrograph. In: *Proceedings of the 3rd IAHR Europe Congress: 14-15 April 2014, Portugal*, pp. 182-193: <http://webpages.fe.up.pt/iahr2014/>.
- Amir M, Nikora VI, Stewart MT. 2014. Pressure forces on sediment particles in turbulent open-channel flow: a laboratory study. *Journal of Fluid Mechanics*. **757**: 459 – 497.
- Arroyo MP, Greated CA. 1991. Stereoscopic particle image velocimetry. *Measurement Science and Technology*, **2** (12): 1181 – 1186.
- Babakaiff CS, Hickin EJ. 1996. Coherent Flow Structures in Squamish River Estuary, British Columbia, Canada. In: Ashworth PJ, Bennett SJ, Best JL, McLelland SJ. (Eds.) *Coherent Flow Structures in Open Channels*. Wiley: Chichester, UK: 321 – 342.
- Bagnold RA. 1960. *Some Aspects of the Shape of River Meanders Some Aspects of the Shape of River Meanders*, USGS Professional Paper: 282-E, Reston, VA.
- Barnhart DH, Adrian RJ, Papen GC. 1994. Phase-conjugate holographic system for high-resolution particle-image velocimetry. *Applied Optics*, **33** (30): 7159 – 7170.
- Bennett S, Best J. 1995. Mean flow and turbulence structure over fixed, two-dimensional dunes: implications for sediment transport and bedform stability. *Sedimentology*, **42**: 491 – 513.
- Berkooz G, Holmes P, Lumley JL. 1993. The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows. *Annual Review of Fluid Mechanics*, **25**: 539 – 575.
- Best JL. 2005. The kinematics, topology and significance of dune related macroturbulence: Some observations from the laboratory and field. In: Blum MD, Marriott SB, Leclair S. (Eds.) *Fluvial Sedimentology VII*. Special Publication of the International Association of Sedimentologists: 41 – 60.
- Best JL, Roy AG. 1991. Mixing-layer distortion at the confluence of channels of different depth. *Nature*, **350**: 411 – 413.
- Best JL, Simmons S, Parsons D, Oberg K, Czuba J, Malzone C. 2010. A new methodology for the quantitative visualization of coherent flow structures in alluvial channels using multibeam echo-sounding

- (MBES). *Geophysical Research Letters*, **37**(6): L06405.
- Biron P, Roy A, Best J. 1996. Turbulent flow structure at concordant and discordant open-channel confluences. *Experiments in Fluids*, **21**: 437 – 446.
- Blanckaert K. 2011. Hydrodynamic processes in sharp meander bends and their morphological implications. *Journal of Geophysical Research*, **116** (F1): F01003.
- Blanckaert K, Kleinhans MG, McLelland SJ, Uijttewaal WSJ, Murphy BJ, van de Kruijs A, Parsons DR, Chen Q. 2013. Flow separation at the inner (convex) and outer (concave) banks of constant-width and widening open-channel bends. *Earth Surface Processes and Landforms*, **38**(7): 696 – 716.
- Blois G, Best JL, Sambrook Smith GH, Hardy RJ. 2014. Effect of bed permeability and hyporheic flow on turbulent flow over bed forms. *Geophysical Research Letters*, **41**: 6435 – 6442.
- Blois G, Barros JM, Christensen KT, Best JL. 2012. An experimental investigation of 3D subaqueous barchan dunes and their morphodynamic processes. *Marine and River Dune Dynamics*, **4**: 35 – 38.
- Bogard DG, Tiederman WG. 1986. Burst detection with single point velocity measurements. *Journal of Fluid Mechanics*, **162**: 389 – 413.
- Bridge JS, Jarvis J. 1977. Velocity Profiles and Bed Shear Stress Over Various Bed Configurations In A River Bend. *Earth surface Processes*, **2**: 281 – 294.
- Buffin-Bélanger T, Roy AG. 2005. 1 Min in the Life of a River: Selecting the Optimal Record Length for the Measurement of Turbulence in Fluvial Boundary Layers. *Geomorphology*, **68** (1-2): 77 – 94.
- Cameron SM, Nikora VI, Albayrak L, Miler O, Stewart M, Siniscalchi F. 2013. Interactions between aquatic plants and turbulent flow: a field study using stereoscopic PIV. *Journal of Fluid Mechanics*, **731** (10): 345 – 372.
- Chakraborty P, Balachandar S, Adrian RJ. 2005. On the relationships between local vortex identification schemes. *Journal of Fluid Mechanics*, **535**: 189 – 214.
- Chickadel CC, Talke SA, Horner-Devine AR, Jessup AT. 2011. Infrared-Based Measurements of Velocity, Turbulent Kinetic Energy, and Dissipation at the Water Surface in a Tidal River. *Geoscience Remote Sensing Letters*, **8** (5): 849 – 853.
- Coleman KM. 1969. Brahmaputra River: Channel processes and sedimentation. *Sedimentary Geology*, **3**: 129 – 239.
- Coleman SE, Nikora VI. 2009. Bed and flow dynamics leading to sediment wave initiation. *Water Resources Research*, **45** (4): DOI: 10.1029/2007WR006741.
- Cooper JR, Tait SJ. 2010. Examining the physical components of boundary shear stress for water-worked gravel deposits. *Earth Surface Processes and Landforms*, **35**: 1240 – 1246.
- Coz J, Le Jodeau M, Hauet A, Marchand B, Le Boursicaud R. 2014. Image-based velocity and discharge measurements in field and laboratory river engineering studies using the free Fudaa-LSPIV software. In: Schleiss AJ, de Cesare G, Franca MJ, Pfister M. *River Flow 2014*. Taylor and Francis Group. London. 1961 – 1967.
- Creutin JD, Muste M, Bradley AA, Kim SC, Kruger A. 2003. River gauging using PIV techniques: a proof of concept experiment on the Iowa River. *Journal of Hydrology*, **277** (3-4): 182 – 194.
- Dantec. 2014. Dantec Website http://www.dantecdynamics.com/docs/products-and-services/fluid-mechanics/piv/Flow_Velocity_and_Turbulence_239.pdf (accessed 29/12/2014).
- Dietrich WE, Smith JD, Dunne T. 1979. Flow and Sediment Transport in a Sand Bedded Meander. *The Journal of Geology*, **87** (3): 305 – 315.
- Dinehart RL. 1992. Evolution of coarse gravel bed forms: Field measurements at flood stage. *Water Resources Research*, **28** (10): 2667 – 2689.
- Drake TG, Shreve RL, Dietrich WE, Whiting PJ, Leopold LB. 1988. Bedload transport of fine gravel observed by motion-picture photography. *Journal of Fluid Mechanics*, **192**: 193 – 217.
- Elsinga GE, Scarano F, Wieneke B, van Oudheusden BW. 2006. Tomographic particle image velocimetry. *Experiments in Fluids*, **41** (6): 933 – 947.
- Fahringer T, Thurow B. 2012. Tomographic Reconstruction of a 3-D Flow Field Using a Plenoptic Camera. *42nd AIAA Fluid Dynamics Conference and Exhibit*: 2826.

- Ferguson RI, Parsons DR, Lane SN, Hardy RJ. 2003. Flow in meander bends with recirculation at the inner bank. *Water Resources Research*, **39** (11): WR001965.
- Ferreira RML. 2011. *Turbulent flow hydrodynamics and sediment transport. Laboratory research with LDA and PIV*. In: Rowinski P. (Ed.) *Experimental Methods in Hydraulic Research*. Polish Academic Science Institute of Geophysics: 67 – 111.
- Ferreira RML, Ferreira L, Ricardo A M, Franca MJ. 2010 Impacts of sand transport on flow variables and dissolved oxygen in gravel-bed streams suitable for salmonid spawning. *River Research and Applications*, **26** (4): 414 – 438.
- Fujita I, Muste M, Kruger A. 1998. Large-scale particle image velocimetry for flow analysis in hydraulic engineering applications. *Journal of Hydraulic Research*, **36** (3): 397 – 414.
- Garcia D. 2011. A fast all-in-one method for automated post-processing of PIV data. *Experiments in Fluids*, **50** (5): 1247 – 1259.
- Gioia G, Guttenberg N, Goldenfeld P, Chakraborty P. 2010. Spectral Theory of the Turbulent Mean-Velocity Profile. *Physical Review Letters*, **105** (18): 184501.
- Gruen AW. 1985. Adaptive Least Squares Correlation: A Powerful Image Matching Technique. In *Proc. ACSM-ASP Convention*, Washington DC: 175 – 187.
- Haller G, Sapsis T. 2011. Lagrangian coherent structures and the smallest finite-time Lyapunov exponent. *Chaos*, **21** (2): 023115.
- Haller G, Yuan G. 2000. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, **147** (3-4): 352 – 370.
- Hardy RJ, Lane SN, Yu D. 2011. Flow structures at an idealized bifurcation: a numerical experiment. *Earth Surface Processes and Landforms*, **36** (15): 2083 – 2096.
- Hardy RJ, Best JL, Lane SN, Carbonneau PE. 2010a. Coherent flow structures in a depth-limited flow over a gravel surface: The influence of surface roughness. *Journal of Geophysical Research*, **115** (F3): F03006.
- Hardy RJ, Best JL, Parsons DR, Keevil GM. 2010b. On determining the geometric and kinematic characteristics of coherent flow structures over a gravel bed: a new approach using combined PLIF-PIV. *Earth Surface Processes and Landforms*, **36** (2): 279 – 284.
- Hardy RJ, Best JL, Lane SN, Carbonneau PE. 2009. Coherent flow structures in a depth-limited flow over a gravel surface: The role of near-bed turbulence and influence of Reynolds number. *Journal of Geophysical Research*, **114** (F1): F01003.
- Hardy RJ, Lane SN, Lawless MR, Best JL, Elliott L, Ingham DB. 2005. Development and testing of a numerical code for treatment of complex river channel topography in three-dimensional CFD models with structured grids. *Journal of Hydraulic Research*, **43** (5): 468 – 480.
- Hjelmfelt AT, Mockros LF. 1966. Motion of discrete particles in a turbulent fluid. *Applied Scientific Research*, **16** (1): 149 – 161.
- Hooke RL. 1975. Distribution of Sediment Transport and Shear Stress in a Meander Bend. *The Journal of Geology*, **83** (5): 543 – 565.
- Jackson RG. 1976. Sedimentological and fluid-dynamic implications of the turbulent bursting phenomenon in geophysical flows. *Journal of Fluid Mechanics*, **77** (3): 531 – 560.
- Jackson RG. 1975. Velocity - bed-form - texture patterns of meander bends in the lower Wabash River of Illinois and Indiana. *Geological Society of America Bulletin*, **86** (11): 1511 – 1522.
- Jodeau M, Hauet A, Paquier A, Le Coz J, Dramais G. 2008. Application and evaluation of LS-PIV technique for the monitoring of river surface velocities in high flow conditions. *Flow Measurement and Instrumentation*, **19** (2): 117 – 127.
- Kadota A, Nezu I. 1999. Three-dimensional structure of space-time correlation on coherent vortices generated behind dune crest. *Journal of Hydraulic Research*, **37** (1): 59 – 80.
- Keane RD, Adrian RJ. 1992. Theory of cross-correlation analysis of PIV images. *Applied Scientific Research*, **46**: 191 – 215.
- Kitzhofer J, Nonn T, Brücker C. 2011. Generation and visualization of volumetric PIV data fields. *Experiments in Fluids*, **51** (6): 1471 – 1492.
- Kostaschuk R, Church MA. 1993. Macroturbulence generated by dunes: Fraser

- River, Canada. *Sedimentary Geology*, **8**(1-4): 25 – 37.
- Lane SN, Biron PM, Bradbrook KF, Butler JB, Chandler JH, Crowell MD, McLelland SJ, Richards KS, Roy AG. 1998. Three-dimensional measurement of river channel flow processes using acoustic Doppler velocimetry. *Earth Surface Processes and Landforms*, **23**: 1247 – 1267.
- Lapointe M. 1992. Burst-like sediment suspension events in a sand bed river. *Earth Surface Processes and Landforms*, **17** (3): 253 – 270.
- Lourenco L, Krothapalli A. 1995. On the accuracy of velocity and vorticity measurements with PIV. *Experiments in Fluids*, **18** (6): 421 – 428.
- Lu SS, Willmarth WW. 1973. Measurements of the structure of the Reynolds stress in a turbulent boundary layer. *Journal of Fluid Mechanics*, **60** (3): 481 – 511.
- Lynch K, Fahringer T, Thurow B. 2012. Three-Dimensional Particle Image Velocimetry Using a Plenoptic Camera. In *50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*: 1056.
- Maddux TB, McLean SR, Nelson JM. 2003. Turbulent flow over three-dimensional dunes: 2. Fluid and bed stresses. *Journal of Geophysical Research*, **108** (F1): 6010.
- Malkiel E. 2003. The three-dimensional flow field generated by a feeding calanoid copepod measured using digital holography. *Journal of Experimental Biology*, **206** (20): 3657 – 3666.
- Marquis GA, Roy AG. 2011. Bridging the gap between turbulence and larger scales of flow motions in rivers. *Earth Surface Processes and Landforms*, **36** (4): 563 – 568.
- Matthes GH. 1947. Macroturbulence in natural stream flow. *Transactions of the American Geophysical Union*, **28**: 255 – 262.
- Meinart R. 1983. *Mesure de champs de vitesse d'écoulements fluides par analyse de suites d'images obtenues par diffusion d'un feuillet lumineux*: Université Libre de Bruxelles.
- Melling A. 1997. Tracer particles and seeding for particle image velocimetry. *Measurement Science and Technology*, **8** (12):1406 – 1416.
- Muste M, Fujita I, Hauet A. 2008. Large-scale particle image velocimetry for measurements in riverine environments. *Water Resources Research*. **44** (4): W00D19.
- Muller A, Gyr A. 1986. On the Vortex Formation in the Mixing Layer Behind Dunes. *Journal of Hydraulic Research*, **24** (5): 359 – 375.
- Nelson J, McLean SR, Wolfe SR. 1993. Mean Flow and Turbulence Fields Over Two Dimensional Bed Forms. *Water Resources Research*, **29** (12): 3935 – 3953.
- Okamoto T, Nezu I. 2013. Spatial evolution of coherent motions in finite-length vegetation patch flow. *Environmental Fluid Mechanics*, **13** (5): 417 – 434.
- Omid MH, Piomelli U. 2013. Large-eddy simulation of three-dimensional dunes in a steady, unidirectional flow. Part 2. Flow structures. *Journal of Fluid Mechanics*, **734**: 509 – 534.
- Parsons DR, Best JL, Orfeo O, Hardy RJ, Kostaschuk R, Lane SN. 2005. Morphology and flow fields of three-dimensional dunes, Rio Parana Argentina: Results from simultaneous multibeam echo sounding and acoustic Doppler current profiling. *Journal of Geophysical Research*, **110**: F04S03.
- Peakall J, Ashworth P, Best J. 1996. Physical Modelling in Fluvial Geomorphology: Principles, Applications and Unresolved Issues. In: Rhoads B, Thorn C. (Eds.) *The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology*: 27 – 29.
- Prasad AK. 2000. Stereoscopic particle image velocimetry. *Experiments in Fluids*, **29** (2): 103 – 116.
- Prasad AK, Adrian RJ. 1993. Stereoscopic particle image velocimetry applied to liquid flows. *Experiments in Fluids*, **15** (1): 49 – 60.
- Raffel M, Willert CE, Wereley ST, Kompenhans J. 2007. *Particle Image Velocimetry. A Practical Guide*. Springer, New York.
- Rao KN, Narasimha R, Badri Narayanan MA. 1971. Bursting Phenomenon in a Turbulent Boundary Layer. *Journal of Fluid Mechanics*, **48**: 339 – 352.
- Ricardo AM, Koll K, Franca MJ, Schleiss AJ, Ferreira RML. 2014. The terms of the turbulent kinetic energy budget within

- reaches of emergent vegetation. *Water Resources Research*, **50**: 4131 – 4148.
- Ryerson WG, Schwenk K. 2012. A Simple, Inexpensive System for Digital Particle Image Velocimetry (DPIV) in Biomechanics. *Journal of Integrative Biology*, **317**: 127 – 140.
- Schmeeckle MW. 2014. The role of velocity, pressure, and bed stress fluctuations in bed load transport over bed forms: numerical simulation downstream of a backward-facing step. *Earth Surface Dynamics*, **2** (2): 715 – 732.
- Schröder A, Geisler R, Staack K, Elsinga GE, Scarano F, Wieneke B, Henning A, Poelma C, Westerweel J. 2011. Eulerian and Lagrangian views of a turbulent boundary layer flow using time-resolved tomographic PIV. *Experiments in Fluids*, **50** (4): 1071 – 1091.
- Seminara G. 2010. Fluvial Sedimentary Patterns. *Annual Review of Fluid Mechanics*, **42** (1): 43 – 66.
- Sheng J, Malkiel E, Katz J. 2008. Using digital holographic microscopy for simultaneous measurements of 3D near wall velocity and wall shear stress in a turbulent boundary layer. *Experiments in Fluids*, **45** (6): 1023 – 1035.
- Soloff SM, Adrian RJ, Liu ZC. 1997. Distortion compensation for generalized stereoscopic particle image velocimetry. *Measurement Science and Technology*, **8** (12): 1441 – 1454.
- Talke SA, Devine ARH, Chickadel CC, Jessup AT. 2013. Turbulent kinetic energy and coherent structures in a tidal river. *Journal of Geophysical Research: Oceans*, **118** (12): 6965 – 6981.
- Theunissen R, Scarano F, Riethmuller ML. 2010. Spatially adaptive PIV interrogation based on data ensemble. *Experiments in Fluids*, **48**: 875 – 887.
- Theunissen R, Scarano F, Riethmuller ML. 2007. An adaptive sampling and windowing interrogation method in PIV. *Measurement Science and Technology*, **18** (1): 275 – 287.
- Thielicke W. 2014. PIVlab - time-resolved particle image velocimetry (PIV) tool. <http://www.mathworks.co.uk/matlabcentral/fileexchange/27659-pivlab-time-resolved-particle-image-velocimetry--piv--tool>.
- Tomkins CD, Adrian RJ. 1999. Characteristics of vortex packets in wall turbulence. In: Banerjee S, Eaton J. (Eds.) *Turbulence and Shear Flow Phenomena*. Begell House, New York: 31 – 36.
- Torrence C, Compo GP. 1998. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, **79** (1): 61 – 78.
- Unger J, Hager WH. 2007. Down-flow and horseshoe vortex characteristics of sediment embedded bridge piers. *Experiments in Fluids*, **42**: 1 – 19.
- Venditti J. 2007. Turbulent flow and drag over fixed two and three-dimensional dunes. *Journal of Geophysical Research: Earth Surface*, **112**: 1 – 21.
- Westerweel J. 1997. Fundamentals of digital particle image velocimetry. *Measurement Science and Technology*, **8**: 1379 – 1392.
- Westerweel J. 1993. Analysis of PIV interrogation with low pixel resolution. In: Cha SS, Trolinger D. (Eds.) *SPIE's 1993 International Symposium on Optics, Imaging, and Instrumentation*: 624 – 635.
- Westerweel J, Scarano F. 2005. Universal outlier detection for PIV data. *Experiments in Fluids*, **39**: 1096 – 1100.
- Wheatstone C. 1838. Contributions to the Physiology of Vision. Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Philosophical Transactions of the Royal Society of London*, **128**: 371 – 394.
- Wiberg PL, Smith JD. 1991. Velocity distribution and bed roughness in high-gradient streams. *Water Resources Research*, **27** (5): 825 – 838.
- Willert CE, Gharib M. 1991. Digital particle image velocimetry. *Experiments in Fluids*, **193**: 181 – 193.
- Yager EM, Schmeeckle MW. 2013. The influence of vegetation on turbulence and bed load transport. *Journal of Geophysical Research: Earth Surface*. **118** (3): 1585 – 1601.
- Yao CS, Adrian RJ. 1984. Orthogonal compression and 1-D analysis technique for measurement of 2-D particle displacements in pulsed laser velocimetry. *Applied Optics*. **23** (11): 1687 – 1689.
- Zinger JA, Rhoads BL, Best JL, Johnson KK. 2013. Flow structure and channel morphodynamics of meander bend chute

cutoffs: A case study of the Wabash River, USA. *Journal of Geophysical Research: Earth Surface*, **118** (4): 2468 – 2487.

3.3.5. Discharge Estimation: Techniques and Equipment

Richard Gravelle¹

¹ Pera Technology, Melton Mowbray, LE13 0PB (r.gravelle@peratechnology.com)



ABSTRACT: Understanding the discharge of a stream or river is essential for many important hydrological and geomorphological uses across a broad range of scales. However, it is crucial that the correct techniques and instrumentation are used depending on the stream environment and flow conditions encountered. This will ensure that discharge estimations are as reliable as possible. This review discusses the processes and instruments by which data are collected to estimate discharge. This includes stream gauging using stage and rating curves, velocity point measurements and current meters (mechanical and electromagnetic), and dilution gauging (both sudden and continuous). Advantages and disadvantages of each technique are presented to ensure that the most appropriate method can be selected.

KEYWORDS: fluvial, stream gauging, discharge, current meter, velocity-area, dilution gauging

Introduction

The discharge (or streamflow) of a river relates to the volume of water flowing through a single point within a channel at a given time. Understanding this information is essential for many important uses across a broad range of scales, including global water balances, engineering design, flood forecasting, reservoir operations, navigation, water supply, recreation, and environmental management. Growing populations and competing priorities for water, including preservation and restoration of aquatic habitat are spurring demand for more accurate, high frequency, and accessible water data (Whiting, 2003; Hirsch and Costa, 2004).

To be most useful, stream flow data must be collected in a standardized manner, with an estimate of associated accuracy and uncertainty (Pelletier, 1988; Herschy, 1995), and for a continuous time period. (Hirsch and Costa, 2004). However, hydrometric networks have limited spatial (and sometimes temporal coverage) and therefore this paper aims to describe and evaluate the processes and instruments by which discharge can be estimated.

This will include stream gauging using stage and rating curves, velocity point

measurements and dilution gauging (both sudden and continuous). Finally, current meters (both mechanical and electromagnetic) used in discharge measurements will be discussed, before a summary which will enable decisions to be made regarding the effectiveness of each technique under certain flow conditions.

Stream Gauging

Stream gauging is the technique used to measure the volume of water flowing through a channel per unit time, generally referred to as discharge.

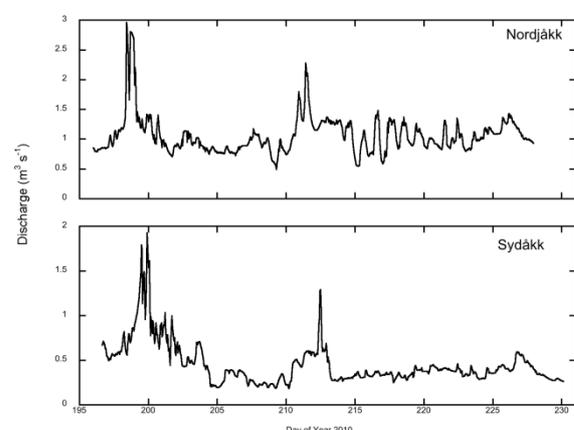


Figure 1. Examples of continuous discharge records collected at two outlet streams of Storglaciären, Sweden during the 2010 melt season. (Gravelle, 2014).

Stream discharge (Figure 1) is determined by the relationship between stream velocity and channel area and as a result, there are a number of variables that need to be recorded before discharge is calculated. Therefore, the measurement of discharge typically requires measurements of water depth, flow velocity and channel characteristics.

Quantifying the relationship between these variables allows continuous records of discharge to be estimated. The first step towards this is the measurement of stage.

Stage measurement and rating curves

Stage describes the depth of water within a channel and is quantified by the height of water at a gauging site above an arbitrary datum (Fenton and Keller, 2001). Stage can be measured in a range of ways from simply reading from a height gauge, or by using instrumentation such as float gauges (Saxon and Dye, 1995), submersible pressure transducers (Liu and Higgins, 2015) or ultrasonic gauges (Tsai and Yen, 2012).

By taking such measurements for a number of different stages and corresponding discharges (as discussed in the next section) over a period of time, a number of points can be plotted on a stage-discharge diagram, and a curve drawn through those points to give a unique relationship between stage and flow; the rating curve (Figure 2; Rantz *et al.*, 1982; Fenton and Keller, 2001; Schmidt, 2002). Understanding this relationship means that when stage is routinely measured, it is assumed that the corresponding discharge can be obtained from the equation of the line that best describes the stage-discharge relationship (Fenton and Keller, 2001).

It is important to note that whilst rating curves are commonly used, they are based on an assumption of the stage-discharge relationship, and have to describe a range of variation from low flow to extreme flood events (Fenton and Keller, 2001). In practice, discharge is often not measured (or, not measured accurately) during flood events, so discharge estimation at the extreme ends of the hydrograph may not be accurate. The effectiveness of rating curves can also be reduced in channels where sediment transport or deposition can alter the channel area. In this situation, a new rating

curve would be required if it was believed that the channel characteristics had changed.

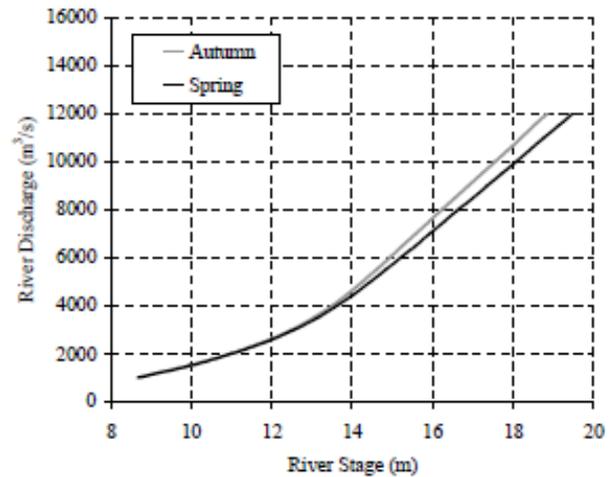


Figure 2. Examples of steady flow rating curves for spring and autumn flow regimes of the Po River, Italy (Di Baldassarre and Montanari, 2009).

Discharge Estimation

Combining continuous stage data with manual discharge measurements to form a rating curve will allow discharge data to be extrapolated to form a longer time series.

Stream discharge can be simply calculated using the equation:

$$Q = VA \quad (1)$$

where: Q is the stream discharge, V is the stream velocity, and A is the cross-sectional area of the channel perpendicular to the predominant flow direction.

The velocity-area method

The most common and direct method of estimating discharge is the velocity-area method. This technique requires measurement of stream velocity, channel width and the depth of water flow at cross-stream vertical sections (Herschy, 1995). Typically, the width of the river channel is divided into at least 20 vertical sections, with each section having no more than 10% of the total flow (see Figure 3; Shaw, 1994; Pelletier, 1988). Discharge is derived from the sum of mean stream velocity, and the channel cross-sectional area.

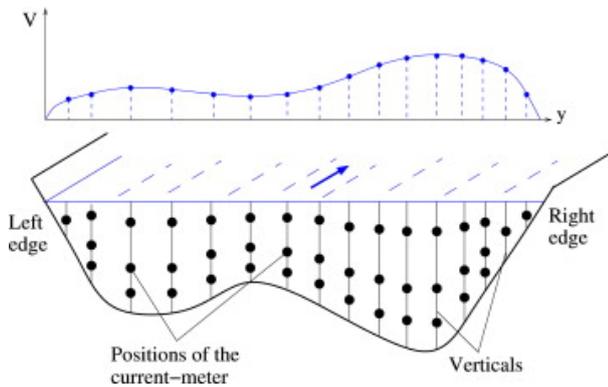


Figure 3. Principle of the velocity–area method: discrete sampling of velocity and depth throughout a cross-section, and the corresponding profile of depth-averaged velocity (Le Coz et al., 2012). Current meter positions are shown at 0.2, 0.6 and 0.8 depth.

The measurement of velocity in rivers is achieved using instruments such as current meters, and these are discussed later in the chapter. However, regardless of the instrument used, there are a number of measurement techniques which can be employed depending on river flow conditions and the time available to carry out measurements.

The most commonly used techniques are (Figure 4; Herschy, 1995):

1. the velocity distribution method;
2. the 0.6 depth method;
3. the two points method; and,
4. the three point method.

The velocity distribution method involves measurements being taken at a number of vertical sections distributed across a river channel (Carter, 1970). Measurements are taken at several points between the channel bed and the water surface, and are used to plot a vertical velocity curve. The mean velocity is obtained by dividing the area beneath the curve with the water depth (Herschy, 1995). The velocity distribution method is most accurate if conducted under steady stage conditions. However, it is generally considered too time consuming for routine gauging (Herschy, 1995).

The 0.6 depth method (sometimes referred to as the sixth-tenth depth method) involves measurements being taken at a single point at each vertical section, at 0.6 (or 60%) of the

water depth (Figure 4; Herschy, 1995). This value is assumed to be the mean velocity for that vertical, based on the analysis of many vertical velocity curves which showed acceptable accuracy of 0.6 depth measurements (Herschy, 1995). Advantages of using this method are found in the diminished number of measurements required in this technique, and therefore the reduced time taken to complete a gauging (Herschy, 1995).

The two points method involves measurements being taken at 0.2 (20%) and 0.8 (80%) depth at each vertical section (Figure 4). Similarly to the 0.6 depth method, the accuracy of this technique has been established through analysis of vertical velocity curves (Shaw, 1994). However, it is recommended that this technique only be used in rivers where the minimum depth of flow is 0.75 m (Herschy, 1995).

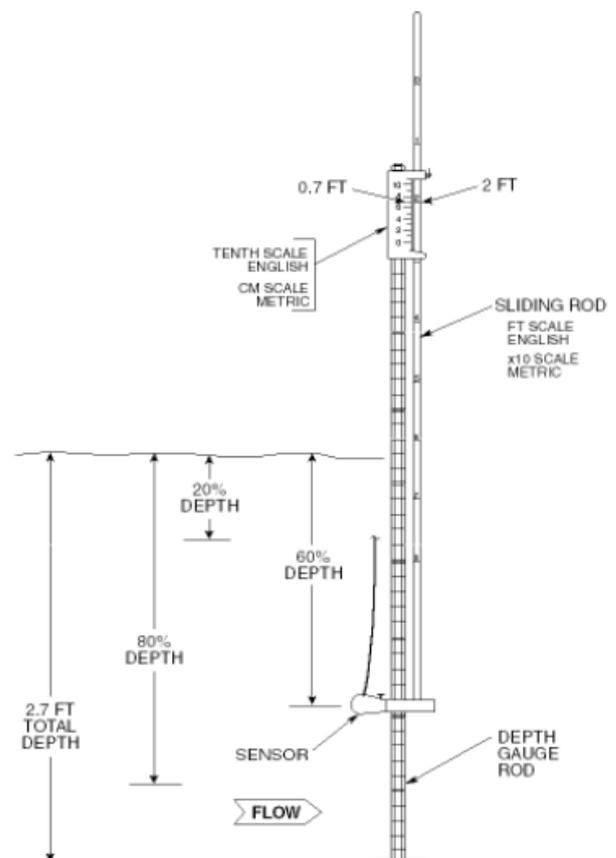


Figure 4. Diagram showing velocity measurement locations at 0.2 (20%), 0.6 (60%) and 0.8 (80%) depth within a channel, and how these are measured using a wading rod and current meter (Nolan and Shields, 2000).

The three point method is effectively an amalgamation of the 0.6 depth, and 0.2 and 0.8 depth methods, requiring measurements at each of the three depths to be taken (Figure 4). The average of the three measurements is taken as the mean vertical velocity (Hersch, 1995). Whilst accurate due to the number of measurements required, this technique is very time consuming and is therefore less widely used than the 0.6 depth method.

Regardless of the technique used, the time required to obtain an accurate velocity measurement is uncertain, with several studies suggesting differing times from 45 seconds (Carter and Anderson, 1963) to 3 minutes (Hersch, 1975). In practice, the time taken to record the measurement should reflect the level of accuracy desired (Pelletier, 1988).

Dilution gauging

Dilution gauging was originally developed as a means to measure discharge in closed channels such as pipes (Allen and Taylor, 1924; John *et al.*, 1982). However, it has become an important technique in streams where turbulent flow or non-uniform channel shape (e.g. in boulder streams) are present, limiting the accuracy of other methods of discharge calculation (e.g. the velocity-area method). For this reason, it is commonly used in studies of glacier hydrology and mountain environments (e.g. Elder *et al.*, 1990; Nienow *et al.*, 1996; Richards and Moore, 2003; Orwin and Smart, 2004).



Figure 5. Photograph showing injection of a fluorescent tracer into a channel (courtesy: http://www.openchannelflow.com/blog/archives_article).

Dilution gauging involves the injection of a tracer into a channel reach and subsequently, detection of the tracer a known distance downstream when complete mixing has occurred (Figure 5). Although this can be carried out using a number of chemical, fluorescent, or radioactive tracers (Wilson *et al.*, 1986), the most commonly used tracer is sodium chloride, in the form of common salt. Sodium chloride is preferable due to its availability and inexpensiveness, and also the ease with which it dissolves. The amount used largely depends on the length of the reach measured and the volume of water it contains, although in natural water with low background conductivity, 0.2 kg of salt per m^3s^{-1} is considered sufficient (Hersch, 1995). Detection of salt as a tracer involves the measurement of electro-conductivity (EC), the ability of a material to conduct an electrical current (Figure 6). EC can be measured using an EC probe which measures conductivity between two electrodes. Tracer injection can occur either as: (1) a single sudden injection; or, (2) a continuous injection. In both techniques, the underlying principle is the conservation of mass. Calculation of discharge is determined by the dilution of the tracer at the sampling point (Hersch, 1995).

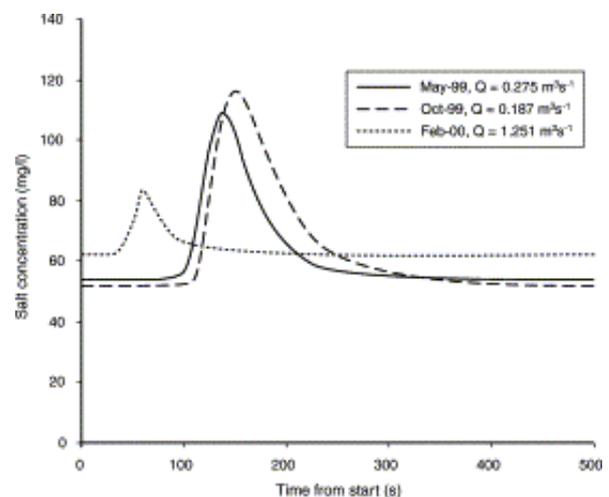


Figure 6. Variations in concentrations of dissolved salt (mg l^{-1}) with time following injection of the salt solution, 50 m downstream of the injection point in the River Holme (Wood and Dykes, 2002).

The sudden injection technique involves a single but steady injection of tracer in solution. Detection of the tracer occurs downstream at a distance sufficient to allow

complete mixing of the tracer. Sampling at this location should also be carried out over a time period which allows the whole volume of tracer to reach the sampling point (Herschy, 1995). After the injected tracer passes the sampling point, discharge can be calculated using:

$$M = vC_1 = Q \int_{t_0}^{\infty} C_2(t) dt \quad (2)$$

where M is the mass of injected tracer, v is the volume of injected tracer solution, C_1 is the concentration of the tracer solution, $C_2(t)$ is the concentration of tracer measured at the sampling point over the time interval dt , t is the elapsed time from injection to measurement, and t_0 is the time interval of the first detection of the tracer at the sampling point (Herschy, 1995).

Using the continuous injection technique, discharge can be calculated using equation 3:

$$Q = \frac{(c_1 - c_0)}{(c_2 - c_0)} q \quad (3)$$

Where c_0 is the background concentration already present in the water, c_1 is the known concentration of tracer added at a constant rate (q), and c_2 is the final concentration of tracer in the flow (Shaw, 1994).

The effectiveness of dilution gauging can be reduced where there is either inflow of water into, or abstraction of water from the reach being measured. Either of these occurrences will result in loss or excessive dilution of the tracer being used (Herschy, 1995). Dead water zones, or infiltration of the tracer into hyporheic pathways (Moore, 2004) may have a similar effect, preventing the tracer from reaching the sampling location. It is therefore crucial that the measurement reach be carefully selected before dilution gauging is attempted (Herschy, 1995). However, both the sudden and continuous injection techniques suffer from problems due to the uncertainty in ensuring adequate mixing of the tracer, and also from environmental concerns associated with the addition of chemical substances into streams (Butterworth *et al.*, 2000; Wood and Dykes, 2002). Generally, the continuous injection technique is used less commonly due to the difficulty of sustaining a constant rate of tracer flow, and the bulkiness of the

equipment required to do so (Butterworth *et al.*, 2000).

Instrumentation

The most common means of measuring stream velocity is through the use of current meters (Whiting, 2003). These can be divided into two general categories: mechanical and electromagnetic. Typically, water depth and magnitude of velocity will determine the selection of current meter for deployment. Whilst both current meter types vary in design, they are generally deployed in similar ways. Deployment of current meters can be carried out by means of a wading rod, although at greater water depths and velocities, cableways, bridges or boats may also be used (Whiting, 2003).

Mechanical current meters

Mechanical current meters (Figure 7) measure flow velocity through the rotation of a bucket wheel or impeller (Whiting, 2003). Measurement is based on the relationship between water velocity and the resulting velocity of the current meter rotor (Herschy, 1995). By counting the number of rotations of the current meter rotor during a given time period, the velocity of water can be established.

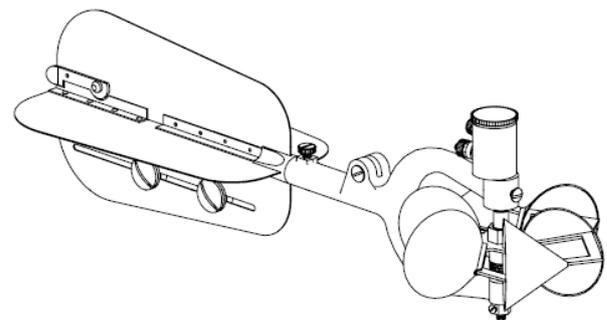


Figure 7. Price Type AA vertical axis current meter (Hubbard *et al.*, 1999).

Mechanical current meters can be classified by whether the instruments rotors are aligned on a horizontal or vertical axis. Current meters with a horizontal axis rotor are referred to as propeller-type current meters. Current meters with a vertical axis rotor are referred to as cup-type current meters (Herschy, 1995). A study comparing vertical- and horizontal-axis current meters (Fulford *et al.*, 1994) concluded that vertical-axis meters perform better in low flow velocities than

horizontal-axis meters. Horizontal-axis meters are also more prone to becoming tangled by weeds and debris, potentially reducing the accuracy of the meter (Rantz *et al.*, 1982; Fulford *et al.*, 1994). Generally however, Fulford *et al.* (1994) suggest that the uncertainties associated with vertical and horizontal-axis meter measurement are very similar, although it is thought that vertical-axis meters have an overall greater accuracy of measurement.

Electromagnetic current meters

Electromagnetic current meters (ECMs) (Figure 8) measure flow velocity using electromagnetic induction, in which the movement of water through a magnetic field generated at the head of the ECM produces a voltage which is linearly proportional to its flow velocity (Herschy, 1995; Whiting, 2003; MacVicar *et al.*, 2007). Electrodes mounted within the probe head detect changes in the electric potential of the water, caused by its movement through the magnetic field. This potential is then amplified, and converted into a readable format (Herschy, 1995).

The use of ECMs can be advantageous in channels where flow direction may be reversed (e.g. in the lee of bedforms). This is due to the ability of ECMs to measure bi-directional flow (albeit with an uncertainty of $\pm 2\%$) (Whiting, 2003). However such bedforms are capable of causing turbulence and flow separation which subsequently result in changes in flow velocity (Robert *et al.*, 1992; Roy *et al.*, 1996). If this is the case, then uncertainty can be introduced into the data collection. Roy *et al.* (1996) therefore suggest that a correction be applied

to ECM results where rotation of the ECM or flow direction is suspected.



Figure 8. Ott electromagnetic current meter (Turnipseed and Sauer, 2010).

One disadvantage of ECM use is the possibility that ECMs can be affected by electrical and magnetic fields. This includes other ECMs within a 0.6 m range (Whiting, 2003). In cold environments where ice formation is likely, ECMs can be affected by the formation of frazil ice (Derecki and Quinn, 1987). Frazil ice formation can coat current meter sensors, reducing sensitivity and producing low velocity readings, even as little as zero.

Summary of Techniques

As outlined in the sections above, each technique described may have specific advantages or disadvantages of use depending on the circumstances. Table 1 gives an indication of where a specific advantage or disadvantage exists for a given technique during certain channel or flow conditions. Although not designed to be exhaustive, this will assist in decision-making where options exist for the use of different techniques or instruments in discharge estimation.

Table 1: Matrix of the techniques and methods and their use under certain conditions. Ticks and Crosses indicate where an advantage or disadvantage has been identified. A blank space indicates that none have been identified.

	ECM	Mechanical Current Meter	Dilution/tracer gauging	Velocity distribution method	0.6 depth	Two-point method	Three-point method
Turbulent flow	x		✓	x			
Steady stage	✓	✓		✓			
Irregular channel	✓	x	✓			x	
Regular channel	✓	✓		✓	✓	✓	✓
Low flow	x	✓			✓	x	x
Reversing flow	✓	x					
Icing	x		✓				
Channel dilution			x				
Time				x	✓	✓	x

Conclusions

This review has presented several methods of stream gauging and discharge estimation with a particular focus on their appropriateness in differing flow conditions and stream settings. Although often assumed to be a single process, the estimation of stream discharge brings together a range of techniques and can be performed using a number of methods and instruments. The accuracy (and therefore uncertainty) of a measurement is minimised when the techniques and equipment are chosen in consideration of channel and flow conditions

Acknowledgements

Early versions of this review were written whilst the author was in receipt of a Loughborough University Faculty Development Studentship. The author would like to thank Richard Hodgkins and Stephen Rice for constructive comments on these drafts. The chapter was improved by the comments of two anonymous reviewers.

References

Allen CM, Taylor EA. 1924. The salt velocity method of water measurement. *Mechanical Engineering* (January): 13-16.

Butterworth JA, Hewitt EJ, McCartney MP. 2000. Discharge measurement using portable dilution gauging flowmeters. *Water and Environment Journal* **14** (6): 436-441.

Carter RW. 1970. Accuracy of current meter measurements. Hydrometry (Proceedings of the Koblenz Symposium, September 1970). *IAHS Publication* **99**: 86-98.

Carter RW, Anderson IE. 1963. Accuracy of current meter measurements. *Journal of the Hydraulics Division. Proceedings of the American Society of Civil Engineers* **8** (HY4, Part 1): 105-115.

Derecki JA, Quinn FH. 1987. Use of current meters for continuous measurement of flows in large rivers. *Water Resources Research* **23** (9): 1751-1759.

Di Baldassarre G, Montanari A. 2009. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Science* **13**: 913-921.

Elder K, Kattelmann R, Ferguson R. 1990. Refinements in dilution gauging for mountain streams. *IAHS Publication* **193**: 247-254.

Fenton JD, Keller RJ. 2001. The calculation of streamflow from measurements of stage. *Cooperative Research Centre for Catchment Hydrology Technical Report* **01/6**: 84pp.

Fulford JM, Thibodeaux KG, Kaehrle WR. 1994. Comparison of current meters used for stream gauging. In: Pugh CA (Ed), *Fundamentals and Advancements in Hydraulic Measurements and Experimentation*. American Society of Civil Engineering. 376-385.

Gravelle RD. 2014. *Temporal variability of meltwater and sediment transfer dynamics at an Arctic glacier, Storglaciären, northern Sweden*. Unpublished Ph.D.thesis, Loughborough University. 249pp. doi: 10.13140/2.1.1740.6081.

Herschey RW. 1995. *Streamflow Measurement* (2nd Edition). Taylor and Francis, London. 524pp.

Herschey RW. 1975. *The accuracy of existing and new methods of river gauging*. Unpublished Ph.D. thesis, Department of Geography, University of Reading.

Hirsch RM, Costa JE. 2004. U.S. stream flow measurement and data dissemination improve. *EOS* **85** (20): 197-203.

Hubbard EF, Thibodeaux KG, Duong MN. 1999. Quality Assurance of U.S. Geological Survey Stream Current Meters: The Meter-Exchange Program 1988-98. *USGS Open-file Report* **99-221**. 19pp.

John PH, Johnson FA, Sutcliffe P. 1982. Two less conventional methods of flow-gauging. *Advances in Hydrometry* (Proceedings of the Exeter Symposium, July 1982). *IAHS Publication* **134**: 141-152.

Le Coz J, Camenen B, Peyrard X, Dramais G. 2012. Uncertainty in open-channel discharges measured with the velocity-area method. *Flow Measurement and Instrumentation* **26**: 18-29.

Liu Z, Higgins CW. 2015. Does temperature affect the accuracy of vented pressure transducer in fine-scale water level measurement? *Geoscientific Instrumentation Methods and Data Systems* **4**: 65-73.

MacVicar BJ, Beaulieu E, Champagne V, Roy AG. 2007. Measuring water velocity in highly turbulent flows: field tests of an electromagnetic current meter (ECM) and an acoustic Doppler velocimeter (ADV). *Earth*

- Surfaces Processes and Landforms* **32**: 1412-1432.
- Moore RD. 2004. Introduction to salt dilution gauging for streamflow measurement Part 2: Constant-rate injection. *Streamline Watershed Management Bulletin* **8**(1): 11-15.
- Nienow PW, Sharp M, Willis IC. 1996. Velocity-discharge relationships derived from dye tracer experiments in glacial meltwaters: implications for subglacial flow conditions. *Hydrological Processes* **10**: 1411-1426.
- Nolan KM, Shields RR. 2000. Measurement of stream discharge by wading. *Water Resources Investigations Report 00-4036*.
- Orwin JF, Smart CC. 2004. Short-term spatial and temporal patterns of suspended sediment transfer in proglacial channels, small River Glacier, Canada. *Hydrological Processes* **18** (9): 1521-1542.
- Pelletier PM. 1988. Uncertainties in the single determination of river discharge: a literature review. *Canadian Journal of Civil Engineering* **15**: 834-850.
- Rantz SE *and others*. 1982. Measurement and computation of streamflow Vol. 1. Measurement of Stage and Discharge. *USGS Water Supply Paper 2175*. 313pp
- Richards G, Moore RD. 2003. Suspended sediment dynamics in a steep, glacier-fed mountain stream, Place Creek, Canada. *Hydrological Processes* **17** (9): 1733-1753.
- Robert A, Roy AG, De Serres B. 1992. Changes in velocity profiles at roughness transitions in coarse-grained channels. *Sedimentology* **39**: 725-735.
- Roy AG, Biron P, De Serres B. 1996. On the necessity of applying a rotation to instantaneous velocity measurements in river flows. *Earth Surface Processes and Landforms* **21**: 817-827.
- Saxon SL, Dye CW. 1995. A simple, inexpensive peak water-level gauge. *Hydrobiologia* **315**: 231-233.
- Schmidt AR. 2002. *Analysis of stage-discharge relations for open-channel flows and their associated uncertainties*. Ph.D. thesis, University of Illinois, Urbana-Champaign.
- Shaw EM. 1994. *Hydrology in Practice* (3rd Edition). Chapman & Hall, London. 569pp.
- Tsai T, Yen P. 2012. Improvement in stage measuring technique of the ultrasonic sensor gauge. *Measurement* **45**(7): 1735-1741.
- Turnipseed DP, Sauer VB. 2010. Discharge measurements at gaging stations. *USGS Techniques and Methods 3-A8*. 87pp.
- Whiting PJ. 2003. Flow measurement and characterization. In: Kondolf GM, Piégay H. (Eds.) 2003. *Tools in Fluvial Geomorphology*. John Wiley & Sons Ltd. Chichester. 688pp.
- Wilson JF, Cobb ED, Kilpatrick FA. 1986. *Fluorometric Procedures for Dye Tracing*. USGS Techniques of Water-Resources Investigations Reports. Book 3: Chapter A12. 43pp.
- Wood PJ, Dykes AP. 2002. The use of salt dilution gauging techniques: ecological considerations and insights. *Water Research* **36**: 3054-3062.

Suspended Sediment Sampling

Matthew T. Perks¹

¹ School of Geography, Politics and Sociology, Newcastle University
(matthew.perks@newcastle.ac.uk)



ABSTRACT: Fine sediment (< 2mm) is of considerable importance in fluvial systems given the physical and ecological impacts caused by elevated levels. Fine sediment is eroded from the landscape and subsequently transported in suspension by rivers and streams. Suspended sediment can be sampled using a range of manual and automated approaches designed to estimate river loads and capture samples for subsequent analysis. The most appropriate method(s) for adoption will be determined by the flow and sediment dynamics, sample/data requirements and resources available. This section presents information on a range of approaches for the direct sampling of suspended sediment in fluvial systems, discussing the advantages and disadvantages of each.

KEYWORDS: fluvial, depth-integrated, single-point, passive, time-integrating, suspended sediment sampling

Introduction

In many river systems, fine material is transported in suspension and is termed the suspended sediment load (Owens *et al.*, 2005). The fine fraction incorporates both the organic (e.g. plankton and detritus) and mineral particles (e.g. sand and silt) of diameters > 0.45 μm and < 2000 μm . Particles within this range account for the majority of material eroded from the landscape and subsequently transported by rivers and streams (Meade *et al.*, 1990). The lower boundary (0.45 μm) traditionally provides the distinction between dissolved and solid material and is somewhat of an arbitrary guideline as defined by analytical procedures (Håkanson, 2008). The upper boundary represents the transition between material typically transported close to the river bed (bed-load) and the material carried in suspension (suspended-load) (Owens, 2008). An additional distinction may also be made between fine sediment and very-fine sediment (< 62.5 μm). The latter is not controlled by the hydraulic characteristics of flow; rather its occurrence is dependent on the upstream supply rate (Khullar *et al.*, 2010). This is commonly termed the 'wash load' and constitutes an important component of the particulate flux from terrestrial surfaces

(Owens, 2008). This material may flocculate to produce much larger composite particles and can be extremely important in the transfer of pollutants and the degradation of water-bodies (Droppo, 2001, Ongley, *et al.*, 1992).

Factors influencing the optimal sampling approach

A wide range of techniques are available for sampling suspended sediment in rivers. The appropriateness of each technique is determined by the flow and sediment dynamics; sample and data requirements; and resources available. These factors will determine the sampling approach adopted and the way in which the sample is handled (transported and stored) following collection. It may therefore be pertinent to give careful consideration to the following question prior to deployment: *Given the inherent temporal variability of suspended sediment flux, cross-sectional variations in sediment transport and the mass of material required for subsequent analysis, what approaches will provide the most representative sampling method?* The sampling techniques described herein can be broadly classified as: (i) manual and, (ii) automatic.

Manual sampling

The most effective and direct means of obtaining a suspended sediment sample is through manual sampling of the river. This is considered the standard against which the accuracy of automated and indirect approaches are compared (Wren *et al.*, 2000). This is often the adopted approach for regulatory assessments, including the General Quality Assessments conducted by the Environment Agency. However, the following points relating to manual sampling should be considered: (1) safe access during high-flows must be guaranteed; (2) financial and time constraints associated with travel to and from the site may limit site visit frequency; and (3) important storm flows are infrequent and often difficult to predict. Given these constraints, it is difficult to attain continuity of high temporal resolution sampling and capture infrequent high-magnitude events using this approach in isolation. Although, this approach is often used in conjunction with automated approaches (see Automated Sampling section). The operational guidelines for manual sampling are specific to the sampling apparatus used, a range of which are now discussed in the following sections.

Depth-Integrated

Given the vertical and horizontal variability of suspended sediment concentrations (SSCs) that often exist in rivers transporting particles $> 63 \mu\text{m}$ in diameter, it is recommended that depth-integrated sampling should be undertaken to ensure a representative sample is collected (Horowitz *et al.*, 1990, Wass and Leeks, 1999). This may be achieved using depth-integrated samplers (such as the D-77 or DH-81), whereby a single discharge-weighted composite sample is collected by moving the sampling device through the stream vertical (Vanoni, 2006). This method is capable of providing vertically representative samples when the sampler is lowered to the stream bed and raised at a uniform rate. Most depth-integrated samplers were developed as part of the US Federal Interagency Sedimentation Project (FISP) and are capable of collecting samples ranging from 0.57 – 6.0 litres in volume. However, several factors limit their deployment including absolute minimum operating depths ranging from 0.08 – 0.24 m,

a maximum operating depth of 4.5 m for rigid bottle samplers and a minimum velocity of 0.45 m s^{-1} to ensure isokinetic sampling (cf. Davis, 2005).

Alternatively, samples may be acquired using point-integrating samplers (e.g. P-46, P-61 or P-72). These provide discrete representative samples of suspended sediment at the measured point. The basis of this approach is that a sufficient number of individual point samples are collected to determine the average value of the property of interest (e.g. SSC – typically reported in mg L^{-1}). Alternatively, when point samples and corresponding point velocities are integrated, it is possible to calculate the flux (total mass transported per unit time) of suspended solids, or bound constituents (Meade and Stevens Jr, 1990). Point-integrating devices are the preferred apparatus when sampling deep rivers (i.e. $> 4.5 \text{ m}$). However, their deployment is restricted by the hydraulic characteristics of the flow, absolute minimum sampling depths of 0.11 – 0.15 m and a limited sampling capacity of 0.57 or 1.14 L, which may be insufficient for some applications (cf. Davis, 2005).

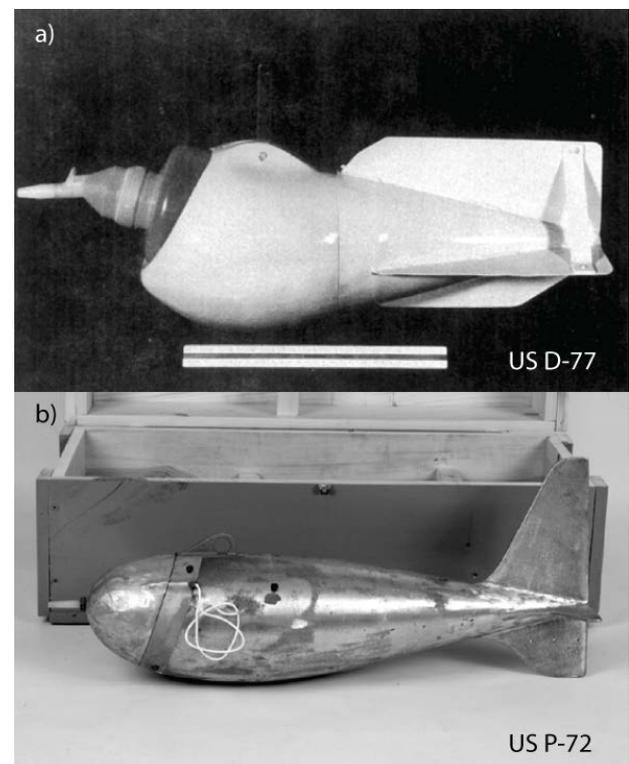


Figure 1: Examples of a) depth-integrating and b) point-integrating suspended sediment samplers designed through the US Federal Interagency Sedimentation Project (FISP). (Modified from www.rickly.com).

To account for horizontal variations in suspended sediment, multiple vertical sampling profiles should be conducted across the channel. An absolute minimum of four verticals is required. There are two commonly adopted approaches for calculating the required number: (1) equal discharge increment (EDI); or (2) equal width increment (EWI). Detailed descriptions of these are provided in Gray *et al.* (2008: available online).

Single-point

Whilst it is best practise to sample across the flow cross-section, it is acknowledged that this may not always be feasible, especially when the river is in spate (Abtey and Powell, 2004). Assessments have therefore been conducted on the representativeness of single-point samples. It has been observed that single representative samples can be obtained in shallow, well mixed streams where suspended sediment is uniformly distributed along the vertical and horizontal planes (Sheldon, 1994). This is achieved by positioning the sampler intake at 60% of the stream depth (Newburn, 1988). However, to ensure representativeness of the single-point sample, *a priori* measurements should be made at more than 10 locations (determined using the aforementioned EWI method) through the cross-section to determine the relation between the average and the point at which sampling is to be undertaken. A coefficient can then be produced to convert future discrete samples to the mean cross-sectional value (Horowitz, 1995). Alternatively, this information can be used to determine the optimal location for sampling (Porterfield, 1977). However, it should also be acknowledged that the relation between the average and the point at which sampling is undertaken and therefore the optimal location may not be constant, becoming modified with changes in bed forms, source and type of sediment.

Automated Sampling

An alternative to manually sampling the river is to deploy apparatus capable of automatically collecting a sample without a field operator being present. This is possible using basic passive samplers, more advanced pump samplers, or time-integrating

sampling devices. Each is assessed in the following sections.

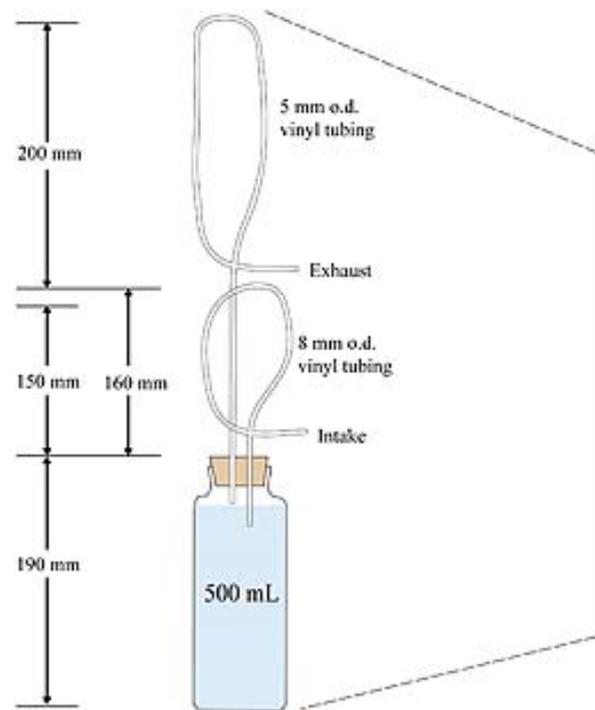


Figure 2: Siphon sampler (modified after Graczyk *et al.* 2000)

Passive

The single-stage sampler (FISP, 1961), also known as a siphon sampler, is an example of a passive sampling device which is fixed at the stage height at which the sample will be taken. Water enters the intake tube as the river level rises. The sample then enters the sampler body upon the creation of a siphon when the water level exceeds the height of the intake tube loop (Figure 2). An airlock is created when the water level in the bottle reaches the exhaust tube, preventing further filling (Mackay and Taylor, 2012). These devices may be staggered vertically at a single point in the channel (e.g. Estrany *et al.*, 2009b), or along a cross-section (e.g. Shellberg *et al.*, 2013). Single-stage samplers may be useful in capturing material for SSC determination (e.g. Estrany *et al.* 2009a, Estrany *et al.* 2009b) and determination of particle size characteristics (e.g. Kostaschuk *et al.* 2003) especially in remote locations. These devices have been reported to produce representative samples in small streams (Collins, 1981) although trials in large rivers have reported differences

between manual and single-stage sampling in the region of 10 - 20% for SSCs (Batalla, 1993). Modifications to the original rising-stage sampler have enabled sampling of the falling-stage; secondary rises and; the collection of sufficient material for geochemical analysis (e.g. Minella *et al.* 2009).

Automated Pump

The collection of suspended sediment samples has been made significantly easier in recent years following the commercial availability of automatic pump samplers. These consist of an intake, sample distributor, pump, bottle container unit and activation system (Gray *et al.* 2008), whereby a sample volume is drawn up from the channel through the creation of a partial vacuum (Newburn, 1988). These samplers have become efficient, lightweight, affordable, and computer controlled, allowing sampling to be triggered remotely (e.g. via SMS) or by external trigger devices (e.g. in response to changes in river flow). This remote activation has enabled greater precision and frequency of sampling during storm events. Most pump sampling equipment takes samples at a fixed point in the river cross-section, although depth-proportional sampling is possible (Eads and Thomas, 1983). During deployment, EWI sampling may be conducted to ensure the representativeness of the discrete or depth-integrated sample. Finally, the intake should be faced upstream (Navratil *et al.* 2011). However, debris fouling and the potential for the purge cycle to be compromised against strong flow often leads to the intake being fixed perpendicular to the flow direction.

Discrete samples collected using automatic samplers have been shown to be comparable with those derived using manual sampling methods (Graczyk *et al.* 2000). However, they operate best in fine grained fluvial environments due to the samplers' inability to collect samples isokinetically (Lewis and Eads, 2008). Where sand-sized material is in transport, the particle size distribution and amount of sediment collected may be compromised (Bent *et al.* 2001).

Time-Integrating

Various time-integrating devices have been designed and used for monitoring purposes (Vanoni, 2006). A popular device is that designed by Phillips *et al.* (2000). It was originally developed to trap sediment through principles of sedimentation to be used for the assessment of physical, geochemical and magnetic properties of transported material in lowland rivers dominated by very-fine suspended sediment (e.g. Phillips *et al.* 2000, Russell *et al.* 2000). If deployed appropriately, the device is subject to the full range of flow conditions over the sampling period, providing a continuous record of suspended sediment flux, which may be representative of all events (Walling, 2005). The device has been used in a variety of fluvial environments for sediment source ascription studies (e.g. Collins *et al.* 2010, Fox and Papanicolaou, 2007, Fukuyama *et al.* 2010) and to assess sediment fluxes (e.g. Schindler Wildhaber *et al.* 2012). The device has also been subject to modifications for optimal operation in upland catchments (Figure 3; Perks *et al.* 2013) and arctic fluvial systems (McDonald *et al.* 2010).

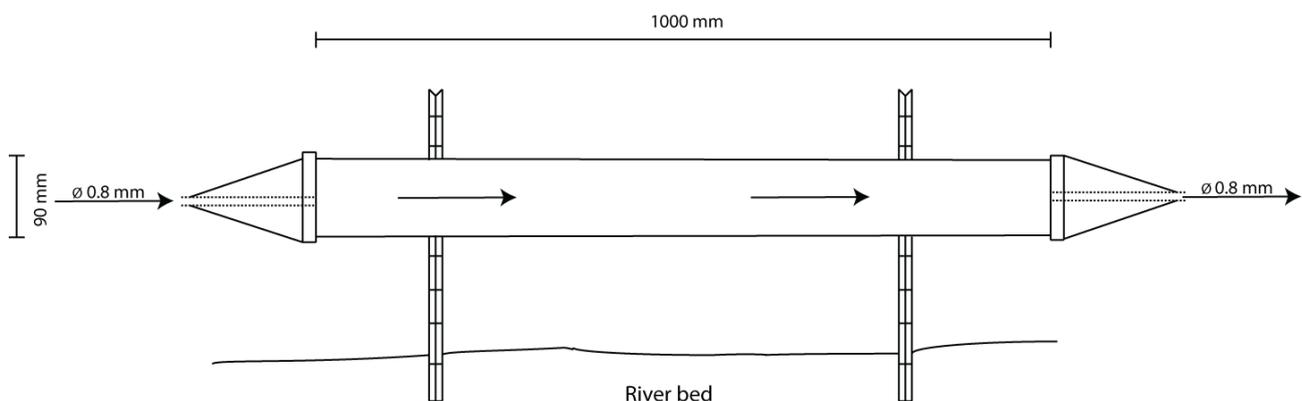


Figure 3: Schematic of a time-integrating mass-flux sampler (after Perks *et al.* 2013)

During deployment, the time-integrating sampler should be installed along a straight river reach with the inlet facing upstream. The device may sample a fixed position in the cross section (e.g. Perks *et al.* 2013, Schindler Wildhaber *et al.* 2012) or may have a variable sampling height (e.g. McDonald *et al.* 2010). The sampler is often left in-situ for a prolonged period (e.g. 30 days) to capture a sufficient mass of sediment for subsequent analysis. At the end of the sampling period, the device is removed from the river and the fine sediment is collected in sufficiently large containers to store the entire sediment-aqueous mix from within the sampler. The sampler should be rinsed and relocated with the samples taken to the laboratory for analysis.

Conclusion

Suspended sediment sampling methods can be categorised into: (1) manual sampling approaches capable of capturing a mass of material which is representative of the sediment flux; (2) devices capable of collecting discrete samples which can be passive or intelligently controlled to sample during events of interest; and (3) devices which are capable of collecting material which is potentially representative of the ambient flux over the entire monitoring period. Devices falling under (1) (i.e. manual samplers) require the presence of a field operative; largely precluding their applicability for studies interested in the dynamic nature of suspended sediment. Devices within group (2) (i.e. passive and pump samplers) are capable of collecting representative samples when the apparatus is appropriately located. They may also be configured to capture a sufficient mass of material for accurate analysis of, for example trace quantities of bound substances. Devices within group (3) (i.e. time-integrating devices) provide the potential means of overcoming the lack of temporal integration associated with (1) and provide a composite sample for analysis. However, further research is required to ensure that fully representative samples are collected using these devices. Each of these approaches may be used very effectively in combination (e.g. Perks *et al.* 2013). Although ultimately, a balanced decision based on the sample and data requirements, resources available and conditions of

deployment must be made as to the most suitable approach and method(s) to adopt.

Acknowledgements

The author would like to thank Dr Jeff Warburton for constructive comments on an earlier draft and also the comments of two anonymous reviewers.

References

- Abtew W, Powell B. 2004. Water quality sampling schemes for variable flow canals at remote sites. *JAWRA Journal of the American Water Resources Association* **40**: 1197-1204
- Batalla R.J. 1993. Sand-bed transport contribution to the sediment budget of a granitic Mediterranean drainage basin. Unpublished PhD Thesis. University of Barcelona.
- Bent GC, Gray JR, Smith KP, Glysson GD. 2001. A synopsis of technical issues for monitoring sediment in highway and urban runoff. In *US Geological Society open-file report*. Reston, VA: USA.
- Collins AL, Walling DE, Stroud RW, Robson M, Peet LM. 2010. Assessing damaged road verges as a suspended sediment source in the Hampshire Avon catchment, southern United Kingdom. *Hydrological Processes* **24**: 1106-1122. DOI: 10.1002/hyp.7573
- Collins MB. 1981. Sediment yield studies of headwater catchments in Sussex, S.E. England. *Earth Surface Processes and Landforms* **6**: 517-539. DOI: 10.1002/esp.3290060603
- Davis BE. 2005. A Guide to the Proper Selection and Use of Federally Approved Sediment and Water Quality Samplers. U.S. Geological Survey: Virginia; 26.
- Droppo IG. 2001. Rethinking what constitutes suspended sediment. *Hydrological Processes* **15**: 1551 - 1564. DOI: 10.1002/hyp.228
- Eads RE, Thomas RB. 1983. Evaluation of a depth proportional intake device for automatic pumping samplers. *JAWRA Journal of the American Water Resources Association* **19**: 289-292. DOI: 10.1111/j.1752-1688.1983.tb05328.x

- Estrany J, Garcia C, Batalla RJ. 2009a. Groundwater control on the suspended sediment load in the Na Borges River, Mallorca, Spain. *Geomorphology* **106**: 292-303. DOI: 10.1016/j.geomorph.2008.11.008
- Estrany J, Garcia C, Batalla RJ. 2009b. Suspended sediment transport in a small Mediterranean agricultural catchment. *Earth Surface Processes and Landforms* **34**: 929-940. DOI: 10.1002/esp.1777
- Federal Interagency Sedimentation Project (1961). Report No. 13 The single-stage sampler for suspended sediment. Washington, D.C.: 105.
- Fox JF, Papanicolaou AN. 2007. The use of carbon and nitrogen isotopes to study watershed erosion processes. *Journal of the American Water Resources Association* **43**: 1047-1064. DOI: 10.1111/j.1752-1688.2007.00087.x
- Fukuyama T, Onda Y, Gomi T, Yamamoto K, Kondo N, Miyata S, Kosugi Ki, Mizugaki S, Tsubonuma N. 2010. Quantifying the impact of forest management practice on the runoff of the surface-derived suspended sediment using fallout radionuclides. *Hydrological Processes* **24**: 596-607. DOI: 10.1002/hyp.7554
- Graczyk DJ, Robertson DM, Rose WJ, Steuer JJ. 2000. Comparison of Water-Quality Samples Collected by Siphon Samplers and Automatic Samplers in Wisconsin. United States Geological Survey: Reston, VA: USA; 4.
- Gray JR, Glysson GD, Edwards TK. 2008. Suspended-sediment samplers and sampling methods. In *Sedimentation engineering: processes, measurements, modeling, and practise*, García MH (ed). ASCE Publications: Reston, Virginia: USA; 320 - 339.
- Håkanson, L. 2008. Suspended particulate matter in lakes, rivers and marine systems. The Blackburn Press, Caldwell, NJ, 331
- Horowitz AJ. 1995. The use of suspended sediment and associated trace elements in water quality studies. International Association of Hydrological Sciences: Wallingford
- Horowitz AJ, Rinella AR, Lamothe P, Miller TL, Edwards TK, Roche RL, Rickert DA. 1990. Variations in suspended sediment and associated trace element concentrations in selected riverine cross sections. *Environmental Science & Technology* **24**: 1313 - 1320. DOI: 10.1021/es00079a003
- Khullar NK, Kothyari UC, Raju KGR. 2010. Suspended Wash Load Transport of Nonuniform Sediments. *Journal of Hydraulic Engineering* **136**: 534-543. DOI: 10.1061/(ASCE)HY.1943-7900.0000223
- Kostaschuk R, Terry J, Raj R. 2003. Suspended sediment transport during tropical-cyclone floods in Fiji. *Hydrological Processes* **17**: 1149-1164. DOI: 10.1002/hyp.1186
- Lewis J, Eads R. 2008. Implementation guide for turbidity threshold sampling: principles, procedures, and analysis. Department of Agriculture, Forest Service, Pacific Southwest Research Station: Arcata, CA: USA.
- Mackay AK, Taylor MP. 2012. Event-based water quality sampling method for application in remote rivers. *River Research and Applications* **28**: 1105-1112. DOI: 10.1002/rra.1504
- McDonald DM, Lamoureux SF, Warburton J. 2010. Assessment of a time-integrated fluvial suspended sediment sampler in a high arctic setting. *Geografiska Annaler. Series A, Physical Geography* **92**: 225 - 235. DOI: 10.1111/j.1468-0459.2010.00391.x
- Meade RH, Stevens Jr HH. 1990. Strategies and equipment for sampling suspended sediment and associated toxic chemicals in large rivers — With emphasis on the Mississippi river. *Science of The Total Environment* **97–98**: 125-135. DOI: 10.1016/0048-9697(90)90235-M
- Meade RH, Yuzyk TR, Day TJ. 1990. Movement and storage of sediment in the rivers of the United States and Canada. In *Surface water hydrology: The Geology of North America*, Wolman MG, Riggs HC (eds). Geological Society of America: Boulder, Colorado.
- Minella JPG, Merten GH, Walling DE, Reichert JM. 2009. Changing sediment yield as an indicator of improved soil management practices in southern Brazil. *Catena* **79**: 228-236. DOI: 10.1016/j.catena.2009.02.020
- Navratil O, Esteves M, Legout C, Gratiot N, Nemery J, Willmore S, Grangeon T. 2011. Global uncertainty analysis of suspended sediment monitoring using turbidimeter in a small mountainous river catchment. *Journal*

- of Hydrology **398**: 246-259. DOI: 10.1016/j.jhydrol.2010.12.025
- Newburn LH. 1988. Modern sampling equipment: Design and application. In *Principles of environmental sampling*, Keith LH (ed). American Chemical Society: Salem, MA: USA; 209 - 220.
- Ongley ED, Krishnappan BG, Droppo G, Rao SS, Maguire RJ. 1992. Cohesive sediment transport: emerging issues for toxic chemical management. *Hydrobiologia* **235-236**: 177-187. DOI: 10.1007/bf00026210
- Owens PN. 2008. Sediment behaviour, functions and management in river basins. In *Sustainable Management of Sediment Resources*, Owens PN (ed). Elsevier; 1-29.
- Owens PN, Batalla RJ, Collins AJ, Gomez B, Hicks DM, Horowitz AJ, Kondolf GM, Marden M, Page MJ, Peacock DH, Petticrew EL, Salomons W, Trustrum NA. 2005. Fine-grained sediment in river systems: environmental significance and management issues. *River Research and Applications* **21**: 693-717. DOI: 10.1002/rra.878
- Perks MT, Warburton J, Bracken L. 2013. Critical assessment and validation of a time-integrating fluvial suspended sediment sampler. *Hydrological Processes*. DOI: 10.1002/hyp.9985
- Phillips JM, Russell MA, Walling DE. 2000. Time-integrated sampling of fluvial suspended sediment: a simple methodology for small catchments. *Hydrological Processes* **14**: 2589 - 2602. DOI: 10.1002/1099-1085(20001015)14:14<2589::AID-HYP94>3.0.CO;2-D
- Porterfield G. 1977. Computation of fluvial-sediment discharge. In *Techniques of Water-Resources investigations of the United States Geological Survey*. USGS; 71.
- Russell MA, Walling DE, Hodgkinson RA. 2000. Appraisal of a simple sampling device for collecting time-integrated fluvial suspended sediment samples. *International Association of Hydrological Sciences*: Wallingford
- Schindler Wildhaber Y, Michel C, Burkhardt-Holm P, Bänninger D, Alewell C. 2012. Measurement of spatial and temporal fine sediment dynamics in a small river. *Hydrology and Earth System Sciences* **16**: 1501 - 1515. DOI: 10.5194/hess-16-1501-2012
- Sheldon LR. 1994. Field guide for collecting and processing stream-water samples for the National Water-Quality Assessment Program. 62.
- Shellberg JG, Brooks AP, Rose CW. 2013. Sediment production and yield from an alluvial gully in northern Queensland, Australia. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp.3414
- Vanoni VA. 2006. Sedimentation engineering: Manuals and Reports on Engineering Practice No. 54. Vanoni VA (ed). American Society of Civil Engineers Publications: Reston, VA.
- Walling DE. 2005. Tracing suspended sediment sources in catchments and river systems. *Science of The Total Environment* **344**: 159-184. DOI: 10.1016/j.scitotenv.2005.02.011
- Wass PD, Leeks GJL. 1999. Suspended sediment fluxes in the Humber catchment, UK. *Hydrological Processes* **13**: 935 - 953. DOI: 10.1002/(SICI)1099-1085(199905)13:7<935::AID-HYP783>3.0.CO;2-L
- Wren DG, Barkdoll BD, Kuhnle RA, Darrow RW. 2000. Field Techniques for Suspended-Sediment Measurement. *Journal of Hydraulic Engineering* **126**: 97-104. DOI: 10.1061/(ASCE)0733-9429(2000)126:2(97)

The Use of Passive Integrated Transponder (PIT) Systems for Tracing Particles and Organisms in Rivers

Matthew F. Johnson¹

¹ Department of Geography, Loughborough University, UK (M.F.Johnson@lboro.ac.uk)



Passive Integrated Transponder (PIT) systems can be used to trace organisms and materials through Earth surface environments. Radio-tagging offers a number of benefits over traditional tracing techniques, particularly because it is a non-destructive method of repeatedly identifying tagged objects. The small size, long battery life and relatively low cost of PIT-tags are key advantages over other types of radio-tracking allowing for a large number of individuals to be repeatedly located over extended time periods (> years). PIT-tags can be detected with a hand-held portable antenna that is swept across an environment or by an automated system that continuously records when tags are in range of a fixed antenna or multiple antennae. Antennae usually have detection ranges of 0.01 m to 3 m which can be a practical disadvantage. However, the detection range also represents the spatial accuracy resulting in tags being located with greater precision than many other radio-tagging techniques. An automated detection system using an array of 16 antennae is described which has been used to continuously record the location of tagged signal crayfish in a small stream in the UK. Whilst PIT-tags are widely used in ecological studies to track organisms they are, as yet, only infrequently used in geomorphological studies despite their great potential for particle tracing studies.

KEYWORDS: passive integrated transponders; radio frequency identification; radio-telemetry, tracers

Introduction

The tracing of materials through Earth surface environments has led to many advances in our knowledge of both geomorphological and biological systems. For example, tracing sediment particles has led to a greater understanding of the spatial and temporal dynamics of sediment flux in fluvial and marine environments (Hassan and Ergenzinger, 2003) and the tracing of organisms has produced insights concerning the movement, migration and habitat use in a range of ecosystems (Gibbins and Andrews, 2004). Tracers vary widely between studies and can range in complexity from marking clasts with paint or dye through to using radio-telemetry. This technical note focuses on the use of Passive Integrated Transponder (PIT) tracing systems.

Passive and active radio-tags

Passive Integrated Transponder (PIT) tags are a type of radio-tag which are located using Radio Frequency Identification (RFID) technology. A distinction can be made between active and passive radio-tags with PIT tags belonging to the passive group. An active radio-tag sends a signal to an antenna array giving a continuous path of movement of the object being tracked (Habersack, 2001). Consequently, active radio-tags require an internal power source. PIT tags do not contain an internal power source. Instead, when an antenna is within range of a tag (usually ~ 1 m) an electromagnetic field is generated from which the tag derives power so that it can transmit information to the antenna and tag reader. This information is usually used to ascertain the point location of the tag at a specific time.

Both PIT tags and antennae are variable in size and design. However, the general principle is the same for all PIT systems. A PIT tag is a microchip which is usually encased in a glass tube approximately 2 mm wide and between 10 – 20 mm long (Figure 1). The range and efficiency of tag detection depends on the electromagnetic field created by the antenna and its attenuation by the object within which it is encased or buried. Consequently, the antenna design and tag size will influence the detection distance with most systems having detection ranges of 0.01 m to 3 m. PIT tags can be fixed onto or placed inside sediment particles and organisms (with virtually no negative impacts on animals; Gibbins and Andrews, 2004).

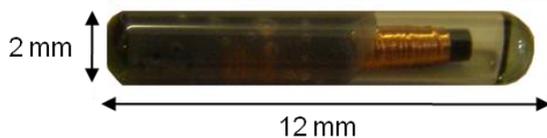


Figure 1: Photograph of a typical PIT-tag

Benefits and weaknesses

Radio-tags versus traditional techniques

Both active and passive radio-telemetry methods have a number of advantages over other forms of tracing. First, it is not necessary for an operator to be able to see the tagged object in order to locate it. This ensures that retrieval rates are typically higher than for visual marking techniques using painted or dyed objects which may be out of view (for example, they may be buried or in turbid water) or difficult to detect because markings have faded or otherwise degraded. Second, identifying an individual tagged object does not depend on an operator being able to physically retrieve it because electronic tags transmit a unique identification code. Again, this is advantageous in comparison to painted or similar identifiers, which are prone to become faded or illegible. It is also an advantage over magnetic tagging. With magnetic tags, locating an object using a magnetometer does not require that it is visible (e.g. buried clasts can be located), but identifying which particular object one has found requires that the object is recovered in order to examine its

markings. Third, the use of traditional tracing techniques requires physically searching for the tagged object, so that retrieval becomes a destructive process, particularly if sediment particles are buried or if organisms are hiding. In contrast, because radio-tagged objects can be located and identified without physical disturbance (for example, they have been detected in buried sediments to a depth of approximately 1 m), radio-tagging can have little or no impact on the system being studied. The main limitation of radio-tagging is the expense in comparison to traditional techniques, which can limit the number of objects being traced. This is important in sediment transport studies where large numbers of observations are required to adequately characterise natural variability.

Passive versus active radio-tags

The key advantage of PIT-tags over active tags is their relatively low cost, although this cost remains greater than more traditional techniques. The antenna array is the largest expense, the PIT tags themselves costing approximately £2 each (2009 prices; see practical issues below for full pricings) allowing for numerous individual clasts or organisms to be tagged once an antenna system has been purchased. Further advantages of PIT tags over active tags are their small size and the fact that tracing is not limited by onboard battery life. A weakness relative to active tags is the small detection range which requires the antenna to be within approximately 1 m of the tag. However, detection distances are increasing with technological advancements and for many applications short-range detection can be viewed as a benefit because it means that the spatial error in locating the tag is small (centimetres) relative to active tags (metres) and errors associated with recording several tags simultaneously are minimised.

Perhaps the greatest benefit of PIT tags is the very high reliability in tag detection (95 – 100%) and reading accuracy (100%) (Gibbins and Andrews, 2004). However, this efficiency can be affected by the speed at which tags are moving, tag angle relative to the antenna, and the number of tags simultaneously within range of an antenna (Castro-Santos *et al.*, 1996). Errors associated with these issues tend to be infrequent and limited, for instance, PIT tags attached to fish have been

detected when travelling at 3.6 m s^{-1} (Prentice *et al.*, 1990) and were detected while moving at up to 8 m s^{-1} in the system used by Downing *et al.* (2001). Errors are also small relative to the large quantity of data obtained using this technique and can be minimised with careful experimental design. However, tag readers cannot register through some substances such as metal (Freeland and Fry, 1995) and not all PIT tags can be read by all readers because they need to be on the same radio frequency for codes to be received.

Specific examples of use

PIT tags were initially used to identify individual animals using handheld tag readers that are held directly over the tags attached to or inserted into the animal, for example in agriculture (Freeland and Fry 1995). Subsequently, many of the advancements in PIT tag usage have been associated with ecological studies. However, PIT tags have been used in some geomorphological studies and they offer significant potential for geomorphic applications.

Portable detectors

Portable detectors are swept across an environment, much like a metal detector (Roussel *et al.*, 2000; Morhardt *et al.*, 2000). These tend to consist of a circular antenna mounted on a pole. Bubb *et al.* (2006) were able to successfully track the movement of tagged signal crayfish through the River Wharfe, UK using such antenna and they have been widely used to locate tagged fish (Roussel *et al.*, 2000; Zydlewski *et al.*, 2001; Hill *et al.*, 2006), including fish in ice covered streams (Limnansaari *et al.*, 2007). More recently, this technique has been used to identify and trace sediment particles in marine, hill-slope, and fluvial environments (Nichols, 2004; Allan *et al.*, 2006; Carré *et al.*, 2007; Lamarre *et al.*, 2008). Nichols (2004) was able to locate 96% of cobble tracers after four major run-off events from hillslopes, whereas only 63% would have been located without radio-tags. A similar technique was used by Allan *et al.*, (2006) to study the transport of cobbles across beaches with a 90% recovery rate after 8 months. After flood events, 87 – 96% of tagged particles (*b* axis

40 – 250 mm) were recovered from Moras Creek, Québec, including those buried up to 0.25 m (Lamarre *et al.*, 2005). Rollet *et al.* (2008) were able to recover 87 – 90% of PIT-tagged particles after two years in a small river (6 m wide) but recovery was reduced to only 36% after one year in a large river (100 m wide). Despite this reduction, most likely due to the inability to locate tags in deep pools and when buried at depths greater than 0.25 m, the recovery rate was still substantially greater than other tracing techniques (Rollet *et al.*, 2008). This highlights that in fluvial environments the use of hand-held antennae is particularly well suited to small-moderate sized streams. The main weakness of using hand-held antennae is that the spatial and temporal extent and resolution of the data obtained is dependent on intensive surveying effort, making it a labour intensive approach.

Automated detection of objects

Rather than sweeping an environment with an antenna, automated detection systems can be used. In this case, one or more antennae are distributed in space (for example, across the river bed) and a record is made every time a tag is within range of an antenna. When using multiple antennae it is necessary for them to be connected to a Multi-Point Decoder (MPD) that sequentially interrogates each antenna in turn. If a tag is detected, its identification number, the antenna number and a time and date is logged and recorded by the MPD. Automated systems are an effective, non-destructive way of monitoring the passage of particles or organisms or for establishing residence times at points within the sampling frame. The main limitation of using automated systems is that antennae must be placed in locations where the tagged objects will come into range of them. This assumes some knowledge of the future location of tagged particles. Consequently, automated tracking is also particularly well suited to the fluvial environment.

Automated systems have been used for a variety of ecological studies but there are no published accounts of use in geomorphic studies. For example, Riley *et al.* (2003; 2006) used an automated system consisting of two MPD connected to 31 antennae placed

on the substrate surface to monitor the location of different fish species within a river channel. A similar technique was used by Johnston *et al.* (2009) who used a grid of 242 buried antennae to detect the location of tagged fish. Large, flat antennae can be constructed across entire channel widths to record every time a tagged fish enters or exits a particular channel reach (Lucas *et al.*, 1999; Greenburg and Giller, 2000). In the Columbia River Basin over 12 million fish have been PIT tagged and monitored since 1987 (www.psmfc.org/PIT_Tag_Information_System_PTAGIS).

With careful consideration of antennae placement, the automated approach has the potential to be used to study the dynamics of tagged particles in rivers and other geomorphological settings. As well as being able to relate high resolution temporal and spatial point locations of sediment grains to hydraulic conditions, it could provide valuable information on the storage of bed material by continuously recording where grains come to rest and how long they remain stationary.

Practical issues

The author used PIT tags to continuously trace the location of signal crayfish in relation to substrate and flow conditions in a small English stream for a period of 150 days in 2009. Multiple antennae were used and each was buried below the surface of discrete habitat units in a meander bend with sites selected to reflect combinations of substrate and hydraulic conditions. This information was used to relate experimental laboratory studies on the reworking of fluvial substrates (Johnson *et al.*, 2010) to a field environment; in essence to show that substrates comprising those gravel sizes which crayfish activity can alter, were occupied by crayfish in the field.

We used 16 circular, 0.25 m diameter antennae (ANT-SP-DISC-250) connected to a Multi-Point Decoder (DEC-MPD-16) with 10 m coaxial cables (figure 2). Tags (12 x 2.2 mm) and all equipment were purchased in the UK from Wyre Micro Design Ltd. The 16 antennae were interrogated over a three second period, with each antenna in turn

being activated for 300 milliseconds. This rapid interrogation removed potential

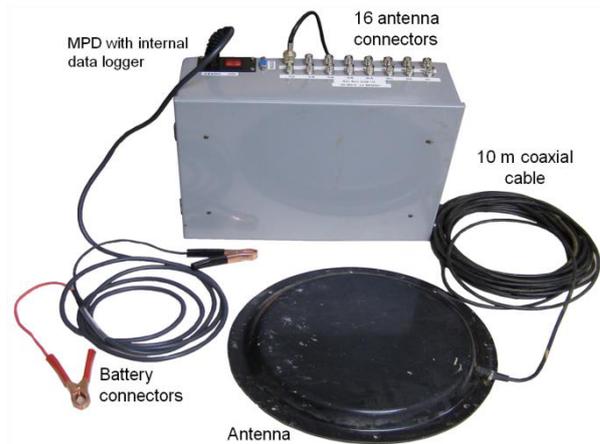


Figure 2: Photograph of the equipment used by the author to track signal crayfish in an English river.

problems of multiple tags being recorded at the same time or interference between antennae positioned close together. Each antennae cost approximately £175 (2009 price) and the MPD cost £1450 (2009 price). Over the tracking period, we successfully recorded over 10,000 point locations of a total of 65 PIT-tagged crayfish. Therefore, we obtained a continuous record of every time a PIT-tagged crayfish was present in a predefined habitat unit over the 150 day study.



Figure 3: Tracking equipment in situ at the field site used by the author. The MPD and two 12V batteries were housed in the black box and antennae were buried under the substrate of the river.

Conclusions

PIT tags are a relatively inexpensive way of accurately tracking large numbers of individual particles or organisms over extended periods. The small size and mass of the tags (< 10 mm, 0.1 grams) means that relatively small clasts or organisms (crayfish, juvenile fish) can be tracked, and their long life makes them ideal for tracking over extended periods (> years). The relatively small detection range between tags and antenna may be seen as a weakness as it necessitates some knowledge of the future location of the individuals being tracked. However, the short detection range makes the technique ideally suited to fluvial environments where the location of sediments and organisms tend to be limited to within the channel. The small detection range also results in a greater spatial accuracy when locating objects. The flexibility of the approach holds great promise for geomorphological studies of particle dynamics.

Links

AVID, Inc. (American Veterinary Identification Devices): www.avidid.com

Biomark, Inc.: www.biomark.com

Lotek Wireless, Inc.:
www.lotek.com/index.htm

Trovan, Ltd.: www.trovan.com

Wyre Micro Design, Ltd:
www.wyremicrodesign.co.uk/

References

Allen JC, Hart R, Tranquili JV. 2006. The use of Passive Integrated Transponder (PIT) tags to trace cobble transport in a mixed sand-and-gravel beach on the high-energy Oregon coast, USA. *Marine Geology* **232**: 63–86.

Bubb DH, Thom TJ, Lucas MC. 2006. Movement patterns of the invasive signal crayfish determined by PIT telemetry. *Canadian Journal of Zoology* **84**: 1202–1209.

Carré DM, Biron PM, Gaskin SJ. 2007. Flow dynamics and bedload transport around paired deflectors for fish habitat enhancement: a field study in the Nicolet River. *Canadian Journal of Civil Engineering* **34**: 761–769.

Castro-Santos T, Haro A, Walk S. 1996. A passive integrated transponder (PIT) tag system for monitoring fishways. *Fisheries Research* **28**: 253–261.

Downing SL, Prentice EF, Frazier RW, Simonson JE, Nunnallee EP. 2001. Technology developed for diverting passive integrated transponder (PIT) tagged fish at hydroelectric dams in the Columbia River Basin. *Aquacultural Engineering* **25**: 149–164.

Freeland WJ, Fry K. 1995. Suitability of passive integrated transponder tags for marking live animals for trade. *Wildlife Research* **22**: 767–773.

Gibbons JW, Andrews KM. 2004. PIT Tagging: Simple Technology at Its Best. *Bioscience* **54**: 447–454.

Greenberg LA, Giller PS. 2000. The potential of flat-bed passive integrated transponder antennae for studying habitat use by stream fishes. *Ecology of Freshwater Fish* **9**: 74–80.

Habersack HM. 2001. Radio-tracking gravel particles in a large braided river in New Zealand: a field test of the stochastic theory of bed load transport proposed by Einstein. *Hydrological Processes* **15**, 377–391.

Hassan MA, Ergenziner P. 2003. Use of tracers in fluvial geomorphology. In *Tools in fluvial Geomorphology*, Kondolf GM, Piegay H. (eds). John Wiley and Sons, Chichester, UK; 397–424.

Hill MS, Zydlewski GB, Zydlewski JD, Gasvoda JM. 2006. Development and evaluation of portable PIT tag detection units: PIT packs. *Fisheries Research* **77**: 102–109.

Johnson MF, Rice SP, Reid, I. 2010. Topographic disturbance of subaqueous gravel substrates by signal crayfish (*Pacifastacus leniusculus*). *Geomorphology* **123**: 269–278.

Johnston P, Bérubé F, Bergeron NE. 2009. Development of a flatbed passive integrated transponder antenna grid antenna grid for continuous monitoring of fishes in natural streams. *Journal of Fish Biology* **74**: 1651–1661.

- Lamarre H, MacVicar B, Roy AG. 2005. Using passive integrated transponder (PIT) tags to investigate sediment transport in gravel-bed rivers. *Journal of Sediment Research* **75**: 736–741.
- Lamarre H, Roy AG. 2008. The role of morphology on the displacement of particles in a step-pool river system. *Geomorphology* **99**: 270–279.
- Limnansaari T, Roussel J-M, Cunjak RA, Halleraker JH. 2007. Efficacy and accuracy of portable PIT-antennae when locating fish in ice-covered streams. *Hydrobiologia* **582**: 281–287.
- Lucas MC, Mercer T, Armstrong JD, McGinty S, Rycroft P. 1999. Use of a flat-bed passive integrated transponder antenna array to study the migration and behaviour of lowland river fishes at a fish pass. *Fisheries Research* **44**: 183–191.
- Morhardt JE, Bishir D, Handlin CI, Mulder SD. 2000. A portable system for reading large passive integrated transponder tags from wild trout. *North American Journal of Fish Management* **20**: 276–283.
- Nichols MH. 2004. A radio frequency identification system for monitoring coarse sediment particle displacement. *Applied Engineering Agriculture* **20**: 783–787.
- Prentice EF, Flagg TA, McCutcheon CS. 1990. Feasibility of using implantable passive integrated transponder (PIT) tags in salmonids. *American Fisheries Society Symposium* **7**: 317–322.
- Riley WD, Eagle MO, Ives MJ, Rycroft P, Wilkinson A. 2003. A portable passive integrated transponder multi-point decoder system for monitoring habitat use and behaviour of freshwater fish in small streams. *Fisheries Management and Ecology* **10**: 265–268.
- Riley WD, Ives MJ, Pawson MG, Maxwell DL. 2006. Seasonal variation in habitat use by salmon, *Salmo salar*, trout, *Salmo trutta* and grayling, *Thymallus thymallus*, in a chalk stream. *Fisheries Management and Ecology* **13**: 221–236.
- Rollet AJ, Macvicar B, Piegay H, Roy A. 2008. L'utilisation de transpondeurs passif pour l'estimation du transport sédimentaire: premiers retours d'expérience. *La Houille Blanche* **4**: 110–116.
- Roussel J-M, Haro A, Cunjak RA. 2000. Field test of a new method for tracking small fishes in shallow rivers using passive integrated transponder (PIT) technology. *Canadian Journal of Fisheries and Aquatic Science* **57**: 1326–1329.
- Zydlewski GB, Haro A, Whalen KG, McCormick SD. 2001. Performance of stationary and portable passive transponder detection systems for monitoring of fish movements. *Journal of Fish Biology* **58**: 1471–1475.

3.4.1. Sampling and describing glacier ice

Mario Toubes-Rodrigo¹, Simon J. Cook¹, David Elliott¹, Robin Sen¹

¹School of Science and the Environment, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, UK (M.Toubes-Rodrigo@mmu.ac.uk)



ABSTRACT: Determination of the physical, chemical and biological properties of glacier ice is essential for many aspects of glaciology and glacial geomorphology. In this chapter, we draw principally on examples of the description and sampling of the basal zone of glaciers where the ice is in direct contact with its substrate, and hence is where a great deal of geomorphological work is achieved. Whilst a pre-determined sampling strategy is essential to inform sampling equipment requirements, flexibility in data collection is necessary because of the dynamic nature of glaciers, and variability of ice exposure. Ice description is best achieved through stratigraphic logging, section drawing and photography. Detailed description can include a variety of information about the nature of layering, structures and sediment distribution; the size, shape and roundness of included debris; ice crystallography; and bubble content. It is common practice to categorise descriptively different ice types into cryofacies, so that comparisons can be made between studies. Sample extraction may be required for more detailed analyses of the physical, chemical and microbiological composition of the ice. We outline the use of a number of tools for ice sample extraction, including chainsaws, ice axes, chisels and ice screws.

KEYWORDS: cryofacies, glacial sediment, ice crystallography, sampling, stratigraphy

Introduction

Glaciers are highly heterogeneous in nature, comprising a wide variety of ice types with different characteristics. Traditionally, glaciologists and geomorphologists have focused on characterizing the physical and chemical nature of glacier and basal ice (e.g. Hubbard and Sharp, 1989; Knight, 1997), although increasingly, biological characteristics are being considered (e.g. Hodson *et al.*, 2008; Montross *et al.*, 2014).

The diversity and complexity of ice types result from factors including flow and strain histories within the glacier; the character of parent snow, ice or water, as well as entrained sediment; sediment availability and processes of entrainment; melting and re-freezing; and many more (Hubbard and Sharp, 1989; Knight, 1997). Accounting for differences in ice composition is important. For example, ice characteristics (e.g. sediment content, structure, presence of chemical impurities) affect its rheological

properties, and hence impact upon ice flow (e.g. Fitzsimons, 2006; Chandler *et al.*, 2008). Equally, the origin and history of the ice can be interpreted from its physical characteristics (e.g. Knight, 1997; Hambrey and Lawson, 2000; Cook *et al.*, 2010, 2011a; Lovell *et al.*, 2015).

For the most part, glacial geomorphologists are interested in the amounts of geomorphological activity (i.e. erosion and sediment transfer) achieved by glaciers (e.g. Hallet *et al.*, 1996), and the deposition of sediment to create landforms and sediments, such as moraines and till (e.g. Cook *et al.*, 2011b). Increasingly, however, there is recognition of the role of glaciers in global biogeochemical cycles (especially carbon), and the discharge of carbon and other nutrients to downstream ecosystems (e.g. Hood *et al.*, 2015). Hence, there is a need to describe, sample and classify the different ice types that exist within glaciers. In this chapter, we outline how this is achieved, focusing on the basal zone of glaciers where

the glacier interacts directly with its substrate, which is of most relevance to geomorphologists.

Selecting locations for ice description and sampling

The fundamental issue facing most researchers is what to describe and sample. As with any study, a carefully designed sampling strategy is important in order not to bias results, and the exact approach will depend on the purpose of the study. The dynamic nature of glaciers can make planned sampling strategies difficult to apply. For example, some cryofacies (i.e. distinct ice types) may be visible on some field visits and not on others, and safe access to sampling locations is always a consideration and limitation. Nonetheless, it is useful to have a sample 'wish list' before leaving for fieldwork (particularly, to inform equipment requirements) but flexibility is often necessary depending on what is visible and accessible in the field. It is also worthwhile consulting recent satellite imagery (e.g. most recent Google Earth or Landsat imagery) before

fieldwork to plan access to the glacier margin and surface. Once in the field, it is worthwhile devoting time (perhaps a few days) to reconnoitering the glacier margin, identifying suitable sampling locations and practicing sampling before embarking on the sampling campaign itself.

Describing and classifying ice

In most cases, the description and sampling of ice are undertaken together in the field, starting with description. Once a suitable location has been selected, it is good practice to document the nature of the site prior to sampling, which usually requires the removal of large quantities of ice and sediment. Take field sketches, photographs (using reference objects, such as ice axes or people, for scale) and any basic measurements (e.g. height) of the undisturbed ice section or surface. Basal ice exposures are commonly covered by a surficial smear of sediment, and it may be necessary to clean the section (e.g. by sluicing it with meltwater) before logging and photographing to reveal ice types more clearly (Figure 1).



Figure 1: (a) Cleaning an ice section by sluicing with meltwater. (b) Example of basal ice section that has been cleaned by sluicing. Note how different ice types (a lower layered ice type and an upper coarse, white ice type) are clearer in the cleaned section compared to the non-cleaned sections either side.

The process of stratigraphic logging starts with the identification of the cryofacies present within the section, which in turn requires ice description. A cryofacies, in the context of the cryospheric sciences, is ice that has a suite of characteristics that enable it to be distinguished from other ice types (e.g. Hubbard *et al.*, 2009).

Several physical characteristics can be explored when producing a stratigraphic log or section diagram, and these characteristics are often used to classify ice into constituent cryofacies (e.g. Lawson, 1979; Hubbard *et al.*, 2009). Examples of visually different cryofacies are shown in Figure 2 and give some impression of how variable basal

cryofacies can be. The characteristics recorded in basal ice stratigraphic logs are in many ways similar to those recorded for studies in sedimentary geology (e.g. Nichols, 2009; Miall, 2016), and include:

- ice layer thickness
- nature of contacts between layers (sharp, gradational, unconformable, etc.)
- horizontal extent of the layer (e.g. whether it is continuous, or pinches-out laterally)
- evidence for tectonism or deformation (e.g. folds, faults, boudins)
- sediment content (as volumetric or gravimetric percentage of sediment in the ice)
- sediment distribution (e.g. solid sediment with interstitial ice, random distribution of clasts or sediment aggregates within an ice matrix, bands of clasts or aggregates)
- nature of the entrained sediment (estimated particle size, lithology, shape and roundness of particles, etc.)
- ice crystal size
- bubble content, size and shape (e.g. spherical, elongate).

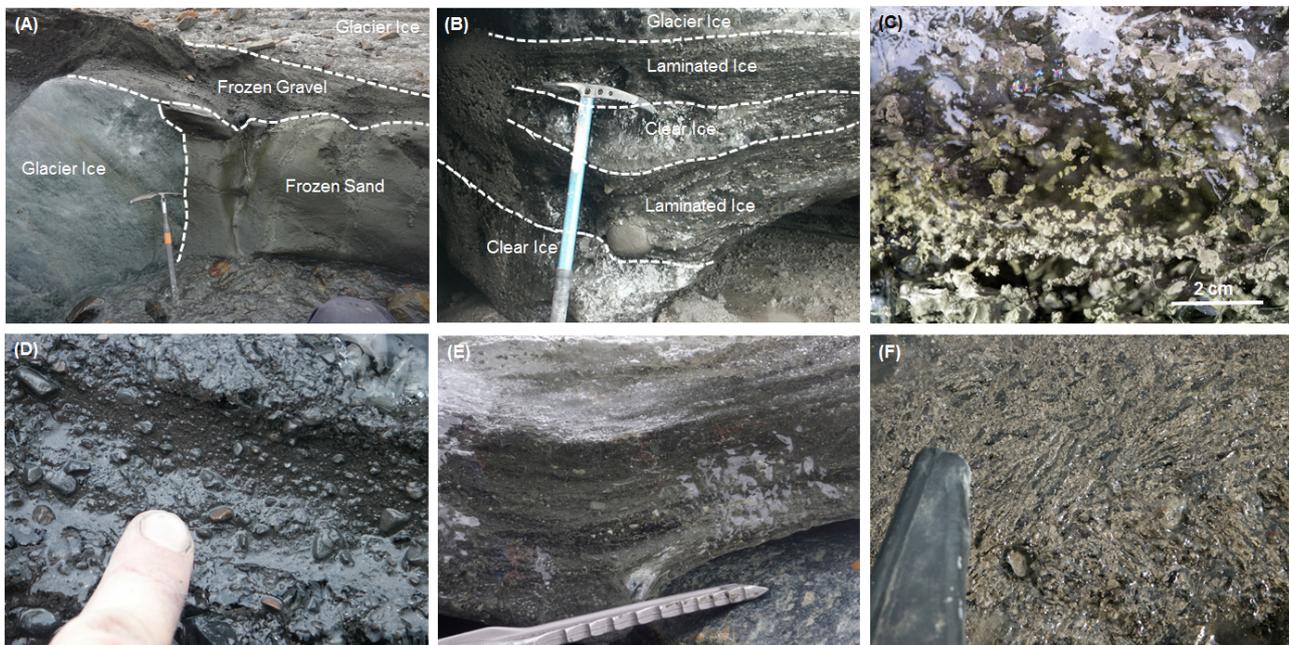


Figure 2: Examples of ice facies variability in section (a and b) and in detail (c to f). (a) Crevasse-fill of frozen sands and gravel outcropping at the surface of Storglaciären, Sweden; (b) Basal ice layer with alternating layers of clear ice and laminated ice at Svínafellsjökull, Iceland; (c) Basal ice with angular debris aggregates at Svínafellsjökull; (d) Basal ice composed of sediment with only interstitial ice at Svínafellsjökull; (e) Gradational zone between white, bubble-rich glacier ice above, and debris-rich, bubble-poor basal ice below, Storglaciären; (f) Distinctive arrangement of ice crystals and sediment associated with formation from supercooled water, Svínafellsjökull.

Whilst many physical characteristics can be estimated visually, further quantification of some properties is possible either in the field, or at a field base or camp. Ice crystal size can often be estimated and measured where ice surface melting accentuates the vein network between crystals. However, measurements can also be made by progressively thinning a small piece of ice on a hot pan lid, heated on a field stove. Once the ice is ~1 mm thick, it can be placed between polarizing filters, held to the light and photographed, ideally with a scale in the

image (Figure 3). Measurements of ice crystal size can then be made on a computer screen at a later date using image analysis software.

Debris volume measurements can be made by returning bagged samples to the field base, allowing them to melt, decanting them into measuring cylinders, and waiting for the sediment to settle, which can take variable amounts of time from hours to days depending on the grain size. In addition, in the case of very fine-grained sediment,

centrifuging samples could be necessary to force the deposition of colloidal-sized suspended sediment. The volume of sediment and water can be measured directly in the cylinder, but it is important to consider the density difference between water and ice; any water volume should be multiplied by 0.9 to allow calculation of the percentage volume of sediment in ice.



Figure 3: Example of a thin section of ice viewed through polarizing filters. (Photo: N. Midgley).

Clast shape, roundness and fabric analyses can also be conducted in the field (e.g. Benn, 2004). Typically, between 30 and 50 clasts per sample would be extracted for these analyses, although this can be particularly time-consuming when clasts must first be removed from the ice with an ice axe or chisel.

Once the ice has been described, it is common to develop a cryofacies classification or naming scheme, or to use a pre-existing classification scheme. The classification of glacier basal ice into distinct cryofacies was first applied by Lawson (1979) at the Matanuska Glacier, Alaska. However, since then, a range of individual classification schemes have been devised for individual studies or glaciers (e.g. Knight, 1987; Sharp *et al.*, 1994; Hubbard and Sharp, 1995; Christoffersen *et al.*, 2006; Cook *et al.*, 2007), which has led to the proliferation of multiple names for descriptively similar cryofacies. This confusing situation has led to attempts to unify existing basal ice facies classification schemes (e.g. Knight, 1994; Hubbard *et al.*, 2009). Given that basal glacier ice and permafrost are essentially the same materials

(a mix of ice and sediment), Waller *et al.* (2009) recommended greater collaboration between glaciologists and permafrost scientists who have been using permafrost classification schemes successfully for some decades (e.g. Murton and French, 1994; French and Shur, 2010). This remains a promising prospect, but has yet to be applied more widely.

The most recent attempt to develop a unified, non-genetic scheme for classifying and naming basal cryofacies was undertaken by Hubbard *et al.* (2009). Their scheme was based principally on the nature of layering and distribution of included sediment. The scheme comprises six primary cryofacies and twelve composite cryofacies (Table 1). Figure 4 illustrates a flow chart of how to classify and name basal cryofacies according to this scheme.

Overall, we recommend either using a unified classification scheme, such as that of Hubbard *et al.* (2009) or, where researchers prefer to develop their own schemes, that comparisons be drawn with cryofacies descriptions in existing studies so that broad similarities and differences can be drawn out between glaciers. Certainly, it would be worth pursuing the suggestion of Waller *et al.* (2009) for closer ties between glacial and permafrost researchers, and the development of common practices between these two sub-disciplines of cryospheric science.

Once described in the field, the documentation of the presence of cryofacies is best achieved through stratigraphic logging, section drawing, or both. Stratigraphic logging involves determining the thickness of layers within a cleaned section, the nature of each of those layers (as discussed above), and the nature of the contacts between individual layers (e.g. sharp, gradational, etc.). An introduction to stratigraphic logging can be found in Evans and Benn (2004). A measuring tape is usually fixed vertically up the cleaned section to provide a reference scale from which to log. Similar principles can be applied on the glacier surface where, for example, there may be layering and changes in ice types and structures along a transect.

Table 1: Basal cryofacies classification and naming scheme of Hubbard *et al.* (2009).

Layered			Uniform		
Typical thickness of primary component layers			Typical debris concentration of basal cryofacies		
Metres or longer	Centimetres to decimetres	Millimetres or shorter	Solid debris	Scattered debris	No debris
Stratified (St)	Banded (B)	Laminated (L)	Solid (So)	Dispersed (D)	Clean (C)

Predominantly layered

Stratified solid (StSo)	Banded solid (BSo)	Laminated solid (LSo)
Stratified dispersed (StD)	Banded dispersed (BD)	Laminated dispersed (LD)

Predominantly uniform

Solid stratified (SoSt)	Dispersed stratified (DSt)
Solid banded (SoB)	Dispersed banded (DB)
Solid laminated (SoL)	Dispersed laminated (DL)

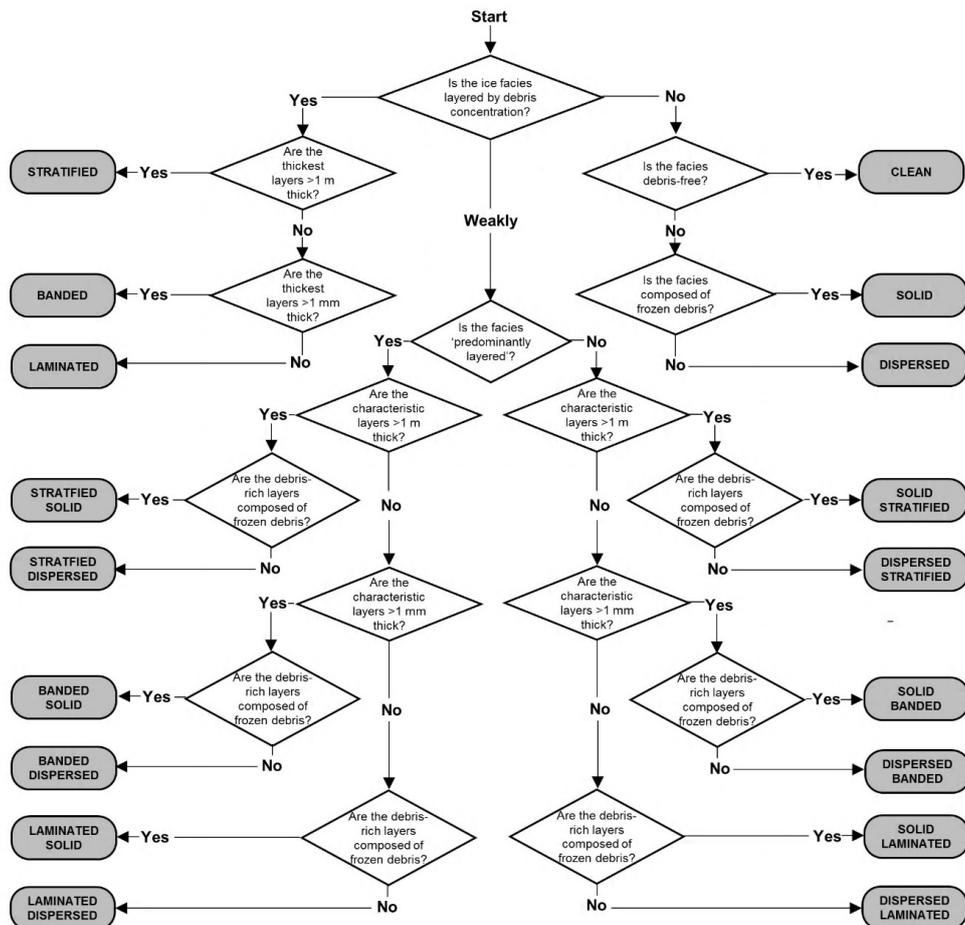


Figure 4: Flow-chart illustrating how to name and classify basal cryofacies (re-drawn from Hubbard *et al.*, 2009).

Sampling ice

Once the ice has been described and logged, consideration should be given to representative sampling. It is common to extract target samples for a variety of further analyses in the field and laboratory such as quantification of debris content, particle size analysis of included sediment, ice crystal size and fabric, geochemistry, stable isotope composition, and cell concentration and extraction of DNA for microbiological studies. The purpose of the analysis will determine the extraction process, since extra care must be taken for some sensitive analytical techniques (e.g. chemical or biological) in order to avoid cross-contamination. Before ice samples are extracted, it is important to remove any ice that may be contaminated with surface water or debris. This is particularly important at the glacier margin and surface of the ablation zone, where several decimetres thickness of ice may be affected by the penetration of surface water and sediment migrating along fractures and intra-crystalline veins. In most cases, removal of the top 20-30 cm of surface ice with an ice axe is sufficient to avoid surface contamination issues.

The choice of tool for ice sample extraction depends on the size of sample required, the nature of the ice being extracted, the need for aseptic microbiological sampling, and the availability of equipment. Where large ice samples are required (e.g. for return of decimetre-size ice blocks to a laboratory for crystal fabric analysis), or for the opening of a channel or tunnel, a chain saw may be the most effective tool (e.g. Fitzsimons *et al.*, 1999; Hubbard and Glasser, 2005). However, as Hubbard and Glasser (2005) note, chainsaw injuries can be extremely serious, and fieldwork is often undertaken far from medical facilities. Hence, extreme care should be taken, and the operator should be trained and properly equipped (i.e. with hardhat, visor or goggles, ear defenders, protective gloves, protective trousers and protective boots) (Hubbard and Glasser, 2005). Other issues include the rapid wearing of chainsaw teeth when cutting through sediment-rich ice, and that chainsaw blades can freeze solid into ice sections at very cold temperatures if left for even relatively short periods of time.

Perhaps the most common tool used for ice sample extraction is the ice axe (Figure 5a). For more precise sample removal, it may be better to use a hammer (or cobble) and chisel (Figure 5b) once the unwanted surface ice has been removed using an ice axe. For even greater precision, and usually much smaller samples, an ice screw can be used (Figure 5c), although for very debris-rich ice or for ice containing clasts, this can be ineffective.



Figure 5: Ice sampling techniques using (a) an ice axe; (b) a chisel; and (c) an ice screw.

Whichever extraction technique is used, it is important to consider contamination issues. For samples that are used to characterize the physical composition of the ice (e.g. sediment content, ice crystal size, bubble content), it is reasonable to expect that surface contamination will have been diluted significantly or completely by the point where the rotten surface ice has been removed. Hence, the same tool can continue to be used to extract the sample. To be certain of low contamination, a different, clean tool could be used for the final sample extraction. For ice geochemistry or microbiology samples, it is desirable to minimise contamination by using clean, sterile equipment. If microbiological samples are being collected, then metal sampling implements should be cleaned and ideally sterilized, typically achieved by flaming with the aid of ethanol (e.g. Skidmore *et al.*, 2005). Note, however, that it is impractical to keep tools sterile during sampling, because the tool will be used to penetrate contaminated surface layers. It is therefore necessary to work in a planned way that does not rely on tool sterility for maintaining sample integrity. Laboratory or medical gloves are useful for minimizing the introduction of salt from sweat, or of microorganisms from skin. For high sensitivity chemical and microbiological applications a useful technique is to chip ice fragments directly into a sample bag without touching by hand or tool. This can be achieved by chiseling extensively around the feature of interest until it can be released by fracturing the back of the sample, ideally with a single chisel blow. Whilst delivering the releasing blow, a sample bag is held firmly inside-out over the feature by an assistant who captures the sample.

It is vital that sample bags and vessels be labelled clearly with a permanent marker and, where samples are to be transported back to a laboratory, it is advised that the sample name be labelled in several locations on the bag or vessel as labels can wear off in transit. Aside from the sample name, it is advised that, space permitting, other useful information, such as GPS coordinates, sample type (e.g. cryofacies), and individual identifying code be recorded on the bag and vessel, as well as in a field notebook. Remember also to mark on the section diagram or log where the sample was taken,

and take a photograph. To avoid cross-contamination and sample loss due to leakage, it is worthwhile double- or triple-bagging the ice samples, since ice can have sharp edges capable of piercing bags.

Several laboratory data analysis techniques require separation of solid and liquid components of the samples, or may require the sample to be returned frozen (e.g. sediment particle size analysis, stable isotope analysis, etc.). Filtering of samples can be achieved easily in a field base or camp. The simplest set-up is the use of filter papers in a funnel to collect sediment (e.g. for laboratory particle size analysis). Perhaps the most effective technique is to use a vacuum pump such that both the sediment and clean water samples can be collected at the same time. Fine filters tend to clog easily with fine sediment, so this can be a time-consuming process for debris-rich samples. Water can be decanted from the collection chamber and bottled as required (e.g. in 8 ml Nalgene screw-cap bottles).

Concluding remarks

The description and sampling of ice is essential for many aspects of the cryospheric sciences. For example, documenting and measuring ice properties can provide vital information about the origin and geomorphological significance of cryofacies, the rheology of ice and its impact on glacier flow, and the potential for ice to serve as a microbial habitat. Much can be learned from visual description and classification of ice in the field, although ice samples are often required for more detailed physical, chemical and biological analyses. A carefully thought-out sampling strategy is important for the success of any campaign involving the description of ice and retrieval of ice samples, and is essential for estimating the appropriate type and amount of sampling equipment. However, the dynamic nature of ice margins and surfaces means that flexibility is required to account for access and safety issues.

Acknowledgments

MTR is in receipt of a Faculty of Science and Engineering PhD Studentship at Manchester Metropolitan University, and gratefully acknowledges a Geographical Club Award

for supporting basal ice description and sampling fieldwork in Iceland. SJC thanks Manchester Metropolitan University for a Research Accelerator award for recent fieldwork in Iceland. We thank David Graham, an anonymous reviewer and Lucy Clarke for their constructive comments that improved the quality of this chapter.

References

- Benn DI. 2004. Clast morphology. In: Evans DJA, Benn DI. (Eds.) *A Practical Guide to the Study of Glacial Sediments*. Arnold, London. DOI: 10.4324/9780203783481
- Chandler D, Hubbard B, Hubbard A, Murray T, Rippin D. 2008. Optimising ice flow law parameters using borehole deformation measurements and numerical modelling. *Geophysical Research Letters*. **35**: L12502. DOI: 10.1029/2008GL033801
- Christoffersen P, Tulaczyk S, Carsey F, Behar AE. 2006. A quantitative framework for interpretation of basal ice facies formed by ice accretion over subglacial sediment. *Journal of Geophysical Research*. **111**: F01017. DOI: 10.1029/2005JF000363
- Cook SJ, Swift DA, Graham DJ, Midgley NG. 2011a. Origin and significance of 'dispersed facies' basal ice: Svínafellsjökull, Iceland. *Journal of Glaciology*. **57**: 710-720. DOI: 10.3189/002214311797409703
- Cook SJ, Graham DJ, Swift DA, Midgley NG, Adam WG. 2011b. Sedimentary signatures of basal ice formation and their preservation in ice-marginal sediments. *Geomorphology*: **125**: 122–131. DOI: 10.1016/j.geomorph.2010.08.018
- Cook SJ, Robinson ZP, Fairchild IJ, Knight PG, Waller RI, Boomer I. 2010. Role of glaciohydraulic supercooling in the formation of stratified facies basal ice: Svínafellsjökull and Skaftafellsjökull, southeast Iceland. *Boreas*. **39**: 24–38. DOI: 10.1111/j.1502-3885.2009.00112.x
- Cook SJ, Knight PG, Waller RI, Robinson ZP, Adam WG. 2007. The geography of basal ice and its relationship to glaciohydraulic supercooling: Svínafellsjökull, southeast Iceland. *Quaternary Science Reviews*. **26**: 2309–2315. DOI: 10.1016/j.quascirev.2007.07.010
- Evans DJA, Benn DI. 2004. Facies description and the logging of sedimentary exposures. In: Evans DJA, Benn DI (Eds.) *A Practical Guide to the Study of Glacial Sediments*. Arnold, London.
- Fitzsimons SJ. 2006. Mechanical behaviour and structure of the debris-rich basal ice layer. In: Knight PG. (Ed.) *Glacier Science and Environmental Change*. Blackwell, Oxford, 329–335.
- Fitzsimons SJ, McManus KJ, Lorrain RD. 1999. Structure and strength of basal ice and substrate of a dry-based glacier: evidence for substrate deformation at sub-freezing temperatures. *Annals of Glaciology*. **28**: 236–240. DOI: 10.3189/172756499781821878
- French H, Shur Y. 2010. The principles of cryostratigraphy. *Earth-Science Reviews*. **101**: 190-206. DOI: 10.1016/j.earscirev.2010.04.002
- Hallet B, Hunter L, Bogen J. 1996. Rates of erosion and sediment evacuation by glaciers: a review of field data and their implications. *Global and Planetary Change*. **12**: 213–235. DOI: 10.1016/0921-8181(95)00021-6
- Hambrey MJ, Lawson W. 2000. Structural styles and deformation fields in glaciers: a review. In: Maltman AJ, Hubbard B, Hambrey MJ. (Eds.) *Deformation of Glacial Materials*. Geological Society, London, Special Publications. **176**: 59–83.
- Hodson AJ, Anesio AM, Tranter M, Fountain AG, Osborn M, Prisco J, Laybourn-Parry J, Sattler B. 2008. Glacial Ecosystems. *Ecological Monographs*. **78**: 41-67. DOI: 10.1890/07-0187.1
- Hood E, Battin TJ, Fellman J, O'Neel S, Spencer RGM. 2015. Storage and release of organic carbon from glaciers and ice sheets. *Nature Geoscience*, **8**(2) 91–96. DOI: 10.1038/ngeo2331
- Hubbard B, Glasser NF. 2005. *Field Techniques in Glaciology and Glacial Geomorphology*. John Wiley: Chichester. DOI: 10.1017/S0016756806232977
- Hubbard B, Sharp MJ. 1995. Basal ice facies and their formation in the Western Alps. *Arctic and Alpine Research*. **27**: 301–310. DOI: 10.2307/1552023
- Hubbard B, Sharp MJ. 1989. Basal ice formation and deformation: a review.

- Progress in Physical Geography*. **13**: 529–558. DOI: 10.1177/030913338901300403
- Hubbard B, Cook S, Coulson H. 2009. Basal ice facies: a review and unifying approach. *Quaternary Science Reviews*. **28**: 1956–1969. DOI:10.1016/j.quascirev.2009.03.005
- Knight PG. 1997. The basal ice layer of glaciers and ice sheets. *Quaternary Science Reviews*. **16**: 975–993. DOI: 10.1016/S0277-3791(97)00033-4
- Knight PG. 1994. Two-facies interpretation of the basal layer of the Greenland ice sheet contributes to a unified model of basal ice formation. *Geology*. **22**: 971–974. DOI: 10.1130/0091-7613(1994)022<0971:TFIOTB>2.3.CO;2
- Knight PG. 1987. Observations at the edge of the Greenland ice sheet: boundary condition implications for modellers. *International Association of Hydrological Sciences Publication*. **170**: 359–366.
- Lawson DE. 1979. Sedimentological Analysis of the Western Terminus Region of the Matanuska Glacier, Alaska. *U.S. Army, Cold Regions Research and Engineering Laboratory*. Hanover, New Hampshire.
- Lovell H, Fleming EJ, Benn DI, Hubbard B, Lukas S, Naegeli K. 2015. Former dynamic behaviour of a cold-based valley glacier on Svalbard revealed by basal ice and structural glaciology investigations *Journal of Glaciology*. **61**: 309–328. DOI: 10.3189/2015JoG14J120
- Miall AD. 2016. *Stratigraphy: A Modern Synthesis*. Springer International Publishing, Switzerland.
- Montross SM, Skidmore M, Christner B, Samyn D, Tison J-L, Lorrain R, Doyle S, Fitzsimons S. 2014. Debris-Rich Basal Ice as a Microbial Habitat, Taylor Glacier, Antarctica. *Geomicrobiology Journal*. **31**: 76–81. DOI: 10.1080/01490451.2013.811316
- Murton JB, French HM. 1994. Cryostructures in permafrost, Tuktoyaktuk coastlands, western Arctic Canada. *Canadian Journal of Earth Sciences*. **31**: 737–747. DOI: 10.1139/e94-067
- Nichols G. 2009. *Sedimentology and Stratigraphy*. Wiley-Blackwell: Chichester.
- Sharp MJ, Jouzel J, Hubbard B, Lawson W. 1994. The character, structure and origin of the basal ice layer of a surge-type glacier. *Journal of Glaciology*. **40**: 327–340. DOI: 10.1002/9780470750636.ch65
- Skidmore M, Anderson SP, Sharp M, Foght J, Lanoil BD. 2005. Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. *Applied and Environmental Microbiology*. **71**: 6986–6997. DOI: 10.1128/AEM.71.11.6986-6997.2005
- Waller RI, Murton JB, Knight PG. 2009. Basal glacier ice and massive ground ice: different scientists, same science? *Geological Society, London, Special Publications*, **320**(1): 57–69. DOI: 10.1144/SP320.

Tracer investigations

Catriona L. Fyffe¹

¹ School of the Environment, University of Dundee (catrionalfyffe@live.com)



ABSTRACT: Tracers can be used to investigate the morphology and evolution of the drainage system that exists within and beneath glaciers. This paper explains how to use fluorescent dyes as tracers within studies of the glacier-hydrological system, detailing the type and quantity of dye required, the use and calibration of fluorometers used to detect the dye, and guidance on conducting tracer experiments. It then goes on to describe the parameters that can be calculated from collected dye breakthrough curves and their interpretation in terms of the likely morphology of the drainage system.

KEYWORDS: glacier-hydrology, dye tracing, fluorometers, rhodamine

Introduction

Investigating the structure and evolution of the drainage system within and beneath glaciers is a challenge because of its inaccessibility. Although some workers (Gulley 2009; Benn et al., 2009; Gulley et al., 2012) have been able to enter englacial and subglacial channels directly, elucidating the form of the glacier-hydrological system is usually performed by injecting fluorescent dyes into moulins or crevasses on the surface of the glacier and detecting the resulting dye return curve at a monitoring site on the proglacial stream. In this paper the methods used to conduct and analyse dye return curves are explained.

At first, understanding the structure of the glacier hydrological system was achieved by tracing injections of sodium chloride (Stenborg, 1969). More recently, fluorescent dyes have become the tracer of choice for glacier hydrology studies due to their greater detectability and the smaller quantities required. If a quantity of dye is injected into a moulin or crevasse, and then detected downstream (to give the dye breakthrough or return curve), the time of travel of the dye within the system, the spread of the dye breakthrough curve and the percentage of dye detected can be calculated. These trace characteristics can then be used to determine the morphology of the glacier-hydrological

drainage system at the time of the test. In this paper the morphology of the system refers to the 'size, shape and roughness' of the drainage system (Willis et al., 2012) (characteristics which determine the efficiency of the system) and the structure of the hydrological system is defined as the 'location, alignment and interconnection' of the drainage routes (Willis et al., 2012).

Dye tracing can therefore be used to determine the drainage structure in terms of identifying the moulins which drain to a particular stream, and can determine changes in the drainage morphology between different parts of a glacier (e.g. the eastern and western sides of Midtdalsbreen, Norway, Willis et al. (1990, 2012)). It can also be used to identify changes in the morphology of the drainage system over the season and how this relates to the position of the snowline (Nienow et al., 1998); and changes in the relationship between the water velocity and discharge within a day, termed velocity-discharge hysteresis (Nienow et al., 1996; Schuler and Fisher, 2009). These studies on alpine glaciers have led to larger-scale studies of the drainage system beneath an outlet glacier of the Greenland ice sheet using both rhodamine dye (Cowton et al., 2013) and sulphur hexafluoride, a gas suitable for tracing over long distances (Chandler et al., 2013).

Instruments and dye

Types of dye

Fluorescence occurs when a substance absorbs electromagnetic radiation at the precise frequency needed to excite its electrons to a higher energy. As the electrons return to their lower energy state, this energy is emitted as light which typically has a longer wavelength than that absorbed. Fluorescent dyes used for tracing transform a high proportion of the absorbed energy to emitted energy, and can be detectable at very low concentrations (e.g. to a minimum of 0.01 ppb rhodamine with a Turner Designs fluorometer (Turner Designs, 2010)). Dyes used for tracing include fluorescein, tinopal CBS-X, rhodamine B powder and rhodamine WT liquid. The rhodamine dyes are most commonly used, but fluorescein or tinopal can be used alongside them if traces are to be conducted simultaneously, since they emit light at different wavelengths to those of the rhodamine dyes (Nienow et al., 1998; Fountain, 1993). Care should be taken not to ingest dye and gloves should be worn to avoid skin contact. Consult the manufacturer's data sheet for details of safety considerations, and follow local guidelines if

dye is to be injected into drinking water supplies or environmentally sensitive areas.

Quantities to use

The quantity of dye required to produce a detectable dye breakthrough curve depends on the discharge of the stream at the monitoring site, the distance to be traced, and importantly the morphology of the drainage system beneath the glacier. The quantity of dye required to trace a distributed system can be an order of magnitude larger than that needed to trace a channelized system. Caution is therefore recommended in the quantity of dye used initially, to prevent the dye being visible at the monitoring site. To estimate the quantity of dye required for tracing in open water streams Equation 1 can be used:

$$V_i = 3.14 \times 10^{-4} \left(\frac{0.305 Q_{max} d}{0.045 u} \right)^{0.94} C_p, \quad (1)$$

where V_i is the volume of 21% rhodamine WT (l), Q_{max} is the maximum stream discharge at the downstream site ($m^3 s^{-1}$), d is the distance to the downstream site (km), u is the mean stream velocity ($m s^{-1}$) and C_p is the peak concentration at the downstream sampling

Table 1: Quantities of dye used by selected investigators. RB is 100% rhodamine B powder, RWT is 21% rhodamine WT liquid, F is fluorescein and T is tinopal CBS-X. Quantities of rhodamine B in grams can be converted to an equivalent of rhodamine WT in ml by multiplying by 4.76. The minimum to maximum quantities used are stated. Proglacial discharge is generally an average between the time of injection and trace peak.

Investigator	Glacier	Distance traced (m)	Proglacial discharge ($m^3 s^{-1}$)	Type of dye	Quantity of dye (g for RB and F, ml for RWT and T)	Time of year
Willis et al., 1990	Midtdalsbreen, Norway	1100-1650	0.2-3.3	RB	140-300	June to August
Willis et al., 2009	Brewster Glacier, New Zealand	~500 -1500	-	RWT	50-70	January to March
Nienow, 1993 (1990 data)	Haut Glacier, d'Arolla, Switzerland	1028-1426	0.5-4.1	RB	30-75	June
		1028-1271	0.6-3.3	F	100-200	June
		999-3300	2.0-6.3	RB	10-400	July
		670-2378	2.1-6.4	F	100-250	July
Fyffe, 2012	Miage Glacier, Italian Alps	997-5867	1.7-10.7	RWT	40-280	June to September
Cowton et al., 2013	Leverett Glacier, Greenland Ice Sheet	1250	<1-300	RWT/RB	300-2000/ 250-1000	April to August
		3600	15-400	RWT/RB	4000-5000/ 750-3000	May to August
		6600	15-400	RWT/RB	9000-10000/ 1500-5000	May to August
Hasnain et al., 2001	Dokriani Glacier, India	930-2300	9-30	RWT	40-50	July to September
Fountain, 1993	South Cascade Glacier, Washington State, USA	485-3335	0.2-0.8	RWT	70-1000	July to September
		485-1355	0.2-0.8	T	1000-3700	July to September
Seaberg et al., 1988	Storglaciären, Sweden	~900-1000	0.1-0.7	RWT	102-208	June to August

site (ppb) (adjusted to SI units from Kilpatrick and Wilson (1989)). The quantity of dye used by other investigators can also be used as a guide, see Table 1.

Measuring fluorescence

A fluorometer is used to detect the dye return curve when it reaches the proglacial stream. Fluorometers supply a beam of radiation at the excitation wavelength, and measure the intensity of light emitted by the fluorescent dye, which is proportional to the concentration of dye in solution.

Two common brands of modern fluorometer are the Seapoint rhodamine fluorometer and Turner Cyclops-7 fluorometer, both of which can be set up to log to a Campbell Scientific data logger or similar. Values should ideally be recorded every 1 to 10 seconds, with data outputted as an average for each minute, especially if sharply peaked return curves are expected, although a lower logging interval is suitable for broader return curves or after the peak has passed through. Measurements should begin well before the trace comes through (preferably before the dye is injected) and continue until well after the fluorescence has returned to background levels. This can be over a day if a less efficient drainage system is traced, or only a few hours if the system is efficient and the trace is conducted close to the proglacial stream outlet.

To position the fluorometer in the stream it can be attached to a length of angle iron (e.g. dexion), which is then secured to the bank and/or river bed (e.g. using expansion bolts, rope or boulders). The fluorometer should be positioned low enough in the water to reduce the influence of sunlight on readings, although using the fluorometer's shade cap also reduces this effect. Note that the fluorescence of dyes can be influenced by the stream pH (the fluorescence of rhodamine WT is stable within a pH range of 5.5 to 11 (Keystone Aniline Corporation, 2002)).

Calibrating a fluorometer

Fluorometers give relative values of fluorescence intensity, which is logged as a value in volts. To convert this to dye concentration the fluorometer must be calibrated, preferably for each dye lot (as

there may be slight differences between batches). The calibration procedure allows the definition of the relationship between the dye concentration and the voltage measured by the fluorometer. The value in volts should vary linearly with dye concentration, apart from at very high or low concentrations. It is important to calibrate the fluorometer in the field to ensure the water temperature is similar to that during a trace, because water temperature is the most significant factor that varies the relationship between fluorescence and dye concentration (Wilson et al., 1986). Furthermore, chlorinated tap water should not be used for calibration as chlorine affects the fluorescence of rhodamine dyes (Wilson et al., 1986). Suspended sediment in the proglacial stream can fluoresce within the same wavelength band as the dye. This means the measured background fluorescence is positive and can be variable (Hubbard and Glasser, 2005), although this is most pronounced using fluorescein (Nienow et al., 1998). The value in volts used in the calibration must therefore be corrected to remove the influence of the background fluorescence.

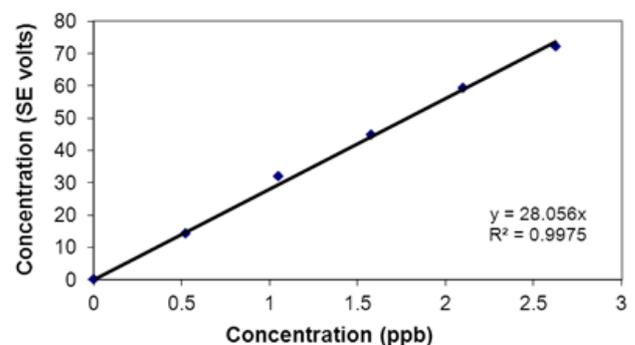


Figure 1: Calibration curve giving the relationship between the voltage measured by the fluorometer and the dye concentration.

Fluorometer calibration involves recording the voltage measured by the fluorometer while it is in solutions with a known dye concentration. Around 10 standards of different dye concentrations should be measured, spanning the measureable range of the fluorometer, and with an emphasis on lower concentrations in case of non-linearity in the calibration relationship. An example calibration procedure for 21% rhodamine WT is as follows:

1. Have ready two completely clean buckets pre-marked with a 5 litre

- level, dye, a μl pipette, and have the fluorometer logging
2. Fill both buckets with 5 litres of stream water, and add the fluorometer to the second bucket
 3. To define the background voltage, record several logged values of the fluorometer voltage (with logged values ideally averages of several readings)
 4. Add 250 μl of dye to the first bucket, and mix it thoroughly
 5. Replace the pipette tip, and then take 250 μl from the first bucket and add this to the second bucket, mixing thoroughly (this gives a concentration in the second bucket of 0.525 ppb rhodamine WT)
 6. Take several readings of the voltage within the second bucket
 7. Replace the pipette tip or wash it out in stream water repeatedly
 8. Add a further 250 μl of the first bucket solution into the second bucket, and mix thoroughly (to give a concentration of 1.05 ppb rhodamine WT)
 9. Repeat steps 6. to 8. until the desired maximum voltage has been reached.

The average background voltage should then be removed from the voltage measurements, with these values plotted against the dye concentration in the second bucket, to give a relationship between voltage and dye concentration (Figure 1). Then, once the background voltage specific to the dye breakthrough curve has been determined (usually by averaging values measured prior to the trace injection), and then subtracted from the logged return curve, the dye concentration of the return curve can be calculated. The background voltage may not be constant (as shown in Figure 3) but can vary over time, necessitating the use of a varying background correction.

Conducting dye trace experiments

Dye should be injected into a supraglacial stream that flows directly into a freely draining crevasse or moulin (Figure 2). Care must be taken to use channels which are ice walled and free of snow (snow should be

removed if it is covering a moulin), and dye should not be injected into standing water. The bottle containing the dye should be well flushed with stream water so that all of the dye enters the stream. When working on a snow-covered glacier (especially in regions of crevasses or moulins), all field members should be roped-up and know what to do in the event of a fall. Consult a qualified professional if unsure.

If the purpose of the study is to determine the seasonal evolution of the glacier-hydrological system, it is advised to conduct dye traces into moulins at a variety of distances from the proglacial stream outlet. Moulins at increasingly higher elevations will become traceable as the season progresses and the snowcover melts. Traces into the selected moulins can be repeated over the season to identify the changes in the resulting return curves. In these cases traces should be made at a similar time each day (e.g. between 11:00 and 15:00, (Willis et al., 2009)), to reduce the influence of diurnal hydrological fluctuations. Conversely, if velocity-discharge hysteresis is the object of the study then traces should be injected into the same moulin at intervals throughout one day (Nienow et al., 1996).



Figure 2: Injecting rhodamine WT liquid dye into a moulin on Miage Glacier, Italian Alps.

Dye trace parameters

Once the dye breakthrough curve values have been converted into dye concentration several parameters can be calculated that describe the shape of the curve and the speed of the dye trace as it travels through the glacier. It is not correct to assume that the values derived will be applicable along the entire length of the drainage network – the channel morphology and water velocity may vary spatially. Note that for broad traces, especially with multiple peaks, the derivation of the parameters described below can be difficult to determine and not accurate (e.g. the difficulty in defining the start and end of a broad return curve such as that shown in Figure 5).

Nevertheless, to analyse a breakthrough curve first find the time to peak (t_m), and time to half the peak on the rising (t_1) and falling (t_2) limb of the breakthrough curve (s). From the time to peak the average trace velocity can be calculated using Equation 2:

$$u = \frac{d}{t_m}, \quad (2)$$

where u is the trace velocity (m s^{-1}) and d is the straight line distance to the injection site (m) (Seaberg et al., 1988, p 222). The value calculated in Equation 2 gives the minimum velocity of the water because the actual distance travelled by the water will be greater than measured due to stream sinuosity.

The dispersion coefficient (D) of the breakthrough curve can be determined (as this is “a measure of the rate at which the dye-concentration peak broadens relative to the rate at which it is transferred through the glacier” (Willis et al., 1990, p93)) using:

$$D = \frac{d^2(t_m - t_i)^2}{4t_m^2 t_i \ln \left[2 \left(\frac{t_m}{t_i} \right)^{\frac{1}{2}} \right]}, \quad (3)$$

(Seaberg et al., 1988, p222) where t_i equals the time at which the dye concentration is half of the peak, either on the rising or falling limb of the dye return curve. The equation is solved iteratively for t_m , allowing D to be found. To do this complete Equation 3 for t_i on the rising limb and subtract Equation 3 completed with t_i on the falling limb, then vary t_m in both variations of the equation until the difference between the equations approaches

zero. At this point take the value of t_m and use it to calculate D for both values of t_i . As the dispersion of the tracer is proportional to its velocity (u), the constant of this proportionality (or the dispersivity (b)) can be calculated (Equation 4).

$$D = bu, \quad (4)$$

(Seaberg et al., 1988, p224). If the drainage system has interlinking passages then the dispersivity is a measure of the length of the passages, so a large dispersivity points to a more distributed network.

The apparent mean cross sectional area (A_{sm} , m^2) of the channel network can also be calculated (Equation 5). If repeat traces are conducted into the same moulin then this can be used to identify changes in the drainage system morphology from a broad, low, distributed system with a large cross-sectional area to a discrete channelized system.

$$A_{sm} = \frac{Q}{u}, \quad (5)$$

where Q is the mean discharge between the injection and detection point, calculated as the average of the supraglacial and proglacial stream discharge over the time of the test (from injection until the end of the trace return curve) (Nienow et al., 1998, p 828). Using the mean discharge decreases the possibility of overestimating the cross sectional area.

To find the volume of dye recovered (V_r , ml) the equation used to calculate the discharge of a stream from the slug-injection method (Kilpatrick and Cobb, 1985, p6) can be rearranged (Equation 6):

$$V_r = \frac{S^{-1} \left(\frac{1}{1.649 \times 10^{-8}} (Q_p A_c) \right)}{C_i}, \quad (6)$$

where S is the specific gravity of the dye used (1.15 for rhodamine WT), Q_p is the average proglacial discharge during the time taken for the dye return curve to pass through ($\text{m}^3 \text{s}^{-1}$), A_c is the area under the dye curve in ppb minute^{-1} , and C_i is the concentration of the dye prior to injection (ppb). A_c can be calculated by summing the positive dye concentrations composing the return curve (once the background has been removed) and multiplying by the logging interval in minutes. Sorption of rhodamine B onto

sediment can reduce the quantity recovered (Hubbard and Glasser, 2005), although rhodamine WT is more resistant to adsorption (Smart and Laidlaw, 1977).

The relationship between discharge and velocity can be used to infer whether the channel is full and pressurised (closed channel conditions) or only partially full and at atmospheric pressure (open channel conditions). These relationships are only valid over short periods (within a diurnal cycle) since changes to the cross-sectional area through melting of the channel sides or ice deformation are less likely to be significant. The relationship is found by fitting a power function (Equation 7) to several measurements of channel velocity at different discharges:

$$u = kQ^m \quad (7)$$

(Nienow et al., 1996, p 1413), where m relates to the conditions in the channel and k is a constant. When the channel is full, discharge can only increase via an increase in velocity (assuming no change in channel dimensions), so velocity and discharge are directly related and $m = 1$. Under open channel conditions, the channel cross-sectional area can increase as well as the velocity, so that $m < 1$. However the application of power functions to determine channel conditions can be erroneous because the tributary moulin and main channel discharge can vary out of phase, or water can be forced to back up in the moulin or tributary channel, decreasing the trace velocity (Nienow et al., 1996).

Interpreting traces

To determine the morphology of the drainage system from dye return curves it is necessary to know the likely trace parameters associated with different channel types. Generally, traces conducted into a system which is efficient and composed of discrete channels will consist of a singular peak. The trace will have a fast velocity (typically $0.3\text{--}0.5\text{ ms}^{-1}$ in small Alpine glaciers (Nienow et al., 1998)), a low dispersion coefficient and dispersivity (typically less than 10 m (Nienow et al., 1998)), and a high percentage dye recovery (Figure 3).

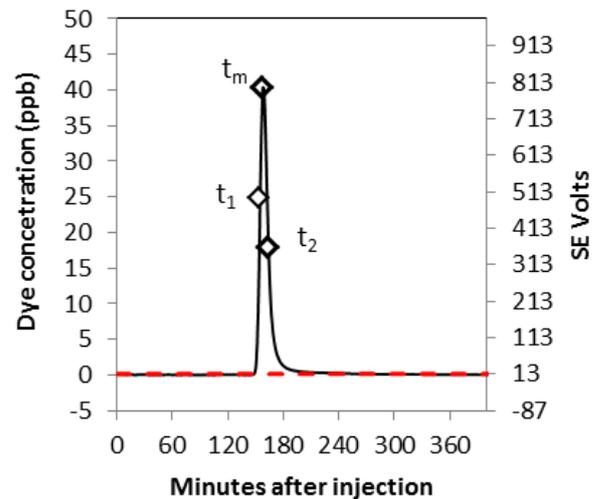


Figure 3: An example of a breakthrough curve indicative of a channelized drainage system, with t_1 , t_m and t_2 shown as diamonds (t_1 and t_2 do not have a concentration that is exactly half of t_m because of the rapid change in dye concentration). The raw SE volts value from the fluorometer is given on the secondary axis, with the dashed red line the background that was removed before the calibration was applied to give the dye concentration.

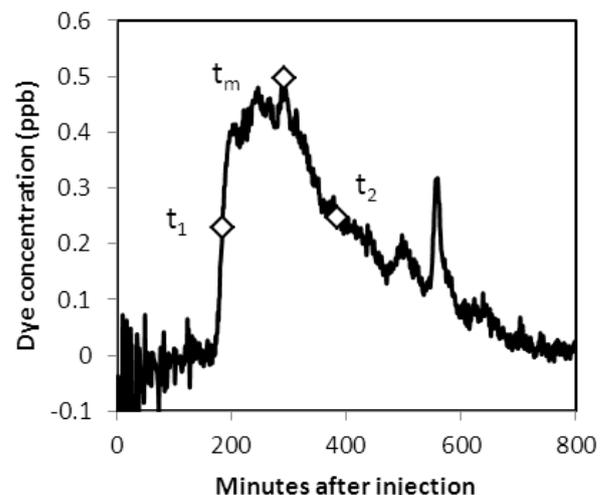


Figure 4: An example of a breakthrough curve with a shoulder on the falling limb, with diamonds showing the position of t_1 , t_m and t_2 .

The slower the trace and the greater the dispersion coefficient and dispersivity, the less efficient the drainage system. However, this may not necessarily imply that the drainage system is distributed, but may indicate that water has become trapped in part of the system (likely during high flows if the input discharge increases at a rate greater than the channel can expand by melting of the channel sides); or that the

conduit roughness is relatively large due to small input discharges (Gulley et al., 2012). Return curves with a pronounced shoulder on the falling limb may indicate storage of water in part of the system (Figure 4).

Traces with a very slow velocity (typically around 0.025 ms^{-1} (Nienow et al., 1998)), especially with multiple peaks, are indicative of a distributed drainage system, usually thought of in terms of a linked cavity network (Figure 5). This type of system is inefficient and gives broad breakthrough curves with low percentage dye returns (signifying that either water has become stored subglacially or that the dye has been returned at too low a concentration to be detected).

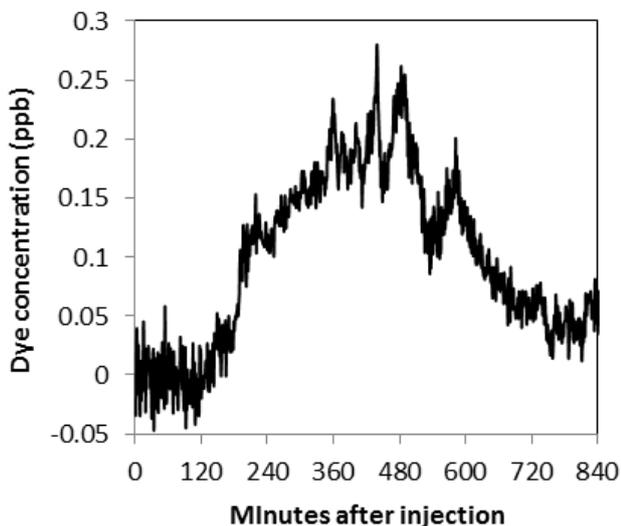


Figure 5: An example of a multi-peaked breakthrough curve indicative of a distributed system. This curve could be seen as just variation in the background fluorescence but subsequent traces into the same moulin confirmed this was a breakthrough curve.

Conclusions

Fluorescent dyes can give an insight into the morphology and seasonal evolution of a glacier's hydrological system. Rhodamine dyes tend to be preferred and the quantity required for a clear breakthrough curve is determined by the morphology of the drainage system, the distance traced and the proglacial discharge. Modern fluorimeters can be installed at the proglacial stream outlet and used to detect the dye breakthrough curve at a high temporal resolution. The resulting return curve can be described using several parameters, which

along with the return curve shape, can be used to identify the likely efficiency and morphology of the drainage system.

Acknowledgements

The author would like to thank the School of the Environment, University of Dundee for her studentship. Thanks to Joe Holden and Ben Brock for the loan of fluorimeters, and to Pete Nienow and Tom Cowton for information on dye quantities. Helpful comments from David Chandler and Ian Willis greatly improved the manuscript.

References

- Benn DI, Gulley J, Luckman A, Adamek A, Glowacki PS. 2009. Englacial drainage systems formed by hydrologically driven crevasse propagation. *Journal of Glaciology* **55**(191): 513-522.
- Chandler DM, Wadham JL, Lis GP, Cowton T, Sole A, Bartholomew I, Telling J, Nienow P, Bagsaw EB, Mair D, Vinen S, Hubbard A. 2013. Evolution of the subglacial drainage system beneath the Greenland Ice Sheet revealed by tracers. *Nature Geoscience* **6**(3): 195-198.
- Cowton T, Nienow P, Sole A, Wadham J, Lis G, Bartholomew I, Mair D, Chandler D. 2013. Evolution of drainage system morphology at a land-terminating Greenlandic outlet glacier. *Journal of Geophysical Research* **118**: 1-13.
- Fountain AG. 1993. Geometry and flow conditions of subglacial water at South Cascade Glacier, Washington State, U.S.A.; an analysis of tracer injections. *Journal of Glaciology* **39**(131): 143-156.
- Fyffe, CL. 2012. *The hydrology of debris-covered glaciers*. Unpublished PhD Thesis: School of the Environment, University of Dundee, UK.
- Gulley J. 2009. Structural control of englacial conduits in the temperate Matanuska Glacier, Alaska, USA. *Journal of Glaciology* **55**(192): 681-690.
- Gulley J, Walthard P, Martin J, Banwell AF, Benn D, and Catania GA. 2012. Conduit roughness and dye-trace breakthrough curves: why slow velocity and high dispersivity may not reflect flow in distributed systems. *Journal of Glaciology* **58**(2011): 915-925.

- Hasnain SI, Jose PG, Ahmad S, Negi DC. 2001. Character of the subglacial drainage system in the ablation area of Dokriani glacier, India, as revealed by dye-tracer studies *Journal of Hydrology* **248**: 216-223.
- Hubbard B, Glasser N. 2005. *Field Techniques in Glaciology and Glacial Geomorphology*. John Wiley and Sons Ltd: Chichester.
- Keystone Aniline Corporation. 2002. *Keyacid rhodamine WT Liquid*. Keystone Aniline Corporation: Chicago.
- Kilpatrick FA, Cobb ED. 1985. *Measurement of discharge using tracers*. United States Government Printing Office: Washington.
- Kilpatrick FA, Wilson JF. 1989. *Measurement of time of travel in streams by dye tracing*. United States Government Printing Office: Washington.
- Nienow PW. 1993. *Dye tracer investigations of glacier hydrological systems*. Unpublished PhD thesis: Cambridge University, UK.
- Nienow PW, Sharp M, Willis IC. 1996. Velocity-discharge relationships derived from dye-tracer experiments in glacial meltwaters: implications for subglacial flow conditions. *Hydrological Processes* **10**: 1411-1426.
- Nienow PW, Sharp M, Willis IC. 1998. Seasonal changes in the morphology of the subglacial drainage system, Haut Glacier d'Arolla, Switzerland. *Earth Surface Processes and Landforms* **23**: 825-843.
- Schuler TV, Fischer UH. 2009. Modelling the diurnal variation of tracer transit velocity through a subglacial channel. *Journal of Geophysical Research* **144**: F04017.
- Seaberg SZ, Seaberg SZ, Hooke R LeB, Wiberg DW. 1988. Character of the englacial and subglacial drainage system in the lower part of the ablation area of Storglaciären, Sweden, as revealed by dye-trace studies. *Journal of Glaciology* **34**: 217-227.
- Smart PL, Laidlaw IMS. 1977. An evaluation of some fluorescent dyes for water tracing. *Water Resources Research* **13**(1): 15-33.
- Stenborg T. 1969. Studies of the internal drainage of glaciers. *Geografiska Annaler* **51A**(1-2): 13-41.
- Turner Designs. 2010. *Cyclops-7 Submersible Sensors*, Turner Designs: Sunnyvale.
- Willis IC, Fitzsimmons CD, Melvold K, Andreassen LM, Giesen RH. 2012. Structure, morphology and water flux of a subglacial drainage system, Midtdalsbreen, Norway. *Hydrological Processes*: doi:10.1002/hyp.8431.
- Willis IC, Lawson W, Owens I, Jacobel B, Autridge J. 2009. Subglacial drainage system structure and morphology of Brewster Glacier, New Zealand. *Hydrological Processes* **23**: 384-396.
- Willis IC, Sharp MJ, Richards KS. 1990. Configuration of the drainage system of Midtdalsbreen, Norway, as indicated by dye-tracing experiments. *Journal of Glaciology* **36**(122): 89-101.
- Wilson JF, Cobb ED, Kilpatrick FA. 1986. *Fluorometric procedures for dye tracing*. United States Government Printing Office: Washington.

3.4.5. Glacier Movement

C. Scott Watson¹ and Duncan Quincey¹

¹ School of Geography and water@leeds, University of Leeds (scott@www.rockyglaciers.co.uk)



ABSTRACT: Quantification of glacier movement can supplement measurements of surface elevation change to allow an integrated assessment of glacier mass balance. Glacier velocity is also closely linked to the surface morphology of both clean-ice and debris-covered glaciers. Velocity applications include distinguishing active from inactive ice on debris-covered glaciers, identifying glacier surge events, or inferring basal conditions using seasonal observations. Surface displacements can be surveyed manually in the field using trigonometric principles and a total station or theodolite for example, or dGPS measurements, which allow horizontal and vertical movement to be quantified for accessible areas. Semi-automated remote sensing techniques such as feature tracking (using optical or radar imagery) and interferometric synthetic aperture radar (InSAR) (using radar imagery), can provide spatially distributed and multi-temporal velocity fields of horizontal glacier surface displacement. Remote-sensing techniques are more practical and can provide a greater distribution of measurements over larger spatial scales. Time-lapse imagery can also be exploited to track surface displacements, providing fine temporal and spatial resolution, although the latter is dependent upon the range between camera and glacier surface. This chapter outlines the costs, benefits, and methodological considerations when using field-based and remote sensing techniques for deriving glacier velocity, and example workflows are presented.

KEYWORDS: glacier velocity, feature tracking, SAR interferometry, time-lapse, surge

Introduction

Glacier movement is spatially and temporally variable, both within glacierised catchments, and between glacierised regions, determined by topographic and climatic conditions. Quantifying movement rates in conjunction with mass balance estimates, helps reveal the response of glacial environments to climate change. Glacier movement can be decoupled from climatic warming, such as during sudden surge events where rapid displacement can occur over a short period, or enhanced glacial retreat that can be initiated at lacustrine- or marine-terminating glaciers. Assessing the future evolution and mass balance of glaciers is essential to determine their contribution to sea level change, seasonal meltwater availability including supra-, en-, sub- and pro-glacial water storage, and glacial lake outburst flood (GLOF) hazard development.

Glacier movement can be attributed to basal sliding, internal ice deformation, and subglacial deformation of the glacier bed; hence there is an interaction between the gravitational driving force of downward ice movement, and the resistance encountered at the bed and margins (Benn and Evans, 2010). Difficulties in accessing the ice-bed interface of a glacier makes quantifying basal movement difficult, though not impossible (e.g. Hubbard, 2002; Kavanaugh and Clarke, 2006) hence glacier surface movement is commonly used as a proxy to represent overall glacier motion. Therefore, methods of quantifying glacier surface movement form the focus of this chapter.

Geomorphological interest

As an agent of significant landscape change, glacier movement and stagnation, both past and present, is of notable interest to geomorphologists at local to regional scales.

Examples include reconstructing past glacial extents (e.g. Benn and Owen, 2002; Nesje, 2009), assessing the proglacial implications of GLOF events following glacial recession and subsequent lake development (e.g. Westoby *et al.*, 2014), and examining glacial morphology and dynamics such as during surge events (e.g. Clarke *et al.*, 1984; Kamb *et al.*, 1985; Quincey *et al.*, 2011), or on debris-covered glaciers exhibiting stagnation and surface lowering (e.g. Immerzeel *et al.*, 2014). The contemporary response of ice sheets to climate change is also receiving significant attention owing to the implications for sea level rise and the potential for resource exploitation. Here, ice movement reflects the export and subsequent loss of ice mass into the ocean, which is coupled to surface meltwater fluxes (Zwally *et al.*, 2002). The likely response of ice sheets during deglaciation can be inferred from interpretations of deglaciated landscapes (e.g. the BRITICE mapping project) (Evans *et al.*, 2005), and used in conjunction with numerical modelling to assess contemporary ice sheet change (Bingham *et al.*, 2010).

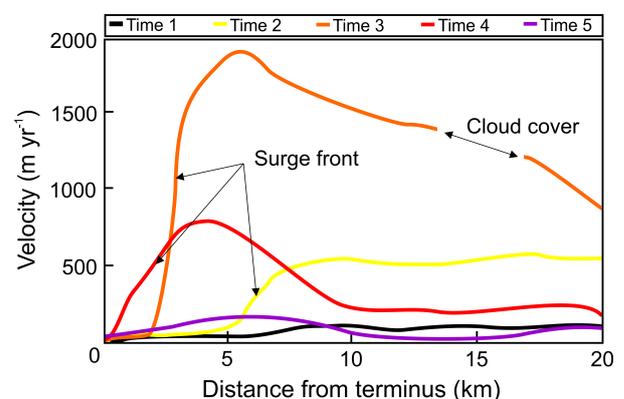
Debris-covered glaciers

Assessments of glacier movement are important when studying debris-covered glaciers owing to a complex response to climate change. Debris-covered glaciers that are stagnating and exhibiting surface lowering are associated with the development of supraglacial ponds (Bolch *et al.*, 2012). The thermal energy stored in ponds and transmitted to the underlying glacier ice can locally increase ablation, and under favourable conditions ponds can also coalesce into larger glacial lakes with subsequent water storage and potentially hazardous implications (Benn *et al.*, 2012). Exposed ice faces often surround supraglacial ponds and can represent areas of enhanced ablation in a melt regime that is largely suppressed by thick debris cover (Sakai *et al.*, 2002; Reid and Brock, 2014). Here, the persistence and evolution of both supraglacial ponds and ice cliffs are closely linked to a low glacier velocity (Quincey *et al.*, 2007).

Surge glaciers

Surge glaciers are characterised by an oscillatory behaviour which switches between

a long quiescent phase (decades to years) and a short active phase (years to decades) of rapid ice movement, where velocities can be ca. 100 times greater (Meier and Post, 1969). These events can bias snap-shot mass balance and glacial extent observations if not accounted for. However, over longer periods this effect may be diminished: Gardelle *et al.* (2013) observed similar mass budgets for surge-type and non-surging glaciers in the Pamir-Karakorum Himalaya over a 10 year period. The spatial and temporal change in glacier velocity during a surge can be observed using remote sensing techniques (e.g. Quincey *et al.*, 2011) (Figure 1), which can help describe the distribution, duration, magnitude, and return period of glacier surges, for which the controlling mechanisms are not fully understood. Evidence of past surge-type behaviour can be derived from characteristic deformation landforms such as thrust moraines, concertina eskers (Evans and Rea, 1999), or deformed medial moraines on the contemporary glacier surface (Meier and Post, 1969).



*Figure 1: Illustrative propagation of a glacier surge front derived from centre line surface velocities from Quincey *et al.* (2011). Note that areas of cloud cover or poor image contrast can lead to data voids.*

Observations of glacier movement

Glacier movement is associated with erosional and depositional landforms, and the transport of sediment down glaciated catchments. This can be important for contemporary glacier dynamics, such as a terminal moraine-dammed lake promoting calving retreat (e.g. Imja Tsho, Nepal), and can also be used to interpret and reconstruct past glacial landscapes. Water-occupied

over-deepened basins epitomise glacial erosion and subsequent glacial retreat.

Seasonal velocity observations can indicate basal thermal conditions. For example, contrasting summer and winter velocities suggest meltwater driven basal sliding is likely, which is linked to bed erosion and ultimately landscape evolution. Studies investigating glacier movement over shorter temporal timescales (e.g. hours, days, and weeks) than is common for remotely sensed observations are able to identify the links between glacier-hydrology and glacier velocity. Seasonal meltwater input into a glacier system can act to reduce basal drag and subsequently increase horizontal and vertical glacier velocity as the sub-glacial drainage system evolves (e.g. Mair *et al.*, 2001; Copland *et al.*, 2003). Hence the water pressure in this drainage system is closely linked to glacier sliding (Iken *et al.*, 1983). This increase in velocity can be short-lived where a subglacial meltwater reservoir outbursts (e.g. Bingham *et al.*, 2006). Depleted surface meltwater inputs and

increasing channelization of the drainage system subsequently reduces subglacial water pressure. However, horizontal velocities may remain raised compared to winter in response to continued meltwater input, despite a more effective drainage system (Bingham *et al.*, 2006).

Summary of techniques

Early observations of glacier movement utilised transects of stakes inserted into the glacier surface, which can be tracked from surrounding moraines using a theodolite and triangulation techniques (e.g. Kodama and Mae, 1976). Alternatively, increasingly utilised remote sensing techniques can enable a fully distributed analysis of glacier surface movement both horizontally and vertically (Quincey *et al.*, 2005), which complements geomorphological observations (e.g. Hambrey *et al.*, 2008), and allows an integrated assessment of glacier dynamics alongside mass balance observations (e.g. Benn *et al.*, 2012).

Table 1. A summary of techniques for assessing glacier surface motion and practical considerations.

Technique	Costs	Benefits	Considerations
Manual survey	<ul style="list-style-type: none"> dGPS, total station or theodolite > 1 field visit usually required 	<ul style="list-style-type: none"> Sub-metre accuracy (sub-centimetre with certain dGPS setups) Horizontal and vertical displacement information Point measurements of velocity, rather than an average for a pixel/cell (e.g. feature tracking) Sub-daily temporal resolution possible 	<ul style="list-style-type: none"> Site accessibility Selecting appropriate visible and persistent markers on a debris-covered surface Spatial coverage on the order of several km² depending upon terrain The NERC Geophysical Equipment has a range of survey equipment available for loan
Time-lapse survey	<ul style="list-style-type: none"> Several cameras generally required dGPS survey of reference targets > 1 field visit usually required 	<ul style="list-style-type: none"> Sub-metre accuracy Horizontal and vertical displacement information Ancillary information e.g. weather conditions Sub-daily temporal resolution possible with the overall mission duration limited by battery power (typically < 1 year) 	<ul style="list-style-type: none"> Number and position of cameras (considering site accessibility) Surface shadows Cloud or snow cover obscuring targets/ rain on the lens Restricting camera movement e.g. bolting down the enclosure Waterproofing of the camera and battery life/ ancillary solar power Camera theft Spatial coverage dependent upon camera quality and viewpoint elevation. Typically < 0.5 km² per camera
Feature tracking	<ul style="list-style-type: none"> Software access and training Imagery costs 	<ul style="list-style-type: none"> Automated workflows Dense velocity field coverage Glacier-scale spatial coverage 	<ul style="list-style-type: none"> Imagery choice: optical or SAR Imagery spatial and temporal resolution (see Figure 2). Unmanned Aerial Vehicle imagery could be sub-daily. However, more commonly used satellite imagery will be on the order of days to months. Cloud covered scenes/ poor image contrast on optical imagery Software choice/ processing algorithm
SAR interferometry	<ul style="list-style-type: none"> Technically more challenging than feature tracking 	<ul style="list-style-type: none"> Automated workflows Higher precision than is achievable using feature tracking with optical or SAR imagery Very small displacements are detectable 	<ul style="list-style-type: none"> The revisit period of imagery relative to the estimated glacier displacement in order to preserve coherence

Commonly used remote sensing techniques for assessing glacier surface movement include manual surveys, time-lapse camera surveys, satellite image feature tracking, and interferometry. A summary of the techniques, and what needs to be considered on a practical level, is provided in Table 1.

Manual survey

Manual surveys of glacier movement can be undertaken using a total station, theodolite or a dGPS to survey stakes inserted perpendicular into the glacier surface. The latter can achieve centimetre accuracy but access is required to the survey locations for all methods. The technique requires repeat measurements of stake positions so that the horizontal and vertical change between them can be calculated. Where debris is thick, a cross on a large boulder could be surveyed instead of stakes. Hubbard and Glasser (2005) provide a useful summary of surveying techniques and a method of transforming survey coordinates into velocities.

Manual survey methods require travel to the glacier itself, and then access to appropriate survey locations, which is time consuming and often hazardous on debris covered glaciers, and may not be possible on heavily crevassed clean ice glaciers. Therefore the spatial distribution of measurements is likely to be sparse, is dependent upon the preservation of the stakes between surveys, and is temporally limited to the duration and revisit period between field campaigns. It does however allow the simultaneous quantification of surface elevation change and glacier velocity (e.g. Pattyn *et al.*, 2003; Zhang *et al.*, 2010; Zhang *et al.*, 2011). This is important on debris covered glaciers where the dominant response to climatic warming is often surface lowering, rather than terminus retreat (Benn *et al.*, 2012). Cavity uplift also produces vertical glacier movement, related to seasonally high subglacial water pressures and development of the glacier drainage system (e.g. Iken *et al.*, 1983; Hooke *et al.*, 1989; Mair *et al.*, 2002; Bingham *et al.*, 2006). Further information on the relationship between glacier strain and vertical velocity, and the calculation of cavity uplift velocities can be found in Iken *et al.* (1983) and Hooke *et al.* (1989). Assessments of vertical change are therefore essential and are otherwise

limited by the high uncertainty often accompanying alternative remotely sensed geodetic DEM differencing methods (Racoviteanu *et al.*, 2008). Point-based measurements are also used to validate remote sensing derived velocity fields (e.g. Quincey *et al.*, 2009b; Herman *et al.*, 2011).

Uncertainty

Three types of error associated with total station/theodolite surveys are discussed by Copland *et al.* (2003). Briefly they are: positional error from determining the location of a stake; a displacement error between two stake surveys (e.g. the sum of positional errors); and, a velocity error, which is the displacement error adjusted for the time interval between surveys. A longer measurement interval reduces the overall velocity error; hence measurements may be subsampled over longer time periods. Mean velocity errors observed by Copland *et al.* (2003) were $\pm 1.1 \text{ cm d}^{-1}$ in the horizontal, and $\pm 0.6 \text{ cm d}^{-1}$ in the vertical.

Uncertainty in dGPS surveys occurs from instrument error of the device and the field setup. For glacier movement applications (e.g. Quincey *et al.*, 2009a), the base station is mounted on stable ground off-glacier, and roving stations are either mounted permanently in the area of interest or used to survey locations intermittently (e.g. stakes or ground control points). The accuracy of a roving survey depends on the occupation time of the unit at each location, and satellite and atmospheric conditions, including the line of site to the base station (Hubbard and Glasser, 2005). Care is also required to maintain a stable antenna rod inclination during each occupation. Horizontal accuracy on the order of centimetres is achievable using a roving survey (e.g. Whitehead *et al.*, 2014). Achieving sub-centimetre accuracy comparable to that of a static survey requires long survey occupation times. However, such precision may not be practical or required.

Time-lapse camera survey

Using a time lapse camera array to observe glacier movement allows continuous data collection between field visits, in addition to ancillary information such as local weather conditions. The camera is used similarly to a theodolite, measuring the horizontal and

vertical trace lines from reference targets to the camera lens, whilst correcting for small camera movements using reference points on stable ground (Harrison *et al.*, 1992).

Implementation

3D point clouds can be automatically generated from photographs taken at different locations using Structure-from-Motion software (e.g. Westoby *et al.*, 2012), allowing both vertical and horizontal displacements to be quantified multi-temporally. This typically requires ca. >7 photographs of the subject taken from different locations (Micheletti *et al.*, 2014). For fewer cameras typical of a time-lapse array, surface movement in 3D space can be determined using traditional photogrammetry applied to cameras observing the glacier surface from an elevated position (e.g. Whitehead *et al.*, 2014). Whitehead *et al.* (2014) used 0.6 m diameter red circles on a white background as reference targets which were GPS surveyed to establish the

coordinate system and reference frame. Daily X, Y, and Z position for targets can then be calculated using the multi-temporal imagery.

Uncertainty

Using a 39 Megapixel camera taking photos every ca. 10 minutes, Maas *et al.* (2008) achieved glacier displacement measurements accurate to several centimetres at distances of up to 5 km from the camera. Whitehead *et al.* (2014) achieved target positions of < 0.1 m at up to 800 m using two 10 Megapixel cameras. Glacier velocity was resolved up to 4 km from a 10.2 Megapixel camera set-up by Ahn and Box (2010). Here, for the Jakobshavn Isbræ averaging 37.6 m d^{-1} the displacement uncertainty was 2 %, whereas the slower moving Umiyako (5.9 m d^{-1}) averaged an uncertainty of 9 %. Uncertainty is related to glacier velocity, since positional uncertainty forms a lower proportion of an overall larger displacement.

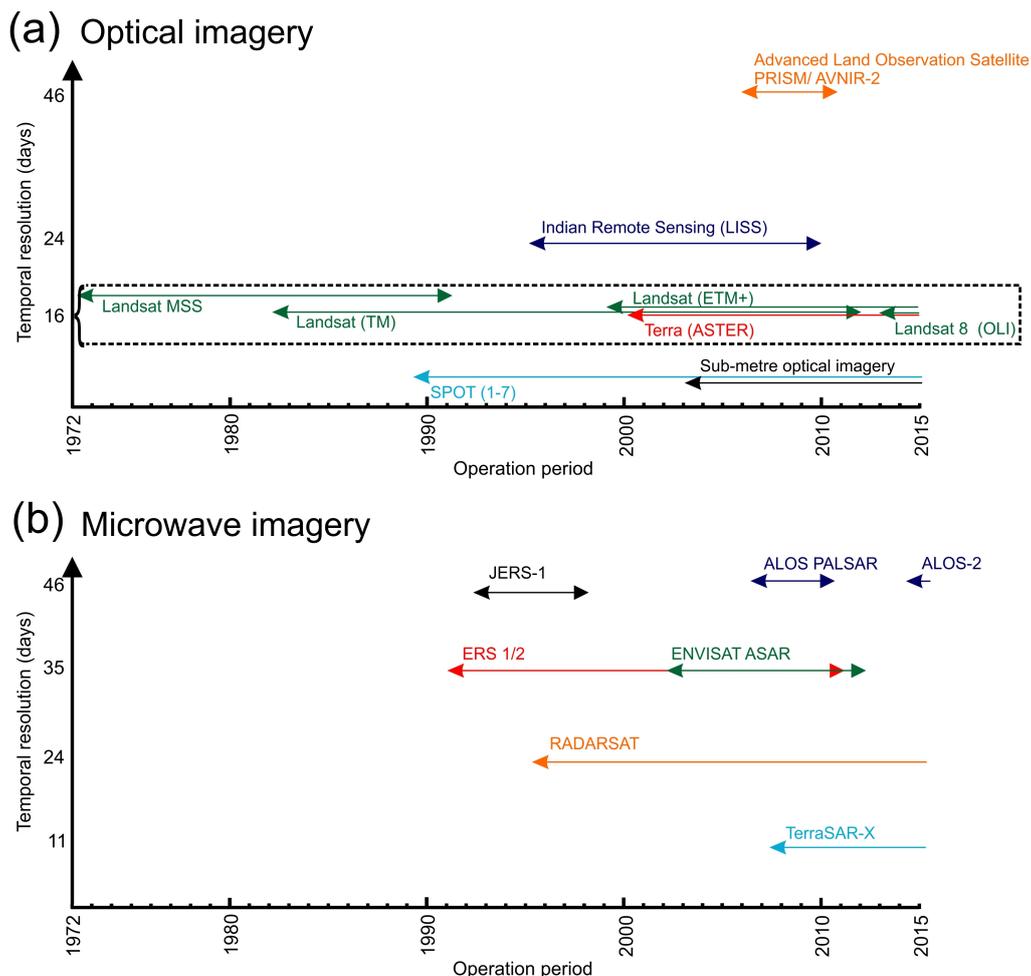


Figure 2. Typical optical and microwave imagery suitable for feature tracking applications. Actual archive coverage varies between regions. The Landsat and Terra (ASTER) archives have a 16 day repeat cycle (dashed box on (a)).

Feature tracking

Automated feature tracking using cross-correlation algorithms can generate spatially distributed velocity fields using multi-temporal optical or microwave remotely sensed imagery (Figure 2, Table 2). The software is used to extract features (e.g. boulders or debris patterns) from multiple images and the movement of the features is used to generate velocity fields. In addition to a greatly increased spatial density of measurements, feature tracking algorithms offer improved accuracy over manual feature tracking using remotely sensed imagery (e.g. Holt *et al.*, 2013), since the images can be co-registered to sub-pixel accuracy before analysis. Feature tracking relies on the preservation of features between scenes, such as boulders or debris patterns and crevasses (Strozzi *et al.*, 2002; Luckman *et al.*, 2003). Its suitability is therefore determined by the imagery resolution and revisit period, relative to glacier displacement between scenes. This can limit its application to fast moving (10 s m d^{-1}) glaciers (Maas *et al.*, 2008), where sufficient temporal resolution imagery is not available or would be prohibitively expensive to obtain.

Implementation

Optimum optical input imagery should have the same sun angle and azimuth, and image viewpoint, such that surface features are similarly represented between scenes (Scambos *et al.*, 1992). For Landsat imagery this can be achieved using the same path and row locations for imagery collected at a similar time of year. Here, the finest resolution 15 m panchromatic band is likely to be used. However, where the panchromatic band is saturated, a combination of several bands may offer the best performance (e.g. Dehecq *et al.*, 2015). When using ASTER imagery, Redpath *et al.* (2013) observed the best performance with band 1 and the worst with band 3, attributed to enhanced band 1 contrast on the debris covered Tasman Glacier. The time interval between imagery should reflect the expected glacier velocity, such that surface features will have moved several pixels. A larger time interval therefore reduces uncertainty up to a threshold point, beyond which deformation in surface features between scenes will act to increase the uncertainty in establishing a match

(Scambos *et al.*, 1992). Derived velocity fields require post-processing to remove spurious points and those that do not conform to the expected glacier flow direction or speed (Redpath *et al.*, 2013; Raup *et al.*, 2007) (Figure 3). Here thresholds for exclusion based on the correlation coefficient and signal to noise ratio can be set.

Optical vs. microwave imagery

Feature tracking techniques can also be implemented on SAR imagery (e.g. Strozzi *et al.*, 2002; Luckman *et al.*, 2007; Quincey *et al.*, 2009b) (Figure 2). SAR imagery does not require cloud free conditions to be suitable for feature tracking, nor is it affected by varying solar illumination, in contrast to optical imagery. Therefore SAR imagery can produce a more consistent dataset for monitoring glacier velocity (Luckman *et al.*, 2003, 2007), where surface features are distinctly visible over image speckle (noise). The procedure uses the same principle of cross-correlating features in image patches between scenes to determine glacier displacement as for optical imagery (Scambos *et al.*, 1992). In cases where coherence (phase of the received signal) is maintained between image pairs, the speckle can also be tracked on a finer scale to yield a displacement field.

Uncertainty

Errors in glacier displacement are associated with changes in the tracked surface features over the selected temporal resolution, geometric transformations of the data, and error in the location of stable reference points, which should be selected to exhibit zero displacement (e.g. bedrock outcrops) (Luckman *et al.*, 2007). Over shorter time periods (e.g. 1 day) the associated uncertainties are larger than compared with a longer (e.g. 35 day) period, where Luckman *et al.* (2003) estimated errors at ca. 1 m d^{-1} and ca. 0.03 m d^{-1} respectively using European Remote Sensing (ERS) satellite SAR imagery applied to several outlet glaciers in East Greenland. Error was reported to be sub-centimetre per day when applied to Himalayan glaciers by Luckman *et al.* (2007), and 1/20 of a pixel in the study of Strozzi *et al.* (2002).

Table 2. Examples of the software and techniques available for determining glacier velocity.

Method	Application	Glacial examples	Details
Feature tracking	CIAS software (2000) Optical imagery	Redpath <i>et al.</i> (2013)	Velocity fields were derived for the Tasman glacier 2009-2010, 2010-2011. Good agreement was observed between GPS measured and satellite derived velocities ($R^2 = 0.97$). ASTER band 1 was shown to offer the best performance for the debris covered glacier, and band 3N the worst (RMSE of 16.5 m a^{-1} and 60.1 m a^{-1} respectively)
		Kääb <i>et al.</i> (2005)	Landsat 7 ETM+ pan band and ASTER 3N band were used to derive velocity fields for the Kronebreen Glacier, Svalbard. Optical imagery feature tracking performed well over higher velocity areas. Interferometric synthetic aperture radar (InSAR) data were most suitable where optical contrast was low (e.g. over snow cover)
	COSI-Corr software (2007) Optical imagery	Herman <i>et al.</i> (2011)	ASTER band 3N imagery separated by 16 days was used to derive glacier velocities for an area of the Southern Alps, New Zealand. Velocities up to ca. 5 m d^{-1} were observed with uncertainties of ca. 0.19 m d^{-1} . Comparison with ground dGPS measurements revealed large discrepancies where displacements were small (e.g. in accumulation areas)
		Scherler <i>et al.</i> (2008)	Glacier-surface velocities were derived from ASTER imagery for two Himalayan glaciers. A quality assessment methodology was demonstrated to improve measurement accuracy. Displacement error was on the order of 2–4 m per correlation
	ImGRAFT software (2014) Time-lapse imagery	Messerli and Grinsted (2015)	The image georectification and feature tracking toolbox (ImGRAFT) tailored for terrestrial oblique imagery was tested on time lapse imagery of the Engabreen Glacier icefall, Norway. A fine resolution DEM (e.g. laser scan) was required for georectification. The imagery revealed local flow characteristics, including extensional and compressional flow in the ice fall
	Satellite radar feature tracking (SRFT). Radar imagery	Luckman <i>et al.</i> (2007)	Repeat pass satellite radar imagery allowed glacier surface patterns to be tracked between scenes. Scene separations of 245 – 1890 days. Suggested error of $< 0.5 \text{ cm d}^{-1}$ over the study period. The technique was most suitable on slow moving, debris-covered glacier areas
Manual feature tracking. Optical imagery	Immerzee <i>et al.</i> (2014)	Manual feature tracking was used to create a velocity field for the Lirung Glacier, Nepal using a ca. 5 cm resolution DEM derived from Unmanned Aerial Vehicle (UAV) surveys in May and October 2013. Distinguishable features (e.g. boulders) ($n = 145$) were identified on multi-temporal ortho-mosaics and used to calculate horizontal displacement	
SAR interferometry	GAMMA software SAR	Quincey <i>et al.</i> (2007)	SAR interferometry was applied to five Himalayan debris-covered glaciers and provided a spatial resolution ($< 20 \text{ m}^2$) and errors of less than 0.01 m d^{-1} for a pair of ERS-1/2 images from 29–30 March 1996. Supraglacial lake development was associated with areas of low glacier displacement
		Schneevogt <i>et al.</i> (2012)	Glacier displacement was derived for Comfortlessbreen, Svalbard, using SAR interferometry applied to ERS-1/2 scenes from April and May 1996. This provided a pre-surge velocity baseline
dGPS	Stake survey	Herman <i>et al.</i> (2011)	Crevasse restricted longitudinal stake coverage during the manual survey. A maximum horizontal error of 0.2 m was suggested between stake surveys (2002-2006). dGPS measurements also revealed vertical glacier surface changes
	Stake survey	Kirkbride and Deline (2013)	Ablation stakes were surveyed annually (2005 – 2008) on the Glacier d'Estellette, Italy, to characterise surface debris transport. Surveying was conducted using a dGPS in real-time stop-and-go mode, which lead to mean positional errors of 0.031 m in the horizontal and 0.026 m in the vertical
Theodolite / total station	Stake survey	Copland <i>et al.</i> (2003)	34 stakes were surveyed in the lower 3 km of the John Evans Glacier in 1998. Reflecting prisms were mounted on the 3 m stakes which were drilled into the glacier surface. Each stake was surveyed daily from two independent locations using a total station theodolite
	Stake survey	Mair <i>et al.</i> (2008)	13 stakes placed on the Haut Glacier d'Arolla, Switzerland, were surveyed over 10 days using a total station situated on a moraine to determine diurnal velocity variation. Two-hourly surveys were also conducted for a three-day period during daylight hours. Stakes were located to within $\pm 3\text{-}5 \text{ mm}$ which translated into velocity errors of $< \pm 0.04 \text{ m d}^{-1}$ when averaged over two hour intervals. Diurnal variability was observed during the two-hourly surveys, although this was on the order of measurement uncertainty

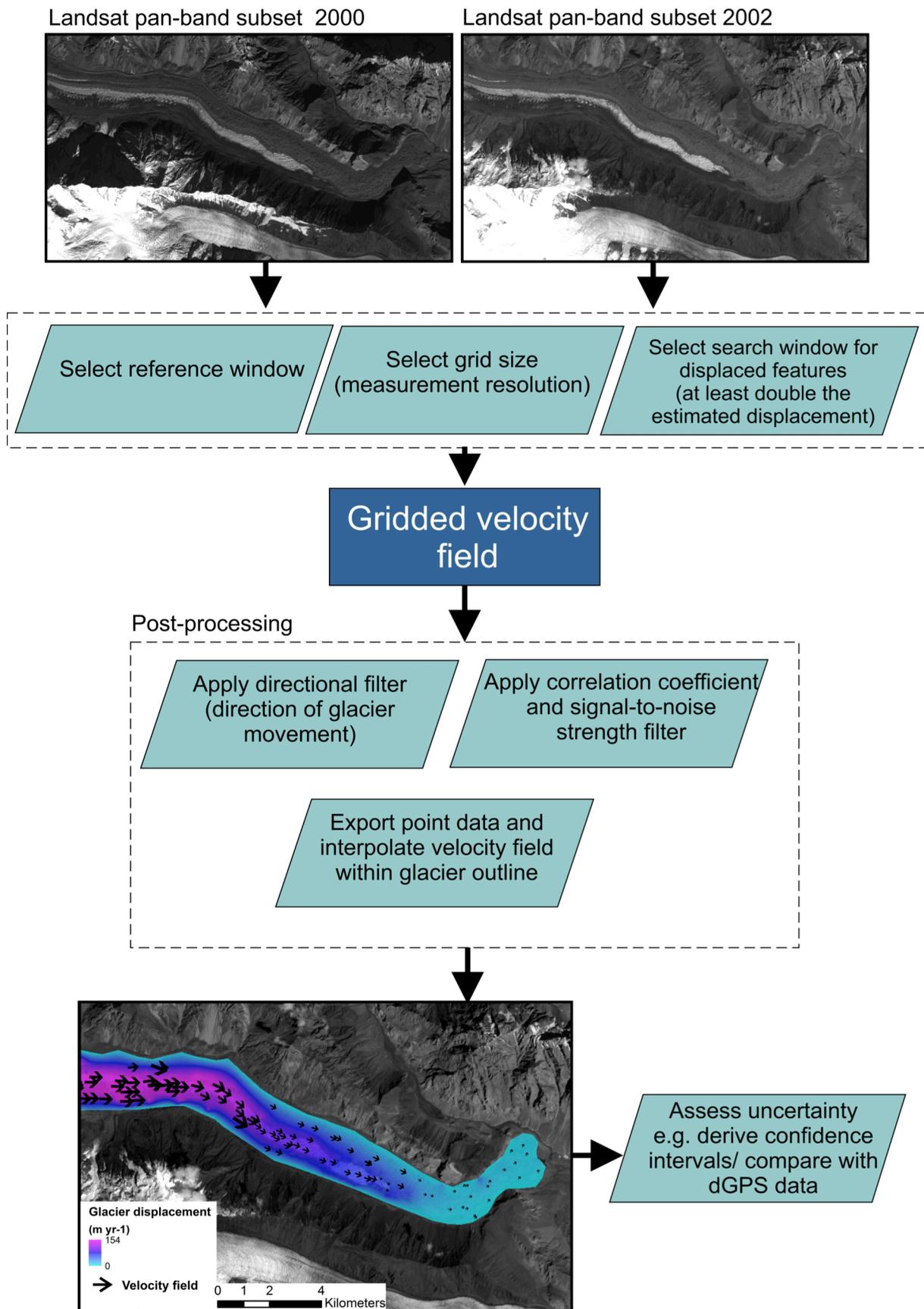


Figure 3. A typical feature tracking workflow for deriving glacier velocity. This example used Landsat 7 ETM+ panchromatic scenes of the Batura Glacier and CIAS software, both available from: <http://www.mn.uio.no/geo/english/research/projects/icemass/cias>

SAR interferometry

SAR interferometry reveals glacier displacement represented through fringe patterns, which correspond to surface displacement (e.g. Figure 4). The process requires coherence to be maintained between images (Massonnet and Feigl, 1998). Coherence can be lost because of:

1. too much displacement between acquisitions due to rapid ice flow or a long satellite revisit period.
2. a change in surface appearance due to surface melt or fresh snow cover.

The requirement of preserved coherence usually limits InSAR measurements to diurnal velocities, which can be beneficial for seasonal or short-term patterns, but becomes unrepresentative if scaled up to annual or longer-term periods. In this case, feature tracking techniques can be beneficial (Strozzi *et al.*, 2002). However, where both short-term and annually averaged velocities are required, both SAR interferometry and feature tracking can be used as complementary methods (e.g. Quincey *et al.*, 2009b).

Implementation

The workflow for deriving glacier displacement using interferometry is outlined in Figure 5 and is extensively reviewed by Massonnet and Feigl (1998), summarised here. In brief, SAR interferometry requires the images to be finely co-registered so that an

interferogram can be produced. This represents, on a pixel by pixel basis, the phase difference between the received signals from each image. This phase difference is dependent on surface topography, the positions of the sensors at the time of image acquisition, a signal relating to the curvature of the Earth, and any surface displacement that has occurred on the ground. The signals relating to the first three of these factors can be simulated (or measured) and then removed from the overall phase to leave only that relating to surface displacement. The remaining glacier displacement signal is then unwrapped from a known stationary point to create a velocity

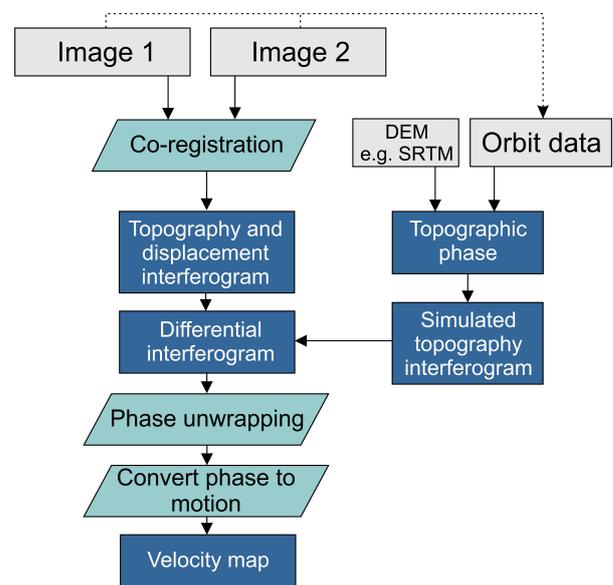


Figure 5. SAR interferometry processing workflow to calculate glacier velocity. Adapted from Rao (2014).

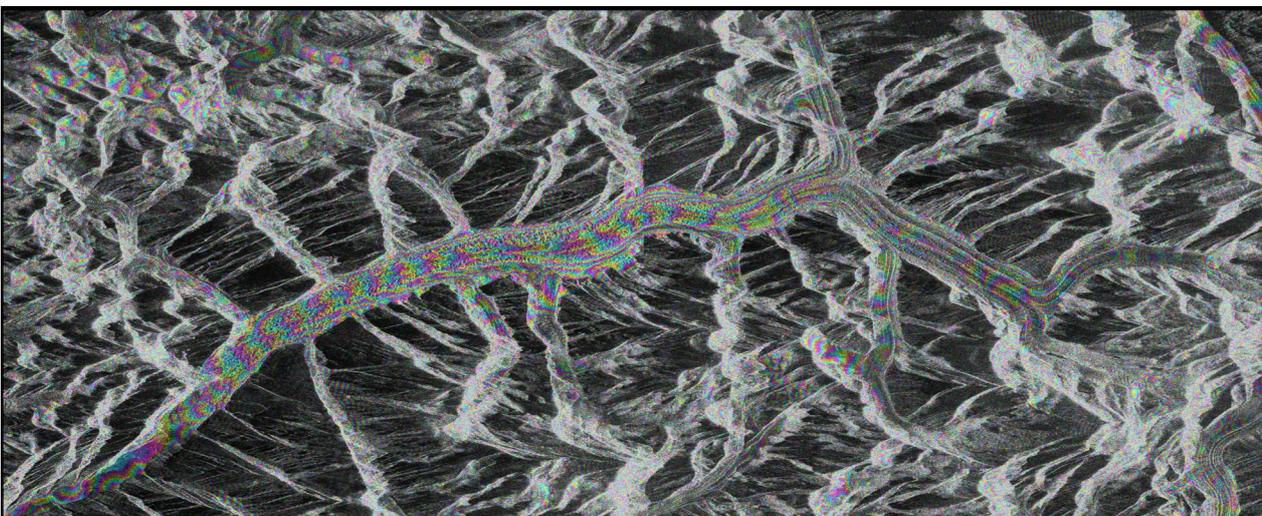


Figure 4. An example interferogram of the Baltoro Glacier. Coloured fringes represent glacier movement.

map, although it should be noted that any vertical surface change (e.g. subsidence, surface melt) will also be incorporated in the final velocity data. Displacements that can be measured are normally of the order of the wavelength of the imagery (i.e. cm's).

Conclusion

dGPS surveying techniques can afford greater precision compared to feature tracking methods but measurements are restricted both spatially and temporally, dictated by site accessibility and field costs. However, dGPS measurements also quantify vertical displacement, important on downwasting debris-covered glaciers, or when assessing cavity uplift. The increasing use of time-lapse imagery can be exploited with semi-automated techniques to derive glacier velocity and reveal local displacement characteristics, whilst also resolving displacements at centimetre resolution at hundreds of metres from the camera positions.

Recent computational developments have allowed spatially distributed velocity fields to be calculated for multiple glaciers with relative ease and with high accuracy, using a wide range of optical and SAR remotely sensed imagery. Feature tracking and SAR interferometry can be used to obtain annually averaged and short-term velocity fields respectively, such that the techniques can be complementary. The greater availability of datasets and broader applicability to both debris-covered and clean-ice glaciers tends towards greater application of feature tracking methods, although interferometry has the clear advantage of being able to yield velocity data over largely featureless surfaces. In either case, use of remotely-sensed imagery allows glacier velocity to be derived for crevassed and other inaccessible areas of the glacier, which cannot be otherwise obtained using manual surveying techniques.

As remote sensing imagery archives continue to expand, observations of glacier velocity will increasingly complement assessments of mass balance. This allows a more integrated assessment of glacier dynamics to assess the response to climate change and other internal feedback mechanisms. This is especially pertinent at lacustrine- and marine-

terminating glaciers where interplay between glacier mass export and calving exists, and also on debris-covered glaciers where surface morphology and glacial lake development are closely linked to glacier stagnation.

Acknowledgements

We thank two reviewers and the editor, Lucy Clarke, whose comments helped improve this article.

References

- Ahn Y and Box JE. 2010. Glacier velocities from time-lapse photos: technique development and first results from the Extreme Ice Survey (EIS) in Greenland. *Journal of Glaciology*. **56**: 723-734.
- Benn DI, Bolch T, Hands K, Gulley J, Luckman A, Nicholson LI, Quincey D, Thompson S, Toumi R and Wiseman S. 2012. Response of debris-covered glaciers in the Mount Everest region to recent warming, and implications for outburst flood hazards. *Earth-Science Reviews*. **114**: 156-174.
- Benn DI and Evans DJA. 2010. *Glaciers and glaciation*. London: Hodder Education.
- Benn DI and Owen LA. 2002. Himalayan glacial sedimentary environments: a framework for reconstructing and dating the former extent of glaciers in high mountains. *Quaternary International*. **97-8**: 3-25.
- Bingham RG, King EC, Smith AM and Pritchard HD. 2010. Glacial geomorphology: Towards a convergence of glaciology and geomorphology. *Progress in Physical Geography*. **34**: 327-355.
- Bingham RG, Nienow PW, Sharp MJ and Copland L. 2006. Hydrology and dynamics of a polythermal (mostly cold) High Arctic glacier. *Earth Surface Processes and Landforms*. **31**: 1463-1479.
- Bolch T, Kulkarni A, Käab A, Huggel C, Paul F, Cogley JG, Frey H, Kargel JS, Fujita K, Scheel M, Bajracharya S and Stoffel M. 2012. The State and Fate of Himalayan Glaciers. *Science*. **336**: 310-314.
- Clarke GKC, Collins SG and Thompson DE. 1984. Flow, thermal structure, and subglacial conditions of a surge-type glacier. *Canadian Journal of Earth Sciences*. **21**: 232-240.

- Copland L, Sharp MJ and Nienow PW. 2003. Links between short-term velocity variations and the subglacial hydrology of a predominantly cold polythermal glacier. *Journal of Glaciology*. **49**: 337-348.
- Dehecq A, Gourmelen N and Trouve E. 2015. Deriving large-scale glacier velocities from a complete satellite archive: Application to the Pamir–Karakoram–Himalaya. *Remote Sensing of Environment*. **162**: 55-66.
- Evans DJA, Clark CD and Mitchell WA. 2005. The last British Ice Sheet: A review of the evidence utilised in the compilation of the Glacial Map of Britain. *Earth-Science Reviews*. **70**: 253-312.
- Evans DJA and Rea BR. 1999. Geomorphology and sedimentology of surging glaciers: a land-systems approach. *Annals of Glaciology*. **28**: 75-82.
- Gardelle J, Berthier E, Arnaud Y and Kaab A. 2013. Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999-2011. *Cryosphere*. **7**: 1263-1286.
- Hambrey MJ, Quincey DJ, Glasser NF, Reynolds JM, Richardson SJ and Clemmens S. 2008. Sedimentological, geomorphological and dynamic context of debris-mantled glaciers, Mount Everest (Sagarmatha) region, Nepal. *Quaternary Science Reviews*. **27**: 2361-2389.
- Harrison WD, Echelmeyer KA, Cosgrove DM and Raymond CF. 1992. The determination of glacier speed by time-lapse photography under unfavorable conditions. *Journal of Glaciology*. **38**: 257-265.
- Herman F, Anderson B and Leprince S. 2011. Mountain glacier velocity variation during a retreat/advance cycle quantified using sub-pixel analysis of ASTER images. *Journal of Glaciology*. **57**: 197-207.
- Holt TO, Glasser NF, Quincey DJ and Siegfried MR. 2013. Speedup and fracturing of George VI Ice Shelf, Antarctic Peninsula. *The Cryosphere*. **7**: 797-816.
- Hooke RL, Calla P, Holmlund P, Nilsson M and Stroeven A. 1989. A three year record of seasonal variations in surface velocity, Storglaciären, Sweden. *Journal of Glaciology*. **35**: 235-247.
- Hubbard B. 2002. Direct measurement of basal motion at a hard-bedded, temperate glacier: Glacier de Tsanfleuron, Switzerland. *Journal of Glaciology*. **48**: 1-8.
- Hubbard B and Glasser NF. 2005. *Field techniques in glaciology and glacial geomorphology*. Chichester: John Wiley.
- Iken A, Rothlisberger H, Flotron A and Haeberli W. 1983. The uplift of Unteraargletscher at the beginning of the melt season – a consequence of water storage at the bed. *Journal of Glaciology*. **29**: 28-47.
- Immerzeel WW, Kraaijenbrink PDA, Shea JM, Shrestha AB, Pellicciotti F, Bierkens MFP and de Jong SM. 2014. High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sensing of Environment*. **150**: 93-103.
- Kääb A, Lefauconnier B and Melvold K. 2005. Flow field of Kronebreen, Svalbard, using repeated Landsat 7 and ASTER data. In: Dowdeswell, J. and Willis, I.C. eds. *Annals of Glaciology, Vol 42, 2005*. pp.7-13.
- Kamb B, Raymond CF, Harrison WD, Engelhardt H, Echelmeyer KA, Humphrey N, Brugman MM and Pfeffer T. 1985. Glacier Surge Mechanism: 1982-1983 Surge of Variegated Glacier, Alaska. *Science*. **227**: 469-479.
- Kavanaugh JL and Clarke GKC. 2006. Discrimination of the flow law for subglacial sediment using in situ measurements and an interpretation model. *Journal of Geophysical Research-Earth Surface*. **111**.
- Kirkbride MP and Deline P. 2013. The formation of supraglacial debris covers by primary dispersal from transverse englacial debris bands. *Earth Surface Processes and Landforms*. **38**: 1779-1792.
- Kodama H and Mae S. 1976. The Flow of Glaciers in the Khumbu Region. Glaciological Expedition to Nepal, Contribution No. 10. *Journal of the Japanese Society of Snow and Ice*. **38**: 31-36.
- Luckman A, Murray T, Jiskoot H, Pritchard H and Strozzi T. 2003. ERS SAR feature-tracking measurement of outlet glacier velocities on a regional scale in East Greenland. *Annals of Glaciology*. **36**: 129-134.
- Luckman A, Quincey D and Bevan S. 2007. The potential of satellite radar interferometry and feature tracking for monitoring flow rates

- of Himalayan glaciers. *Remote Sensing of Environment*. **111**: 172-181.
- Maas H-G, Schwalbe E, Dietrich R, Bässler M and Ewert H. 2008. Determination of spatio-temporal velocity fields on glaciers in West-Greenland by terrestrial image sequence analysis. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*. **37**: 1419-1424.
- Mair D, Hubbard B, Nienow P, Willis I and Fischer UH. 2008. Diurnal fluctuations in glacier ice deformation: Haut Glacier d'Arolla, Switzerland. *Earth Surface Processes and Landforms*. **33**: 1272-1284.
- Mair DWF, Sharp MJ and Willis IC. 2002. Evidence for basal cavity opening from analysis of surface uplift during a high-velocity event: Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology*. **48**: 208-216.
- Mair D, Nienow P, Willis I and Sharp M. 2001. Spatial patterns of glacier motion during a high-velocity event: Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology*. **47**: 9-20.
- Massonnet D and Feigl KL. 1998. Radar interferometry and its application to changes in the earth's surface. *Reviews of Geophysics*. **36**: 441-500.
- Meier MF and Post A. 1969. What are glacier surges? *Canadian Journal of Earth Sciences*. **6**: 807-817.
- Messerli A and Grinsted A. 2015. Image georectification and feature tracking toolbox: ImGRAFT. *Geosci. Instrum. Method. Data Syst.* **4**: 23-34.
- Micheletti N, Chandler JH and Lane SN. 2014. Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*. n/a-n/a.
- Nesje A. 2009. Latest Pleistocene and Holocene alpine glacier fluctuations in Scandinavia. *Quaternary Science Reviews*. **28**: 2119-2136.
- Quincey DJ, Braun M, Glasser NF, Bishop MP, Hewitt K and Luckman A. 2011. Karakoram glacier surge dynamics. *Geophysical Research Letters*. **38**: L18504.
- Quincey DJ, Copland L, Mayer C, Bishop M, Luckman A and Belò M. 2009a. Ice velocity and climate variations for Baltoro Glacier, Pakistan. *Journal of Glaciology*. **55**: 1061-1071.
- Quincey DJ, Luckman A and Benn D. 2009b. Quantification of Everest region glacier velocities between 1992 and 2002, using satellite radar interferometry and feature tracking. *Journal of Glaciology*. **55**: 596-606.
- Quincey DJ, Richardson SD, Luckman A, Lucas RM, Reynolds JM, Hambrey MJ and Glasser NF. 2007. Early recognition of glacial lake hazards in the Himalaya using remote sensing datasets. *Global and Planetary Change*. **56**: 137-152.
- Quincey DJ, Lucas RM, Richardson SD, Glasser NF, Hambrey MJ and Reynolds JM. 2005. Optical remote sensing techniques in high-mountain environments: application to glacial hazards. *Progress in Physical Geography*. **29**: 475-505.
- Racoviteanu AE, Williams MW and Barry RG. 2008. Optical remote sensing of glacier characteristics: A review with focus on the Himalaya. *Sensors*. **8**: 3355-3383.
- Rao YS. 2014. Synthetic Aperture Radar (SAR) Interferometry for Glacier Movement Studies. In: Singh, V. Singh, P. and Haritashya, U. eds. *Encyclopedia of Snow, Ice and Glaciers*. Springer Netherlands, pp.1133-1142.
- Raup B, Kaab A, Kargel JS, Bishop MP, Hamilton G, Lee E, Paul F, Rau F, Soltesz D, Khalsa SJS, Beedle M and Helm C. 2007. Remote sensing and GIS technology in the global land ice measurements from space (GLIMS) project. *Computers & Geosciences*. **33**: 104-125.
- Redpath TAN, Sirguey P, Fitzsimons SJ and Kääh A. 2013. Accuracy assessment for mapping glacier flow velocity and detecting flow dynamics from ASTER satellite imagery: Tasman Glacier, New Zealand. *Remote Sensing of Environment*. **133**: 90-101.
- Reid TD and Brock BW. 2014. Assessing ice-cliff backwasting and its contribution to total ablation of debris-covered Miage glacier, Mont Blanc massif, Italy. *Journal of Glaciology*. **60**: 3-13.
- Sakai A, Nakawo M and Fujita K. 2002. Distribution characteristics and energy balance of ice cliffs on debris-covered glaciers, Nepal Himalaya. *Arctic Antarctic and Alpine Research*. **34**: 12-19.

Scambos TA, Dutkiewicz MJ, Wilson JC and Bindschadler RA. 1992. Application of image cross-correlation to the measurement of glacier velocity using satellite image data. *Remote Sensing of Environment*. **42**: 177-186.

Schneevoigt NJ, Sund M, Bogren W, Kaab A and Weydahl DJ. 2012. Glacier displacement on Comfortlessbreen, Svalbard, using 2-pass differential SAR interferometry (DInSAR) with a digital elevation model. *Polar Record*. **48**: 17-25.

Strozzi T, Luckman A, Murray T, Wegmuller U and Werner CL. 2002. Glacier motion estimation using SAR offset-tracking procedures. *Geoscience and Remote Sensing, IEEE Transactions on*. **40**: 2384-2391.

Westoby MJ, Glasser NF, Brasington J, Hambrey MJ, Quincey DJ and Reynolds JM. 2014. Modelling outburst floods from moraine-dammed glacial lakes. *Earth-Science Reviews*. **134**: 137-159.

Westoby MJ, Brasington J, Glasser NF, Hambrey MJ and Reynolds JM. 2012. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*. **179**: 300-314.

Whitehead K, Moorman B and Wainstein P. 2014. Measuring daily surface elevation and velocity variations across a polythermal arctic glacier using ground-based photogrammetry. *Journal of Glaciology*. **60**: 1208-1220.

Zwally HJ, Abdalati W, Herring T, Larson K, Saba J and Steffen K. 2002. Surface Melt-Induced Acceleration of Greenland Ice-Sheet Flow. *Science*. **297**: 218-222.

Tracing Fine-Sediment Using Artificial Radionuclides

Philip Greenwood¹

¹Geography, College of Life and Environmental Science, University of Exeter, UK, (greenwoodphil@yahoo.co.uk)



ABSTRACT: The environmental characteristics of many artificial gamma-emitting radionuclides provide a number of distinct advantages over other forms of fine-sediment tracer that are readily exploitable when undertaking erosion investigations. This makes them particularly effective for accurately measuring the movement of small quantities of fine-sediment, especially in low-energy environments where soil movement is commensurately slow, or for measuring erosion rates over short (i.e. event-based) timescales. The major advantages and disadvantages associated with their use as tracers are outlined and briefly discussed. Three examples of tracing applications using caesium-134 (¹³⁴Cs) and cobalt-60 (⁶⁰Co) are summarised in order to highlight their versatility and demonstrate the diverse ways in which they can be used in field- and laboratory-based investigations.

KEYWORDS: artificial radionuclide; sediment-tracer; fine-sediment; caesium-134; cobalt-60.

Introduction

Fine-sediment (< 2 mm dia.) acts as a vector for the transfer of nutrients and contaminants and its removal can ultimately lead to a loss of agricultural productivity (Walling and Quine, 1992), as well as pollute terrestrial and aquatic ecosystems (Walling et al., 2003; Walling, 2004). Against this background, numerous approaches have been developed for quantifying soil movement through catchments. One approach is through the use of artificially applied sediment tracers. A tracer of this type can be defined as any identifiable substance which, by examining its behaviour over a given time-frame, may provide information on the behaviour of the host sediment (Sauzay, 1973; Evans, 1983). To be suitable, a tracer must meet certain requirements. The most fundamental include the ability to faithfully replicate the behaviour of the medium being traced, yet remain distinguishable from its surroundings in order to facilitate its re-identification and measurement (Sauzay, 1973; Evans, 1983; Guirese & Revel, 1995; McCubbin &

Leonard, 1995; Foster & Lees, 2000). Many different materials and substances have previously been used as sediment tracers. Their success in providing representative data, however, relies on an assumption that the tracer will behave in the same manner as the host sediment. In order for this assumption to be met, the physical characteristics of the tracer (i.e. approximate aggregate size, bulk density, etc.) must match the host material as closely as possible (Evans, 1983; Parsons et al., 1993; Ventura et al., 2002). Achieving this has proven to be extremely difficult (*cf.* Parsons et al., 1993; Riebe, 1995; Ventura et al., 2002; Zhang et al., 2003; Polyakov and Nearing, 2004), owing to the physical characteristics of fine-sediment and its tendency to be dispersed, often over relatively long distances and sometimes in multiple directions (Fullen, 1982). The use of artificial radionuclides offers a convenient way in which this obstacle can be overcome.

Radionuclides as fine-sediment tracers

Each radionuclide (or radioisotope) emits high-energy ionising radiation at a unique wavelength. This characteristic 'signature' can be used to identify an individual radionuclide and the radiation emission represents a powerful tool that can be used to trace physical and chemical pathways (Jones and Atkins, 2000). Gamma rays represent the most energetic form of ionising radiation because their high energy and extremely small mass allow them to pass through most materials. For those reasons, radionuclides that undergo radioactive decay by gamma radiation arguably represent one of the most effective fine-sediment tracers.

In the context of this technical report, an artificial radionuclide is defined as being of anthropogenic origin, but excludes those deposited by globally-distributed fallout, such as caesium-137 (^{137}Cs). In this respect, artificial radionuclides do not occur naturally in the wider environment (with the exception of licensed or accidental releases in local or regional areas). Typical methods for introducing an artificial radionuclide into a given environment when tracing fine-sediment are by: 1.) directly labelling surface-sediment in situ over a predefined study area; 2.) introducing quantities of 'pre-labelled' sediment onto or over a chosen study site; or 3.) conducting tracing experiments under a confined (i.e. laboratory) environment.

Health and safety

Due to inherent dangers associated with radioactive material, particularly unsealed sources, health and safety precautions must be followed in order to minimise the exposure-time to ionising radiation. A comprehensive guide to radiation protection is provided in Connor et al. (2007). In summary, the main risks associated with over-exposure can be alleviated by: 1.) shielding the radioactive source within an enclosed (i.e. lead-lined) container until required for use; 2.) minimising the time spent in proximity to a radioactive source; 3.) maintaining as much distance from the radioactive source as is practicable; 4.) wearing personal safety equipment and adhering to the Control of Substances Hazardous to Health (COSHH) protocol (HSE, 2012); and 5.) partitioning and diluting to a lower activity concentration (e.g. 100-200

Bq ml^{-1}) a sufficient quantity of radionuclide material from the stock supply to meet the needs of a given investigation.

Measuring radionuclides

The most common method for accurately measuring low-level radioactivity is by gamma spectrometry. This is a relatively simple procedure that involves minimal sample preparation and is non-destructive, which allows repeat measurements to be undertaken on individual samples. Field-based areal activity measurements can be performed *in-situ* using a field gamma spectrometer, shown in Figure 1. Although no sample preparation is required, the crucial factor for obtaining reliable results is to ensure that the radiometric protocol adopted at the beginning of an investigation remains consistent throughout the monitoring campaign (*cf.* He et al., 2002; Greenwood et al., 2008).



Figure 1. Using an *in-situ* gamma spectrometer to measure areal activity.

In contrast, estimating the mass activity of individual sediment samples is typically undertaken using laboratory-based, or static, gamma spectrometers (Figure 2). The key stages of sample preparation for this form of analysis involve drying, gently disaggregating and screening the sediment through a 2 mm dia. sieve. The material is then weighed usually to an accuracy of 0.1 g (Pennock and Appleby, 2002). For reasons of precision, samples presented to a detector should be of similar mass (i.e. within \pm ca. 0.1 g) and

radiometric assays should always be performed in containers of identical dimensions so that the sample-geometry, and hence the distance from the radioactive source to the detector-head, remain constant.



Figure 2. The interior of a static gamma spectrometer. The weighed sediment sample is placed, in a container, on the detector-head (fitted with a white end-cap).

Failure to comply with this protocol can profoundly influence the precision of individual measurements and ultimately produce erroneous results. A fuller explanation of the importance of sample presentation, and of other factors affecting analytical precision in gamma spectrometry is provided in Wallbrink et al. (2002). Where the mass, and hence the geometry, of samples varies, the ability of the detector(s) to 'capture' and record gamma-photons can also vary. In order to compensate and account for this effect, the geometry-efficiency of the detector must be re-calculated in order that sufficient analytical precision between samples of different mass is obtained (Wallbrink et al., 2002). This can be achieved by preparing a quantity (i.e. a set) of 'in-house' geometry-calibration standards. Each standard should increase in mass at regular (e.g. 5 g) intervals and the range of each set should be sufficient to correspond to the lightest and heaviest sediment samples obtained from a particular investigation. Geometry-calibration standards should, ideally, be labelled with the same radionuclide used in the investigation and at a similar mass activity (i.e. Becquerel per gram (Bq g^{-1})) as the soil(s) under investigation. Each standard should be analysed a minimum of, for instance, three

times over a suitable range of counting times (i.e. 1, 2 and 3 hours) to derive an accurate result. All results should be decay-corrected to a common date of analysis in order to account for the short half-life and the radioactive decay process. The mean activity value derived from each standard is then used as basis for re-calculating the geometry-efficiency of the detector at each mass. The appropriate detector efficiency value should be selected from the particular calibration standard whose mass most closely corresponds to the mass of the sample under analysis.

Benefits and advantages

In order to be of maximum value as a sediment tracer, an artificial radionuclide must possess certain environmental characteristics. Of particular importance is the need for a radionuclide to express a rapid and strong affinity for soil particles, in order to limit the effect of leaching, and have a relatively short half-life, in order to minimise environmental contamination (Fullen, 1982). Gamma-emitting radionuclides offer five distinct advantages that are readily exploitable when undertaking erosion investigations. The first relates to the relative ease with which energetic gamma-photons can be detected with a high degree of analytical sensitivity (Lang, 2008). The second relates to the highly penetrative nature of gamma-photons; coupled with the fact that only a relatively few are needed to identify the radionuclide and estimate the relative activity (Fullen, 1982). The third and potentially most powerful advantage relates to the extremely small mass of individual gamma-photons (Courtois, 1973; Sauzay, 1973; McCubbin and Leonard, 1995). In this respect, sediment labelled with a suitable artificial radionuclide for the purpose of tracing its movement arguably represents one of the most effective ways of precisely matching the tracer characteristics to the sediment being traced. Exploiting this attribute provides a (theoretical) means of obtaining an absolute indication of the movement of the eroded sediment without having unduly perturbed or adversely influenced its behaviour (Wooldridge, 1965; Fullen, 1982; Showler et al., 1988; McCubbin and Leonard, 1995). Once the correct environmental permits relating to the release

of unsealed sources of radioactive material have been obtained (refer to EA, 2012), a fourth advantage is that the initial dose-rate of an artificial radionuclide released over a given area can be tailored, within reason, to suit the objectives of an investigation. For example, if an objective is to measure erosion rates over very short (i.e. event-based) timescales, or in low-energy environments where soil movement is likely to be commensurately small, a high dose rate will allow radiometric assays to be performed relatively quickly and accurately, even on small samples, and also provide the best opportunity of obtaining results within an acceptable error range (i.e. usually $\pm 5-10\%$) (Sutherland, 1994; Wallbrink et al., 2002; Parsons and Foster, 2011). Lastly, an initially high dose rate applied directly to the soil-surface (at predetermined locations) within a given study-site can provide a means of lengthening the duration of an investigation so that erosion rates can continue to be documented over one or more seasons (see Wooldridge, 1965; Fullen, 1982; Syversen et al., 2001).

Taking account of the characteristics outlined above, artificial radionuclides offer a significant number of advantages over many other forms of tracer. With careful experimental design, these can be exploited and used to develop novel tracing-techniques that can be applied in a range of erosion scenarios where other tracers or existing tracing-techniques would be ineffective or lack the necessary level of sensitivity needed to provide reliable data.

Disadvantages and limitations

Arguably, one of the most vexing problems when using artificial radionuclide relates to a dearth of established techniques for directly labelling the surface-soil uniformly (Syversen et al., 2001). This limitation has two major implications. The first is that field-based investigations are typically limited to studying soil movement over areas equivalent to a few m^2 (Wooldridge, 1965; Fullen, 1982; Quine et al., 1999; Syversen et al., 2001). This constraint can be overcome, to a certain extent, by applying the radionuclide to small areas, or 'plots', within a larger study-site. This approach overcomes one of the major spatial constraints associated with using artificial radionuclides, since it represents a

convenient and effective way of obtaining contiguous erosion data from plots strategically sited at key areas, or on particular landform-features within a study-site (e.g. Greenwood et al., 2008).

The second limitation associated with the non-uniform application of radionuclide material means that the pre-event areal activity at any location within the labelled area is unknown and must, as a vital prerequisite, be accurately measured before the next erosion event. Importantly, all post-event measurements must thereafter be performed at precisely the same locations as the pre-event measurements, in order to permit changes in the areal activity, and hence, in the movement or redistribution of the labelled soil, to be accurately quantified.

Arguably, one of the most profound disadvantages with using artificial radionuclides arises from the stigma associated with radioactivity, and the fact that field-based investigations inevitably involve deliberately releasing radioactive material into an environment. This problem can be mitigated against, to a certain extent by, firstly, adhering to the precept that any exposure to ionising radiation should be kept as low as reasonably possible (Connor et al., 2007), and secondly, by selecting an artificial radionuclide with an appropriate (i.e. short) half-life. Artificial radionuclides with a suitably contrasting half-life range that have been successfully used in sediment-tracing investigations include iron-59 (^{59}Fe ; $t_{0.5} = 45$ d) (Wooldridge, 1965; Fullen, 1982), caesium-134 (^{134}Cs ; $t_{0.5} = 2.06$ yrs.) (Quine et al., 1999; Syversen et al., 2001; Greenwood et al., 2008), and cobalt-60 (^{60}Co ; $t_{0.5} = 5.27$ yrs) (Toth and Alderfer, 1960; Greenwood et al., 2008).

Any researcher considering using artificial radionuclides in sediment tracing investigations must also ensure that their research institute has the necessary radioisotope permits and adequate facilities to purchase, securely store and safely dispose of radioactive waste. The researcher will also need access to comprehensive analytical facilities, or be prepared to pay for analytical time on radiometric facilities elsewhere, and be trained to handle and partition the unsealed source material to an appropriate concentration. Lastly, a vital requirement when conducting field-based investigations using artificial radionuclides is

the need to identify landowners willing to allow radioactive material to be deliberately introduced onto their land.

Practical applications

Examples of three field / laboratory-based tracing techniques using the artificial radionuclides, ^{134}Cs and ^{60}Co , are now briefly described, in order to demonstrate the novel or diverse ways in which they can be applied.

Tracing eroded earthworm casts on pasture

The presence of earthworm casts on soil-surfaces has led many workers to speculate that the dispersed sediment can potentially contribute to soil erosion (Le Bayon and Binet, 1999, 2001; Le Bayon et al., 2002). Despite efforts to confirm or refute those claims, a notable degree of uncertainty still exists and is attributed, as described earlier, to the difficulty of accurately tracing fine-sediment. Capitalising on this continuing uncertainty, a technique was successfully developed for labelling earthworm casts with either ^{134}Cs or ^{60}Co , for the purpose of tracing the dispersed sediment. Labelling was achieved by immersing the intact air-dried casts into a solution of water, mixed with a known concentration of either ^{134}Cs or ^{60}Co . After labelling, the casts were deployed across a small prepared area of pasture in a semi-moist condition, which was considered to be most representative of the stability, and hence, the potential erodibility, of casts of a wide range of ages (Binet and Le Bayon, 1999). In order to provide information on relative travel distance of the dispersed sediment, a batch of ^{134}Cs -labelled casts was evenly distributed across the upslope half of the plot, at a distance of $> 0.3\text{ m}$ from the channel outlet, and a batch of ^{60}Co -labelled casts was evenly distributed across the downslope half of the plot, at a distance of $< 0.3\text{ m}$ from the channel outlet (Figure 3). All casts were subjected to natural weather events over a two-month period. Runoff samples were collected and the eroded sediment was separated from the liquid, processed following the method described earlier and radiometrically assayed. A mass balance was used to partition and quantify the relative proportions of labelled sediment from the unlabelled surface soil. The

response of both tracers indicate that, in proportional terms, 14.1% and 4.5% of the original mass of ^{60}Co and ^{134}Cs -labelled casts were transported downslope by surface runoff over distances $< 0.3\text{ m}$ and $> 0.3\text{ m}$, respectively. The results provide essentially unique information that confirms that earthworm casts can contribute to soil erosion, certainly over relatively short distances, on vegetated hillslopes.

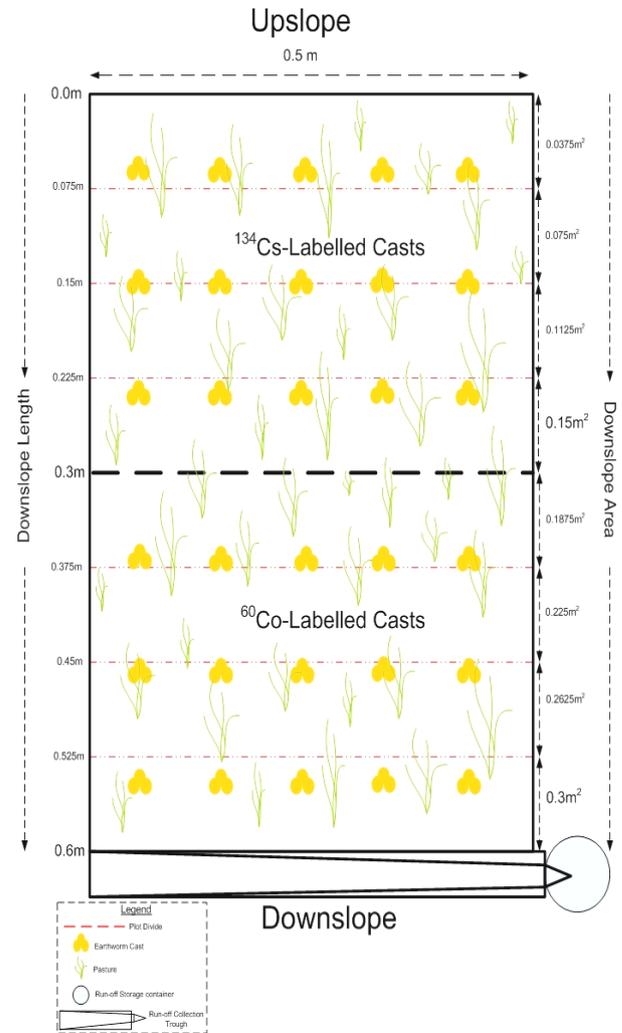


Figure 3. The design-layout and relative positions of each batch of labelled casts within the erosion plot (not to scale).

Measuring soil redistribution on livestock-poached pasture

Recent developments in fingerprinting and sediment-source tracing techniques have demonstrated that the contribution of fine-sediment from pasture is considerably greater than originally assumed (Collins et al., 1997; Russell et al., 2001). Stimulated by

those findings, numerous studies seeking to identify potential sediment-sources (Brazier et al., 2007; Bilotta et al., 2008) have singled-out the effect of poaching by livestock (Warren et al., 1986a, 1986b; Pietola et al., 2005; Haygarth et al., 2006; Bilotta et al., 2007, 2008). Poaching occurs where livestock congregate, such as on the banks of watercourses, around feed-troughs and in gate-entrances (Figure 4). This increases the susceptibility of the soil to erosion, particularly by rainsplash and surface runoff, due to its degraded structure (Kauffman and Krueger, 1984; Warren et al., 1986b).



Figure 4. An example of a heavily poached gateway.

Little work has been done to quantify the contribution of sediment from poached areas, however, owing to a paucity of tracing-techniques capable of measuring its movement at the required spatial and temporal scales (Granger et al., 2007). A field-based technique was thus developed to quantify rates of soil redistribution on small areas of poached pasture over a series of rainfall events. The approach was adapted from an established technique for estimating catchment-scale erosion rates in undisturbed soils using the fallout radionuclide, ^{137}Cs (Walling and Quine, 1990, 1992; Walling et al., 2002). The approach involved directly labelling small plots of poached soil with a known activity of either ^{134}Cs or ^{60}Co . Predetermined measuring points were established across each plot and the areal activity at each point was estimated before and after periods of rainfall. By documenting

spatial variations in the areal activity before and after periods of rainfall, quantitative soil redistribution values could be estimated at each of the measuring points within each plot using an established mathematical accounting procedure (Loughran, 1989; Walling and Quine, 1990, 1992; Walling et al. 2002).

By the end of the 65 d monitoring period, all plots recorded a mean net reduction in the overall soil-depth ranging between -6.8 to -15.2 mm. In quantitative terms, these values represent an equivalent mean soil loss of between 0.134 to 0.302 kg d⁻¹. Overall, these findings provide evidence that areas of livestock-poached pasture can represent a substantial source of eroded sediment.

Determining the grain-size composition of inter-rill and rill-eroded sediment

Many attempts have been made to determine whether contrasting hydraulic conditions associated with inter-rill and rill erosion are reflected in the grain-size selectivity and hence, in the particle-size composition of eroded material (e.g. Young and Onstad, 1978; Alberts et al., 1980; Farenhorst and Bryan, 1995; Fox and Bryan, 1999; Chaplot and Le Bissonais, 2000; Basic et al., 2002; Legu dois and Le Bissonais, 2004; Yang et al., 2006). Attributing the erosion of sediment to a particular process is extremely difficult, however (Yang et al., 2006), and where attempts have been made, the results have often been contradictory. Against this uncertainty, a stratified tracing technique, using both radionuclides simultaneously, was developed to identify changes in sediment-source during the transition from inter-rill to rill erosion. The approach was adapted from techniques developed by Wallbrink and Murray (1993), Wallbrink et al. (1999), Whiting et al. (2001) and Yang et al. (2006), all of whom exploited the contrasting depth-distribution of varying fallout radionuclides to determine the depth-origin of eroded sediment and predict the mobilisation process responsible. Conducting a similar investigation using artificial radionuclides, however, afforded a level of convenience that permitted testing eight different agricultural soils under laboratory-conditions relatively rapidly, and under a standard suite of eroding conditions. The approach involved filling a small erosion plot with a soil, pre-labelled

with a known activity of ^{134}Cs , to a depth of ca. 0.1 m. This was covered with an additional 0.01 m layer of the same soil, pre-labelled with a known activity of ^{60}Co . The principle behind the 'stratified' configuration, seen in Figure 5, capitalises on erosion as a time-variant process (Yang et al., 2006), during which, the early stages are typically dominated by inter-rill erosion which preferentially removes fine material almost exclusively from the surface due to its low transport capacity (Alberts et al., 1980; Abrahams et al., 1998; Parsons et al., 1998; Legu dois and Le Bissonais, 2004; Ahmadi et al., 2006).

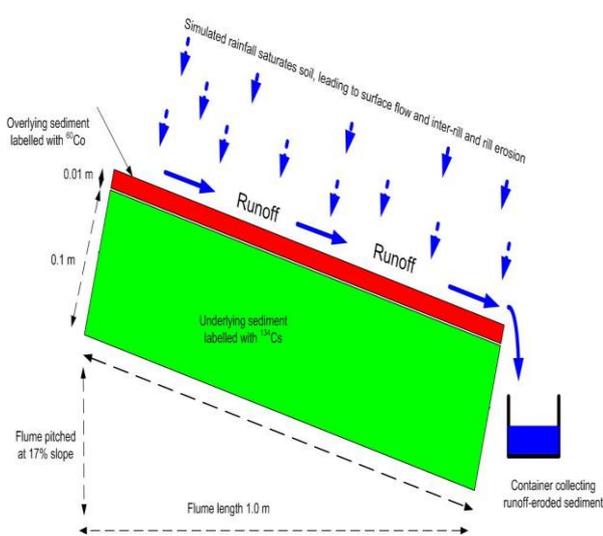


Figure 5. The stratified configuration of the two layers of labelled soil in the erosion plot.

As the rainfall event proceeds in time, surface-flow becomes concentrated, rill erosion becomes the predominant mobilisation process (Figure 6) and the increased transport capacity erodes coarser material when compared with inter-rill eroded sediment. In order to test this hypothesis, simulated rainfall was applied to each soil at a sufficient intensity and duration to initiate inter-rill erosion leading to rill erosion. Runoff was collected continuously over equal time-intervals and the sediment was separated from the liquid, processed following the method described earlier and radiometrically assayed. By examining changes in the radiometric signals of individual samples, it was possible to identify the transition from inter-rill to rill erosion and hence, the change from surface to sub-surface sediment-source. The radiometric data provided a basis by

which a pair of samples were selected from each soil that were considered to be most characteristic of each mobilisation process.

The particle size composition of each pair of samples from each soil was determined and the grain-size was measured at corresponding 10 percentile intervals and statistically tested to determine whether the size-difference was significant. The results indicate that rill-eroded sediment was significantly coarser than inter-rill eroded material for the majority of soils investigated.



Figure 6. A partially developed rill-system.

Conclusion

Artificial gamma-emitting radionuclides possess a number of unique characteristics that are readily exploitable when undertaking sediment tracing investigations. With careful experimental design, these can be utilised to develop novel or diverse tracing techniques to investigate certain erosion scenarios where other tracers or existing tracing-techniques would be ineffective or lack the sensitivity needed to provide accurate data at the required spatial and temporal scales. This conclusion is supported by the findings from the three investigations described.

References

- Abrahams AD, Li G, Krishnan C, Atkinson JF. 1998. Predicting sediment transport by interrill overland flow on rough surfaces. *Earth Surface Processes and Landforms* **23**: 1087-1099.
- Ahmadi SH, Amin S, Keshavarzi AR, Mirzamostafa N. 2006. Simulating watershed outlet sediment concentration using the ANSWERS Model by applying two sediment transport capacity equations. *Biosystems Engineering* **94**: 615-626.
- Alberts EE, Moldenhauer WC, Foster GR. 1980. Soil aggregates and primary particles transported in rill and interrill flow. *Soil Science Society of America Journal* **44**: 590-595.
- Basic F, Kistic I, Mesic M, Butorac A. 2002. Particle size distribution (texture) of eroded soil material. *Journal of Agronomy and Crop Science* **188**: 311-322.
- Bilotta GS, Brazier RE, Haygarth PM. 2007. The impacts of grazing animals on the quality of soils, vegetation and surface waters in intensively managed grasslands. *Advances in Agronomy* **94**: 237-280.
- Billota G, Brazier RE, Haygarth PM, Macleod CJA, Butler P, Granger S, Krueger T, Freer J, Quinton J. 2008. Rethinking the contribution of drained and un-drained grasslands to sediment-related water quality problems. *Journal of Environmental Quality* **37**: 906-914.
- Binet F, Le Bayon RC. 1999. Space-time dynamics in situ of earthworm casts under temperate cultivated soils. *Soil Biology and Biochemistry* **31**: 85-93.
- Brazier RE. 2004. Quantifying soil erosion by water in the UK: a review of monitoring and modelling approaches. *Progress in Physical Geography* **28**: 340-365.
- Chaplot V, Le Bissonnais Y. 2000. Field measurements of interrill erosion under different slopes and plot sizes. *Earth Surface Processes and Landforms* **25**: 145-153.
- Collins AL, Walling DE, Leeks GJL. 1997. Source type ascription for fluvial suspended sediment based on a quantitative composite fingerprinting technique. *Catena*, **29**: 1-27.
- Connor KJ, McLintock IS, Heaton B, Moseley TJ. 2007. *Radiation Protection: Handbook for Laboratory Workers* (ed. TJ Moseley). Association of University Radiation Protection Officers, University of Sheffield, UK.
- Courtois G. 1973. The use of artificial radioactive tracers in France. In: *Tracer Techniques in Sediment Transport*. Technical Report Series No. **145**: 71-80. International Atomic Energy Agency (IAEA), Vienna, Austria.
- EA (Environment Agency). 2012. *Radioactive Substances Regulation (RSR) – Environmental Permitting Guidance*. The Environment Agency, UK. <http://www.environment-agency.gov.uk/business/sectors/117039.aspx> (accessed 05/01/12).
- Evans GV. 1983. Tracer techniques in hydrology. *The International Journal of Applied Radiation and Isotopes* **34**: 451-475.
- Farenhorst A, Bryan RB. 1995. Particle size distribution of sediment transported by shallow flow. *Catena* **25**: 47-62.
- Foster IDL, Lees JA. 2000. Tracers in geomorphology: Theory and application in tracing fine particulate sediments. In: *Tracers in Geomorphology* (ed. IDL Foster), **Chpt. 1**: 3-21. John Wiley & Sons Ltd. Chichester, England.
- Fox DM, Bryan RB. 1999. The relationship of soil loss by interrill erosion to slope gradient. *Catena* **38**: 211-222.
- Fullen MA. 1982. Laboratory and field studies in the use of the isotope ⁵⁹Fe for tracing soil particle movement. *Earth Surface Processes and Landforms* **7**:285-293.
- Granger SJ, Bol R, Butler PJ, Haygarth PM, Naden P, Old G, Owens PN, Smith BPG. 2007. Processes affecting transfer of sediment and colloids, with associated phosphorous, from intensively farmed grasslands: tracing and organic matter. *Hydrological Processes* **21**: 417-422.
- Greenwood P, Walling DE, Quine TA. 2008. Assessing the remobilisation of recently deposited sediment from river flood plains during single overbank flood events, using caesium-134 and cobalt-60 as tracers. *Proceedings of the Symposium on Sediment Dynamics in Changing Environments*, Christchurch, New Zealand. IAHS Publication No. **325**: 13-22.
- Guirese M, Revel SC. 1995. Erosion due to cultivation of calcareous clay soils on hillsides

in South-West France. II. Effect of ploughing down the steepest slope. *Soil & Tillage Research* **35**: 157-166.

Haygarth PM, Bilotta GS, Bol R, Brazier R.E, Butler PJ, Freer J, Gimbert J, Granger SJ., Kreuger T, Macleod CJA, Naden P, Old G, Quinton JN, Smith B, Worsfold P. 2006. Processes affecting transfer of sediment and colloids, with associated phosphorous, from intensively farmed grasslands: an overview of key issues. *Hydrological Processes* **20**: 4407-4413.

He Q, Walling DE, Wallbrink PJ. 2002. Alternative Methods and Radionuclides for Use in Soil-Erosion and Sedimentation Investigations. Chpt. 9: 185-215. In: *Handbook for the Assessment of Soil Erosion and Sedimentation Using Environmental Radionuclides* (F. Zapata, ed.). Kluwer Academic Publishers: London.

HSE (Health & Safety Executive). 2012. Control of Substances Hazardous to Health (COSHH). <http://www.hse.gov.uk/coshh/> (accessed 05/01/12).

Jackson ARW, Jackson JM. 2000. *Environmental Science: The Natural Environment and Human Impact* 2nd ed. Pearson Education Ltd. Essex: UK.

Jones L, Atkins P. 2000. *Chemistry: Molecules, Matter and Change* 4th ed. WH Freeman & Co. New York.

Kauffman JB, Krueger WC. 1984. Livestock impacts on riparian ecosystems and streamside management implication: A review. *Journal of Range Management* **3**: 430-438.

Lang A. 2008. Recent advances in dating and source tracing of fluvial deposits. *Proceedings of the Symposium on Sediment Dynamics in Changing Environments*, Christchurch, New Zealand, Dec. 2008. IAHS Publication No. **325**: 3-12.

Le Bayon RC, Binet F. 1999. Rainfall effects on erosion of earthworm casts and phosphorous transfers by water runoff. *Biological Fertility of Soils* **30**: 7-13.

Le Bayon RC, Binet F. 2001. Earthworm surface casts affect soil erosion by run-off and phosphorous transfer in a temperate maize crop. *Pedobiologia* **45**: 430-442.

Le Bayon RC, Moreau S, Gascuel-Oudou C, Binet F. 2002. Annual variations in earthworm

surface-casting activity and soil transport by water run-off under a temperate maize agroecosystem. *Geoderma* **106**: 121-135.

Legu dois S, Le Bissonais Y. 2004. Size fractions resulting from an aggregate stability test, interrill detachment and transport. *Earth Surface Processes and Landforms* **29**: 1117-1129.

Loughran RJ. 1989. The measurement of soil erosion. *Progress in Physical Geography* **13**: 216-233.

McCubbin D, Leonard KS. 1995. Use of radiotracers for studies of metal sorption behaviour. *Science of the Total Environment* **173/174**: 259-266.

Parsons AJ, Foster IDL. 2011. What can we learn about soil erosion from the use of ¹³⁷Cs? *Earth Science Reviews* **108**: 101-113.

Parsons AJ, Stromberg SGL, Greener M. 1998. Sediment-transport competence of rain-impacted interrill overland flow. *Earth Surface Processes and Landforms* **23**: 365-375.

Pennock DJ, Appleby PG. 2002. Sample Processing, Chpt. 4: 59-65. In: *Handbook for the Assessment of Soil Erosion and Sedimentation Using Environmental Radionuclides* (F Zapata, ed.). Kluwer Academic Publishers, London, UK.

Pietola L, Horn R, Yli-Halla M. 2005. Effects of trampling by cattle on the hydraulic and mechanical properties of soil. *Soil & Tillage Research* **82**: 99-108.

Polyakov VO, Nearing MA. 2004. Rare earth oxides for tracing sediment movement. *Catena* **55**: 255-276.

Quine TA, Govers G, Poesen J, Walling DE, van Wesemael B, Martinez-Fernandez J. 1999. Fine-earth translocation by tillage in stony soils in the Guadalentin, south-east Spain: an investigation using caesium-134. *Soil & Tillage Research* **51**: 279-301.

Riebe B. 1995. Monitoring the translocation of soil particles using a neutron activated tracer. In: *Soil Structure: Its Development and Function* (eds. K. H. Hartge & B. A. Stewart): 277-294. CRC Press Inc., Florida, USA.

Russell MA, Walling DE, Hodgkinson RA. 2001. Suspended sediment courses in two small lowland agricultural catchments in the UK. *Journal of Hydrology* **252**: 1-24.

- Sauzay G. 1973. Tracer techniques in sediment transport: Report of the panel. In: *Tracer Techniques in Sediment Transport*. Technical Report Series No. **145**: 3-8. International Atomic Energy Agency (IAEA). Vienna, Austria.
- Sutherland RA. 1994. Spatial variability of ¹³⁷Cs and the influence of sampling on estimates of sediment redistribution. *Catena* **21**: 57-71.
- Syversen N, Øygarden L., Salbu B. 2001. Cesium-134 as a tracer to study particle transport processes within a small catchment with a buffer zone. *Journal of Environmental Quality* **30**: 1771-1783.
- Toth SJ, Alderfer RB. 1960. Formation and breakdown of Co60-tagged water-stable soil aggregates in a Norton Silt Loam Soil. *Soil Science* **90**: 232-238.
- Ventura E, Nearing MA, Amore E, Norton LD. 2002. The study of detachment and deposition on a hillslope using a magnetic tracer. *Catena* **48**: 149-161.
- Wallbrink PJ, Murray AS. 1993. Use of fallout radionuclides as indicators of erosion processes. *Hydrological Processes* **7**: 297-304.
- Wallbrink PJ, Murray AS, Olley JM. 1999. Relating suspended sediment to its original soil depth using fallout radionuclides. *Soil Science Society of America Journal* **63**: 369-378.
- Wallbrink PJ, Walling DE, He Q. 2002. Radionuclide Measurement Using HPGe Gamma Spectrometry. Chpt. 5: 67-96. In: *Handbook for the Assessment of Soil Erosion and Sedimentation Using Environmental Radionuclides* (F. Zapata, ed.). Kluwer Academic Publishers: London.
- Walling DE. 2004. Quantifying the fine sediment budgets of river basins. *National Hydrology Seminar*: 9-20.
- Walling DE, Owens PN, Carter J, Leeks GJL., Lewis S, Meharg AA, Wright J. 2003. Storage of sediment-associated nutrients and contaminants in river channel and floodplain systems. *Applied Geochemistry* **18**: 195-220.
- Walling DE, Quine TA. 1992. The use of caesium-137 measurements in soil erosion surveys. In: *Erosion and Sediment Transport Monitoring Programmes in River Basins*. International Association of Hydrological Sciences (IAHS), Publication No. **210**: 143-152.
- Walling DE, Quine TA. 1990. Calibration of caesium-137 measurements to provide quantitative erosion rate data. *Land Degradation & Rehabilitation* **2**: 161-175.
- Warren S D, Nevill MB, Blackburn WH, Garza NE. 1986a. Soil response to trampling under intensive rotation grazing. *Soil Science Society of America Journal* **50**: 1336-1341.
- Warren SD, Thurow TL, Blackburn WH, Garza NE. 1986b. The influence of livestock trampling under intensive rotation grazing on soil hydraulic characteristics. *Journal of Range Management* **39**: 491-495.
- Whiting PJ, Bonniwell EC, Matisoff G. 2001. Depth and areal extent of sheet and rill erosion based on radionuclides in soils and suspended sediment. *Geology* **29**: 1131-1134.
- Wooldridge DD. 1965. Tracing soil particle movement with Fe-59. *Soil Science Society of America Journal* **29**: 469-472.
- Yang MY, Walling DE, Tian JL, Liu PL. 2006. Partitioning the contributions of sheet and rill erosion using Beryllium-7 and Cesium-137. *Soil Science Society of America Journal* **70**: 1579-1590.
- Young RA, Onstad C A. 1978. Characterization of rill and interrill eroded soil. *Transactions of the American Society of Agricultural Engineers* **21**: 1126-1130.
- Zhang XC, Nearing MA, Polyakov VO, Friedrich JM. 2003. Using rare-Earth oxide tracers for studying soil erosion dynamics. *Soil Science Society of America Journal* **67**: 279-288.

3.6.1. Karst Landform Classification Techniques

John McIlroy de la Rosa¹

¹School of Geography, Archaeology and Palaeoecology, Queen's University Belfast (jmcilroydelarosa01@qub.ac.uk)



ABSTRACT: Karst landform classification is fraught with problems because of a combination of synonymic terminology and the polygenetic nature of many karren features. The genetic classification of Bögli (1960) and morphogenetic classification proposed by Ford and Williams (1989; 2007) remain the most comprehensive and widely embraced by the academic community. A morphometric approach has predominated within karst studies in a bid to characterise and classify karst landforms at a variety of scales, as well as to understand and simulate their development. However, discrepancies between methods of measurement create difficulties in the global comparison of data between authors. Future morphometric techniques in karst studies are likely to evolve from the collection of data via complex and laborious field measurement methodologies, to the use of ergonomic, fast and sub-millimetre accurate LIDAR systems with associated GIS analysis.

KEYWORDS: Karst, Karren, Morphometry, Classification.

Introduction

The objective of this paper is twofold. Firstly, this paper seeks to briefly explore how karst landforms have been classified, the inherent difficulties involved in their classification, and the benefits and shortfalls of the most common classification systems.

Secondly, common techniques employed in the mapping and morphometric analysis of karst landforms at a variety of scales are outlined. Research techniques to date are summarised with some potential research directions and promising research techniques identified. The methods and techniques used in karst studies vary greatly due to the range of scales of investigation within which karst studies may be undertaken. Scientists have investigated karstic morphologies from the larger landscape scale down to the individual mineral scale, with an increasingly reductionist approach in recent years. As such, this review of analytical techniques will adopt a scaled approach, within which relevant techniques are organised into the

scale boundaries in which they are most applicable.

Due to their relevance to geomorphological studies, particular emphasis is placed on surface karst features, in particular karren landforms. In order to familiarise the reader with noteworthy research to date, reference is made to noteworthy literature throughout this paper. For wider reading, key texts by Ford and Williams (2007) and Ginés *et al.*, (2009) review the process of karstification in detail.

The classification of karst landforms

The nature of karst systems dictates that most dissolution is expended in the epikarst, i.e. at the interface between karst rocks and the atmosphere, the source of meteoric water. Geomorphologists, concerned with landforms and the earth surface processes that form them, seek to understand a range of karst landforms and landform assemblages at a variety of scales.

Table 1. Simple classification of karren and karst surface landforms according to scale.

Scale Boundary	Karst Terminology	Examples in Karst
Macro (c.>10m)	Large Scale Karst Landforms	Polje (Gams, 1973; 1978; Bonacci, 2004), Uvala (Ćalić, 2011), Doline (Sauro, 2003; Ford and Williams, 2007).
	Karrenfield Landscapes and Karren Assemblages	Limestone pavements (Williams, 1966; Zseni <i>et al.</i> , 2003), Megaausgleichsfläche (Kunaver, 2009), Pinnacle karst (Knez <i>et al.</i> , 2003), Ruiniform karst (Perna and Sauro, 1978), Schichttreppenkarst (Bögli, 1980), Corridor / Labyrinth karst (Jennings and Sweeting, 1963; Brook and Ford, 1978), Tower karst (Sweeting, 1995), Cockpit / Cone karst (Xiong, 1992; Lyew-Ayee <i>et al.</i> , 2007; Fleurant <i>et al.</i> , 2008).
Meso (c.1cm-10m)	Karren	Rillenkarren (Glew and Ford, 1980; Lundberg and Ginés, 2009), Rinnenkarren (Bögli, 1976; Veress, 2009a), Kluftkarren (Williams, 1966; Goldie, 2009), Flachkarren (Williams, 1966; Goldie, 2009), Mäanderkarren (Veress, 2000; 2009b), Kamenitze (Bögli, 1960; Cucchi, 2009; Eren and Hatipoglu-Bagci, 2010), Ausgleichsfläche (Bögli, 1980), Trittkarren (Veress, 2009c), Grübchenkarren (White, 1988; Ginés and Lundberg, 2009), Rundkarren (Zseni, 2009), Tumuli (Calaforra and Pulido-Bosch, 1999).
Micro (c.1mm-1cm)	Microkarren	Dissolutional: Microrills (Ford and Lundberg, 1987; Gómez-Pujol and Fornós, 2009; Grimes, 2007), Rillensteine (Laudermilk and Woodford, 1932), Rainpits (Ginés and Lundberg, 2009). Biological: Biotroughs (McIlroy de la Rosa <i>et al.</i> , 2011).
Nano (c.<1mm)	Nanokarren	Dissolutional: V-in-V etching, Stepped etching, Crystal widening, Cleavage widening (Moses <i>et al.</i> , 1995; Viles and Moses, 1998). Biological: Endolithic lichen biopits (McIlroy de la Rosa <i>et al.</i> , 2011), Filament-shaped trenches, Circular etch pits (Viles and Moses, 1998).

Table 2. Genetic classification of karren forms after Bögli (1960).

Free Karren	Half-Free Karren	Covered Karren
<i>Rillenkarren</i> (solution flutes) <i>Trittkarren</i> (heelprint karren) <i>Rinnenkarren</i> (solution runnels) <i>Mäanderkarren</i> (meandering decantation runnels) <i>Wandkarren</i> (decantation runnels) <i>Kluftkarren</i> (grikes) <i>Flachkarren</i> (clints) <i>Grübchenkarren</i> (rainpits)	<i>Kamenitza</i> (solution basins / pans) <i>Korrosionskehlen</i> (solution notches) <i>Hohlkarren</i> (undercut solution runnels)	<i>Kavernösen Karren</i> (cavernous subsoil weathering) <i>Rundkarren</i> (rounded solution runnels) <i>Geologische Orgeln</i> (solution pipes, pits and shafts)

Table 3. Morphogenetic classification of karren forms after Ford and Williams (1989; 2007).

Circular plan forms
<p><i>Micropits and etched surfaces</i> – wide variety of pitting and differential etching forms commonly less than 1.0cm in characteristic dimension.</p> <p><i>Pits</i> – circular, oval, irregular plan forms with rounded or tapering floors, >1cm in diameter.</p> <p><i>Pans</i> – rounded, elliptical, to highly irregular plan forms: planar, usually horizontal floors in bedrock or fill, >1cm in diameter.</p> <p><i>Heelprints or trittkarren</i> – arcuate headwall, flat floor, open in downslope direction. Normally 10-30cm in diameter.</p> <p><i>Shafts or wells</i> – connected at bottom to proto caves / small caves draining into epikarst. Great range of form.</p>
Linear forms (fracture controlled)
<p><i>Microfissures</i> – microjoint guided, normally tapering with depth. May be several centimetres long but rarely more than 1cm deep. Transitional to:</p> <p><i>Splitkarren</i> – joint-, stylolite-, or vein-guided solution fissures. Taper with depth unless adapted for channel flow. From centimetres to several metres in length, centimetres deep. Closed type terminates on fracture at both ends. Open type terminates in other karren at one or both ends.</p> <p><i>Grikes or kluffkarren</i> – Major joint-, or fault-guided solutional clefths. Normally 1-10m in length. Master features in most karren assemblages, segregating clint blocks (<i>flachkarren</i>) between them. Scale up to karst bogaz, corridors, streets etc. Subsoil forms are termed cutters.</p>
Linear forms (hydrodynamically controlled)
<p><i>Microrills</i> – as on <i>rillensteine</i>. Rill width is c.1mm. Flow is controlled by capillary forces and / or gravity and / or wind.</p> <p><i>Gravitomorphic solution channels</i></p> <p><i>Rillkarren</i> – packed channels commencing at crest of slope; 1-3cm wide. Extinguish downslope. Rainfall-generated, no decantation.</p> <p><i>Solution runnels</i> – Hortonian channels commencing below a belt of no channelled erosion. Sharp-rimmed on bare rock (<i>Rinnenkarren</i>), rounded if subsoil (<i>Rundkarren</i>). Channels enlarge downslope. Normally 3-30cm wide, 1-10m long. Linear, dendritic or centripetal channel patterns.</p> <p><i>Decantation runnels</i> – solvent is released from an upslope, point-located store. Channels reduce in size downslope. Many varieties and scales up to 100m in length, e.g. wall karren (<i>Wandkarren</i>), <i>Mäanderkarren</i>.</p> <p><i>Decantation flutings</i> – solvent is released from a diffuse source upslope. Channels are packed; may reduce downslope. 1-50cm wide.</p> <p><i>Scallop forms or solution ripples</i> – ripple-like flutes oriented normal to direction of flow. A variety of scallop. Prominent as a component of <i>cockling patterns</i> on steep, bare slopes.</p>
Polygenetic forms
<p>Mixtures of solution channels with pits, pans, wells and splitkarren. Subsequent development of <i>Hohlkarren</i>, <i>Spitzkarren</i> and subsoil <i>pinnacles</i>. Superimposition of small forms (microrills, rillkarren, small pits) upon larger forms</p>
Assemblages of karren
<p><i>Karrenfeld</i> – general term for exposed tracts of karren.</p> <p><i>Limestone pavement</i> – a type of karrenfeld dominated by regular clints (<i>flachkarren</i>) and grikes (<i>kluffkarren</i>). Stepped pavements (<i>Schichttreppenkarst</i>) when benched.</p> <p><i>Pinnacle karst</i> – pinnacle topography on karst rocks, sometimes exposed by soil erosion, Arête-and-pinnacle, stone forest ,etc. with pinnacles up to 45m high and 20 wide at base.</p> <p><i>Ruiniform karst</i> – wide grike and degrading clint assemblage exposed by soil erosion. Transitional to <i>tors</i>.</p> <p><i>Corridor karst</i> – (or <i>labyrinth karst</i>, <i>giant grikeland</i>): scaled-up clint-and-grike terrains with grikes several metres or more in width and up to 1km in length.</p> <p><i>Coastal karren</i> – distinctive coastal and lacustrine solutional topography on limestone or dolomite. Boring and grazing marine organisms may contribute. Includes intertidal and subtidal notches, and dense development of pits, pans and micropits.</p>

The term *karren* encompasses a complex group of small to medium-sized exokarstic (karst features developed on surfaces exposed to direct precipitation) and cryptokarstic (karst features developed beneath permeable sediments such as soil or till) landforms, showing a great variety of characteristic shapes and forms (Ginés, 2009). Many of these forms are believed to be polygenetic in nature. Consequently, like karst terminology, the classification of karren forms is complex.

A simple classification of karst forms may be devised according to scale. Table 1, compiled by the author, displays karst features of interest to the geomorphologist organised into the major scale boundaries. References to key texts dealing with specific landforms are also included. The proposed classification organises karst landforms into categories of decreasing size and complexity. Macroscale features include karrenfield landscapes and assemblages of karren after Ginés (2009), as well as large scale karst landforms (non-karren). Mesoscale forms encompass the elementary karren forms described by Ginés (2009), Bögli (1980) and Ford and Williams (2007). At the nano and microscales, surface features are differentiated into those which are dissolutional in origin and those which develop due to biological activity. At larger scales, this biotic / abiotic differentiation becomes blurred as many karren features are believed to be largely dissolution-induced but may sometimes develop with biological contribution, for example rillenkarren (Fiol *et al.*, 1996) and kamenitze (McIlroy de la Rosa *et al.*, 2012). A scaled approach to classification permits a useful organisation of landforms according to dimension, but cannot easily encompass landform morphology or genesis. An alternative classification of karstic forms according to scale, with some reference to landform genesis, is provided by Ginés (2009).

According to Bögli (1980), the multiplicity of possible karren landforms makes a morphological classification system endless, while a genetic one allows a meaningful grouping. Bögli (1960) produced the most comprehensive genetic classification system of karren forms. He distinguishes between features that form on bare, exposed rock (free karren), partly covered rock (half free

karren), or rock completely covered by soil or dense vegetation (covered karren) (Table 2).

White (1988) also adopts a genetic approach to classification based on the angle of slope, rock structure, and whether the karst surface is covered or bare. He classifies karren features according to their relationship to water flow. However, only individual karren forms fit within his categories, with polygenetic and coalescing forms unable to be included. Whilst agreeing with the benefits of a genetic classification over a morphological one, Ford and Williams (1989; 2007) argue that the development of many karren features is not yet sufficiently understood to permit a purely genetic classification. Consequently, they propose a morphological classification with five primary classes broadly defined in terms of genetic criteria (Table 3). To conclude, the classification of karren landforms is inherently difficult due to the polygenetic origin and morphology of many forms. Ginés (2009) argues that a variety of criteria constitute a valid base for classification, and that one classification or another should be adopted depending on the problem being addressed.

Methods in karst landform classification

The use of morphometric techniques in karst research has revealed karren features and karst landform assemblages to be highly organised systems, and not chaotic as was previously believed (Ford and Williams, 2007). As a simple classification of karst landforms may be drawn up according to scale (Table 1), scale divisions are also a useful way to categorise the methods and techniques used in karst landform classification. Figure 1, compiled by the author, shows the applicability of various analytical methods and techniques used in karst studies within the major scales of investigation. By their very nature, some methods may only be used at a defined scale, while others may traverse scale boundaries and be utilised at a variety of scales.

grikes on limestone pavements or assemblages of limestone pinnacles.

Ground surveys at the macroscale have traditionally involved field measurements in order to complement and verify cartographic sources and remotely sensed data. Today, manual field measurement and the use of optical theodolites have largely been replaced by the use of total stations and terrestrial LIDAR systems, with the latter having significant morphometric potential in karst studies. Terrestrial laser scanning is a portable, non-destructive, accurate and fast method of 3D data capture. Already employed in the scanning and 3D digital representation of karst cave systems and other subsurface cavities (Rüther *et al.*, 2009; Canevese *et al.*, 2011), terrestrial LIDAR has, as yet, under-explored possibilities in the digital documentation and derivation of high resolution morphometry of karren features, assemblages of karren features and karst landforms such as dolines, with few studies having applied the technology to this research area (Siart *et al.*, 2011). High resolution digital elevation models derived from laser scanning allow the production of detailed topographical analyses, providing a detailed insight into macrokarren, mesokarren and microkarren morphometry. Crucially, the integration of laser-scanned data into Geographical Information Science (GIS) software allows the accurate derivation of topographical parameters such as slope gradient, volume, aspect or surface drainage patterns (Siart *et al.*, 2011). Indeed, such a GIS based approach is applicable at a range of scales of investigation from the macroscale to the microscale.

Finally, at the macroscale, resistivity surveys and ground penetrating radar (GPR) may also be used to identify and map subsoil karren features. GPR has been successfully applied to the locating of hazardous buried collapse dolines prior to full subsidence (Montané, 2001; Kruse *et al.*, 2006; Pueyo-Anchuela *et al.*, 2009; Gutiérrez *et al.*, 2011).

Mesoscale (c.1cm-10m)

Mesokarren features in particular have been the subject of numerous morphometric analyses, not only in order to characterise and classify these landforms through diagnostic measurement parameters, but as

a means to understand their formation and simulate how they develop. The selection of individual mesokarren forms for morphometric analysis often involves sampling strategies based along transects (Veress *et al.*, 2001) or square-grid maps (Tóth, 2009) situated amongst mesokarren assemblages or on karren outcrops. Morphometric analyses of mesokarren forms have traditionally been performed with a carpenter profile gauge in the case of small measurements (c.1-15cm), and survey tapes for larger features (c.0.15-10m). The precise protocol for mesokarren morphometry varies for different landforms. Some features, such as rillenkarrren, have been extensively studied and have an established and detailed morphometric protocol (Lundberg and Ginés, 2009), while other features, such as rainpits, have ill-defined morphological protocols.

Morphometric studies (Mottershead *et al.*, 2000) and modelling (Glew and Ford, 1980) of rillenkarrren typically encompass records of solution flute width, depth, slope angle and length, and are often recorded with a carpenter profile gauge whose pins conform to the shape and size of the rills (Lundberg and Ginés, 2009). Similar studies of kamenitze comprise solution basin depth, width, length and micro notch measurements (Cucchi, 2009; Tóth, 2009), while trittkarrren analysis generally involves measurements of slope, riser height and width, tread angle and length, and foreground width (Vincent, 1983; Veress, 2009c). Mäanderkarrren morphometric records include maximum width, horizontal width, horizontal middle and maximum depth, and large and small cross-sectional area (Hutchinson, 1996), which are similar to the parameters applied to river channels (Veress, 2009b).

Successful mesokarren morphometry involves the comparison of different parameters of the same form, allowing the researcher to deduce how the topographical and stratigraphical position of the karst terrain may influence the development of mesoscale forms. According to Tóth (2009), for any given karren form, characteristic parameters such as width, depth and length should be chosen and compared with measurements of the same karren form under different conditions of, for example, slope angle, exposure and precipitation. Adopting such an approach to research has led to the idea of

karren forms as ecological indicators, since the precise morphology and morphometry of karren features indicates a particular series of ecological and environmental conditions. For example, Ginés (1990) found a negative correlation between altitude (precipitation) and rillenkarren length in the Sierra Tramuntana, Mallorca, while Glew and Ford (1980) identified a positive correlation between slope angle and rillenkarren length, and that wider rillenkarren evolve on karst limestones than on evaporitic karstic materials. Similarly, data comparison by Zeller (1967) identified the sinuosity of mäanderkarren to be greater than that of fluvial meanders and glaciers.

Apart from the slow and laborious nature of manually measuring and deducing mesokarren morphometry, discrepancies between methods of measurement often create difficulties in the accurate morphometrical description of mesokarren landforms and in the global comparison of data between authors (Lundberg and Ginés, 2009). As with macroscale investigations, the widespread application of terrestrial laser and optical scanning systems to mesoscale morphometric analysis in the near future is likely to significantly increase the accuracy and speed of data capture. In a recent study, Mottershead *et al.* (2008) successfully employ a scanning total station and resultant digital elevation models to examine morphological change and determine solutional surface lowering on rocksalt slopes over an eight month period. As highlighted by the authors, the study of highly soluble karstic substrates, such as saltrock, has the potential to shed light upon dissolutional processes and morphological development on less soluble karstic materials such as gypsum and limestone within a much shorter experimental time period.

Terrestrial LIDAR may have further applications in karst modelling to document the evolution and development of mesokarren forms such as rillenkarren under controlled settings by, for example, scanning gypsum plaster blocks subject to rainfall at intervals. Potentially, the technology could also be employed to provide accurate, three dimensional morphometric data of natural karren forms which could then be faithfully recreated at 1:1 scale in gypsum plaster, exposed to runoff and monitored for further

morphological development. Crucially, laser scanning may also resolve issues of data comparability and consistency between authors, as raised by Lundberg and Ginés (2009), by stipulating a new protocol and allowing morphometric data to be directly comparable on a global scale.

Microscale (c.1mm-1cm)

Descending a scale boundary, investigations into karstic surface morphologies at the microscale commonly require the use of optical microscopical techniques, such as stereomicroscopy or thin-section microscopy, although some authors continue to opt for traditional mesoscale morphometric techniques, albeit at a smaller scale.

Millimetre to centimetre-sized rainpits are commonly measured with a carpenter profile gauge, although no protocol has ever been established for their measurement (Ginés and Lundberg, 2009). A measure of surface roughness at the microscale may also be achieved through the use of a carpenter gauge placed vertically along the surface of a rock and photographed. Digitisation of the images allows the height difference between consecutive points to be determined and a measure of surface roughness to be calculated (Crowther, 1996).

Dealing with millimetre-sized microrills, Gómez-Pujol and Fornós (2009) adopt an optical stereomicroscopy approach to morphometric analysis. They derive microrill width and geometry from digital image processing software, allowing them to calculate the lateral distance between microrill crests. McIlroy de la Rosa *et al.* (2012) employ thin-section optical microscopy to measure the width and depth of biotroughs, microscale biokarstic features formed by the coalescing and enlarging of nanoscale endolithic lichen-induced biopits.

Investigations into microscale surface morphology inevitably enter the field of weathering studies, as surface micromorphologies are a product of a karstic substrates interaction with the weathering environment. As such, our understanding of karst surface micromorphologies is derived not only from geomorphological studies but also from studies of karst substrates utilised as dimension stone within buildings and

monuments. Highly soluble saltrock and gypsum yield observable morphological change over short experimental periods. However, the slow rate of natural limestone denudation inevitably entails that human observation and quantification of limestone surface loss occurs at a small, millimetre to centimetre scale, over the course of an extremely short, in geological terms, data recording period. In order to estimate decadal, centennial and millennial rates of limestone surface loss, data are often extrapolated from recording periods usually lasting between one and six years (Cucchi *et al.*, 1995; Urushibara-Yoshino *et al.*, 1999), with some up to 30 years (Stephenson *et al.*, 2010).

A variety of techniques may be employed in order to quantify dissolutional surface loss from karst substrates. Long-term studies from a defined base point may be undertaken using equipment such as the micro-erosion meter, which is fixed to the rock with studs and measures surface loss accurately with a dial gauge (Coward, 1975; Cucchi *et al.*, 1995). A modified version, the traversing micro-erosion meter was developed by Trudgill *et al.* (1981) and allows multiple individual measurements per erosion meter bolt site (Stephenson and Finlayson, 2009; Stephenson *et al.*, 2010). Micro-erosion metres have traditionally been employed on slowly-dissolving karst substrates such as limestone. The rate of surface lowering for more rapidly dissolving karst substrates, such as gypsum or saltrock, may be assessed using plastic (Bruthans *et al.*, 2008) or brass (Mottershead *et al.*, 2007) erosion pins (c.3mm in diameter and c.50mm long) inserted into the rock until the upper end is flush with the surface. As the surface dissolves, the pins emerge allowing a rate of dissolution to be quantified. Alternatively, surface loss may be quantified by chemical studies of runoff from karst surfaces (Dunkerley, 1983; Vleugels, 1992; Fiol *et al.*, 1992; 1996), or by simulated laboratory experiments under accelerated weathering environments (Goudie, 1999; Thornbush and Viles, 2007). Another method, often applied to slowly-dissolving substrates such as limestone, involves the weighing of rock tablets before and after exposure in order to derive solutional losses (Gams, 1981; Urushibara-Yoshino *et al.*, 1999; Viles *et al.*, 2002). In nature, pedestals formed beneath

glacial erratics on limestone pavements have also been used to calculate solution rates since the last glacial maxima of the Pleistocene, when the pavements were scoured by ice and the erratics deposited (Jennings, 1987). Finally, optical and laser scanning (LIDAR) technology allows the precise quantification of surface loss and its spatial variation on rock surfaces. High-resolution laser scanners are capable of detecting sub-millimetre changes in surface morphology, making them highly relevant to investigations into microtopographical changes on karst surfaces over time. Optical and laser scanning approaches have a number of advantages over other techniques as they are non-intrusive, requiring no direct contact with the karstic surface under investigation, unlike the micro-erosion meter (Spate *et al.*, 1985). They may also be used *in situ* in the field, without the need to remove rock surfaces for laboratory analysis. At present, the use of laser and optical scanning in the assessment of microscale surface roughness and mineral loss over time from karst substrates has been largely restricted to the study of stone decay in a cultural setting (Birginie and Rivas, 2005; Gomez-Heras *et al.*, 2008; Meneely *et al.*, 2009). However, the techniques and methodologies employed on dimension stone are equally applicable to the study of stone surface micromorphologies and morphological change over time in the natural environment. Gómez-Pujol *et al.* (2006) adopt a similar laser scanning approach to Mottershead *et al.* (2008) in order to quantify surface roughness on coastal carbonate rocks, with a scanning resolution of 0.1mm (± 0.05 mm error). At its current level of development, the applicability of traditional optical and laser scanning to karst landform morphometric analysis and classification generally ends at the microscale. Descending a scale boundary to the nanoscale, where observations are often made in microns, surpasses the resolution and accuracy of current laser scanning devices.

Nanoscale (c.<1mm)

In recent years, weathering studies have adopted an increasingly reductionist approach, with research evolving towards morphological investigation at increasingly smaller scales. The smallest scale of observation for weathering morphologies is

the nanoscale. The classification of nanotopographical features on karstic surfaces is problematic in that features at this scale are often difficult to differentiate and are surrounded by a degree of equifinality or convergence, as similar nanomorphologies may develop as a result of different processes (Viles and Moses, 1998).

Inevitably, investigations at the nanoscale enter the realm of crystallography and petrography, requiring the use of microscopical and spectroscopical techniques. Thin-section light microscopy is still applicable at this scale and has been applied to both biotic and abiotic weathering studies (Del Monte and Sabbioni, 1987; Bolívar and Sánchez-Castillo, 1997; Garcia Vallès *et al.*, 2000; Bungartz *et al.*, 2004; Favero-Longo *et al.*, 2009; Mcllroy de la Rosa *et al.*, 2012). However, detailed investigations of surface features smaller than c.0.5mm are often pursued through the use of electron microscopy.

Scanning electron microscopy (SEM) has been successfully used to identify and attempt to classify carbonate stone weathering nanomorphologies on both experimental blocks sprayed with dilute acid, natural surfaces and building materials. Nanomorphologies such as stepped etching, etch pits, cleavage widening and v-in-v etching among others, have been observed on calcite crystals and are associated with crystal dissolution (Gillott, 1978; Moses *et al.*, 1995; Moses and Viles, 1996; Viles and Moses, 1998; Thornbush and Viles, 2007).

SEM has also been applied to the identification of biologically-induced nanomorphologies on calcite. Circular etch pits, etch tunnels and filament-shaped trenches are the most widely documented (Moses *et al.*, 1995; Moses and Viles, 1996; Viles and Moses, 1998), although morphologies produced by the excretion of organic acids are often difficult to distinguish from those produced by simple dissolution. SEM has also been successfully applied to the identification of algae cells (Fiol *et al.*, 1996; Smith *et al.*, 2000), fungal hyphae (Thornbush and Viles, 2006; Concha Lozano *et al.*, 2012) and salts, such as calcium oxalate (Wadsten and Moberg, 1985; Russ *et al.*, 1996; Giordani *et al.*, 2003; Edwards, 2007) and gypsum (Delalieux *et al.*, 2001; De

Graef *et al.*, 2005), within the rock matrix or at the interface between lithobionts and calcareous karst substrates. Variants of SEM successfully applied to the investigations of nanoscale interactions between lithobionts and karst substrates include transmission electron microscopy (TEM), favoured for the observation of biological structures (De los Ríos *et al.*, 2002; 2009), low temperature scanning electron microscopy (LTSEM) (Ascaso *et al.*, 2002; De los Ríos *et al.*, 2009), and scanning electron microscopy with backscattered electron imaging (SEM-BSE) (Sanders *et al.*, 1994; Ascaso *et al.*, 1998; 2002; Bolívar and Sánchez-Castillo, 1997; Bungartz *et al.*, 2004). Atomic force microscopy (AFM), a three-dimensional imaging and measuring approach, has also been employed to quantify limestone surface roughness at the micron scale (Fornós *et al.*, 2011).

The recent reconsideration of the role of biological agents in karst landform development (e.g. Fiol *et al.*, 1996; Moses and Viles, 1996; Mottershead *et al.*, 2000; Mcllroy de la Rosa *et al.*, 2012), has opened the way for geomicrobiological techniques in biokarstic studies. Increasingly, we are seeing contributions from, and interdisciplinary research between, geomorphologists, microbiologists, lichenologists and mycologists, with the innovation in methods and techniques that this brings.

Finally, in an attempt to characterise the geochemical and mineralogical characteristics of weathered surfaces, weathering studies of karst substrates frequently involve spectroscopical techniques such as Fourier Transform infrared and Raman spectroscopy (Edwards *et al.*, 1997; Holder *et al.*, 2000; Monte, 2003; Frost, 2004), x-ray diffraction (Goudie *et al.*, 1997; Ascaso *et al.*, 1998; Giordani *et al.*, 2003) and atomic absorption spectrophotometry (Goudie *et al.*, 1997) among others.

Conclusion

The classification of karst landforms is a complex combination of synonymic terminology and polygenetic karren features. Classifications by Bögli (1960) and Ford and Williams (1989; 2007) remain the most

comprehensive and widely embraced by the academic community.

Karst surface morphologies exist at a variety of scales. Inevitably then, a wide variety of geomorphological techniques and methods are employed in karst morphometric analysis. Some techniques are applicable only at a particular scale while others may traverse scale boundaries. Karst and karren studies are at a methodological crossroads, with future morphometric techniques likely to evolve from the collection of data via complex and laborious manual field measurement methodologies, to the use of ergonomic, fast, non-intrusive and high resolution airborne / terrestrial LIDAR systems, the production of accurate digital elevation models and associated GIS analysis. Such an integrated approach is applicable from the macro to the microscale and allows morphometric data to be globally comparable.

References

- Ascaso, C., Wierzchos, J., Castello, R. (1998) "Study of the biogenic weathering of calcareous litharenite stones caused by lichen and endolithic microorganisms". *International Biodeterioration and Biodegradation*. 42: 29-38.
- Ascaso, C., Wierzchos, J., Souza-Egipsy, V., De los Ríos, A., Delgado Rodrigues, J. (2002) "In situ evaluation of the biodeteriorating action of microorganisms and the effects of biocides on carbonate rock of the Jeronimos Monastery (Lisbon)". *International Biodeterioration and Biodegradation*. 49: 1-12.
- Birginie, J.M. and Rivas, T. (2005) "Use of a laser camera scanner to highlight the surface degradation of stone samples subjected to artificial weathering". *Building and Environment*. 40: 755-764.
- Bögli, A. (1960) "Kalklösung und Karrenbildung". *Zeitschrift für Geomorphologie, Supplement-band*. 2: 4-21.
- Bögli, A. (1976) Die wichtigsten Karrenformen der Kalkalpen. In: *Karst Processes and Relevant Landforms*. Department of Geography. Ljubljana. p 141-149.
- Bögli, A. (1980) *Karst Hydrology and Physical Speleology*. Berlin: Springer-Verlag.
- Bolívar, F.C. and Sánchez-Castillo, P.M. (1997) "Biomineralization processes in the fountains of the Alhambra, Granada, Spain". *International Biodeterioration and Biodegradation*. 40: 205-215.
- Bonacci, O. (2004) Poljes. In: Gunn, J. (Ed.), *Encyclopedia of Caves and Karst Science*. New York: Fitzroy Dearborn. p. 599-600.
- Brook, G.A. and Ford, D.C. (1978) "The origin of labyrinth and tower karst and the climatic conditions necessary for their development". *Nature*. 275: 493-496.
- Bruthans, J., Asadi, N., Filippi, M., Vilhelm, Z., Zare, M. (2008) "A study of erosion rates on salt diaper surfaces in the Zagros Mountains, SE Iran". *Environmental Geology*. 53 (5): 1079-1089.
- Bungartz, F., Garvie, L.A.J., Nash III, T.H. (2004) "Anatomy of the endolithic Sonoran Desert lichen *Verrucaria rubrocincta* Breuss: implications for biodeterioration and biomineralisation". *The Lichenologist*. 36(1): 55-73.
- Calaforra, J.M. and Pulido-Bosch, A. (1999) "Genesis and evolution of gypsum tumuli". *Earth Surface Processes and Landforms*. 24: 919-930.
- Čalić, J. (2011) "Karstic uvala revisited: Toward a redefinition of the term". *Geomorphology*. 134: 32-42.
- Canevese, E.P., Forti, P., Naseddu, A., Ottelli, L., Tedeschi, R. (2011) "Laser scanning technology for the Hypogean Survey: The case of Santa Barbara karst system (Sardinia, Italy)". *Acta Carsologica*. 40 (1): 65-77.
- Concha-Lozano, N., Gaudon, P., Pages, J., de Billerbeck, G., Lafon, D., Eterradosi, O. (2012) "Protective effect of endolithic fungal hyphae on oolitic limestone buildings". *Journal of Cultural Heritage*. 13 (2): 120-127.
- Coward, J.M.H. (1975) *Paleohydrology and streamflow simulation of three karst basins in southeastern West Virginia*. McMaster University. Canada. PhD Thesis.

- Crowther, J. (1996) Roughness (mm-scale) of limestone surfaces: examples from coastal and mountain karren features in Mallorca. In: Fornós, J.J. and Ginés, A. (Eds.) *Karren Landforms*. Universitat de les Illes Balears, Palma de Mallorca. 149-159.
- Cucchi, F., Forti, F., Marinetti, E. (1995) Surface degradation of carbonate rocks in the Karst of Trieste (Classical Karst, Italy). In: Fornós, J.J. and Ginés, A. (Eds.) *Karren Landforms*. Universitat de les Illes Balears, Palma de Mallorca. 41-51.
- Cucchi, F. (2009) Kamenitzas. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- De Graef, B., Cnudde, V., Dick, J., De Belie, N., Jacobs, P., Verstraete, W. (2005) "A sensitivity study for the visualisation of bacterial weathering of concrete and stone with computerised X-ray microtomography". *Science of the Total Environment*. 341: 173-183.
- De los Ríos, A., Ascaso, C., Grube, M. (2002) "An ultrastructural, anatomical and molecular study of the lichenicolous lichen *Rimularia insularis*". *Mycological Research*. 106: 946-953.
- De los Ríos, A., Cámara, B., García del Cura, M.A., Rico, V.J., Galván, V., Ascaso, C. (2009) "Deteriorating effects of lichen and microbial colonization of carbonate building rocks in the Romanesque churches of Segovia (Spain)". *Science of the Total Environment*. 407: 1123-1134.
- Delalieux, F., Cardell, C., Todorov, V., Dekov, V., Van Grieken, R. (2001) "Environmental conditions controlling chemical weathering of the Madara Horseman monument, NE Bulgaria". *Journal of Cultural Heritage*. 2: 43-54.
- Del Monte, M. and Sabbioni, C. (1987) "A study of the patina called 'scialbatura' on imperial Roman marbles". *Studies in Conservation*. 32 (3): 114-121.
- Dunkerley, D.L. (1983) "Lithology and microtopography in the Chillagoe karst, Queensland, Australia". *Zeitschrift für Geomorphologie*. 27 (2): 191-204.
- Edwards, H.G.M. (2007) "A novel extremeophile strategy studied by Raman spectroscopy". *Spectrochimica Acta Part A*. 68: 1126-1132.
- Edwards, H.G.M., Russell, N.C., Seaward, M.R.D. 1997. "Calcium oxalate in lichen biodeterioration studied using FT-Raman spectroscopy". *Spectrochimica Acta Part A*. 53: 99-105.
- Eren, M. and Hatipoglu-Bagci, Z. (2010) "Karst surface features of the hard laminated crust (caliche hardpan) in the Mersin area, southern Turkey". *Acta Carsologica*. 39 (1): 93-102.
- Favero-Longo, S.E., Borghi, A., Tretiach, M., Piervittori, R. (2009) "In vitro receptivity of carbonate rocks to endolithic lichen-forming aposymbionts". *Mycological Research*. 113: 1216-1227.
- Fiol, Ll., Fornós, J.J., Ginés, A. (1992) "El rillenkarrén: un tipus particular de biokarrén? Primeres dades". *Endins*. 17-18. 43-49.
- Fiol, Ll., Fornós, J.J., Ginés, A. (1996) "Effects of biokarstic processes on the development of solutional rillenkarrén in limestone rocks". *Earth Surface Processes and Landforms*. 21: 447-452.
- Fleurant, C., Tucker, G.E., Viles, H.A. (2008) "A model of cockpit karst landscape, Jamaica" *Géomorphologie: relief, processus, environnement*. 1: 3-14.
- Ford, D.C. and Lundberg, J.A. (1987) "A review of dissolutional rills in limestone and other soluble rocks". *Catena Supplement*. 8: 119-140.
- Ford, D.C. and Williams, P. (1989) *Karst Hydrology and Geomorphology*. London: Unwin Hyman.
- Ford, D.C. and Williams, P. (2007) *Karst Hydrogeology and Geomorphology*. Chichester: Wiley.
- Fornós, J.J., Gómez-Pujol, L., Cifre, J., Hierro, F. (2011) "First steps in limestone weathering and erosion: an atomic force

- microscopy (AFM) and scanning electron microscopy (SEM) approach". *Acta Carsologica*. 40 (2): 275-282.
- Frost, R.L. 2004. "Raman spectroscopy of natural oxalates". *Analytica Chimica Acta*. 517: 207-214.
- Gams, I. (1973) The Terminology of the Types of Polje. Slovenska Kraska Terminologija, Zveza Geografskih Institucij Jugoslavije, Ljubljana. pp.60-70.
- Gams, I. (1978) "The polje: the problem of its definition". *Zeitschrift für Geomorphologie*. 22: 170-181.
- Gams, I. (1981) "Comparative research of limestone solution by means of standard tablets". *Proceedings of the 8th International Congress of Speleology, Bowling Green, Kentucky*. Vol 1: 273-275.
- Garcia-Vallès, M., Urzì, C., De Leo, F., Salamone, P., Vendrell-Saz, M. (2000) "Biological weathering and mineral deposits of the Belevi marble quarry (Ephesus, Turkey)". *International Biodeterioration and Biodegradation*. 46: 221-227.
- Gillot, J.E. (1978) "Effect of deicing agents and sulphate solutions on concrete aggregate". *Quarterly Journal of Engineering Geology*. 11: 177-192.
- Ginés, A. (1990) "Utilización de las morfologías de lapiaz como geoindicadores ecológicos en la Sierra de Tramuntana (Mallorca)". *Endins*. 16: 27-39.
- Ginés, A. (2009) Karrenfield landscapes and karren landforms. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing. 13-24.
- Ginés, A. and Lundberg, J. (2009) Rainpits: an outline of their characteristics and genesis. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing. 169-183.
- Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Giordani, P., Modenesi, P., Tretiach, M. 2003. "Determinant factors for the formation of the calcium oxalate minerals, weddellite and whewellite, on the surface of foliose lichens". *Lichenologist*. 35 (3): 255-270.
- Glew, J.R. and Ford, D.C. (1980) "A simulation study of the development of rillenkarren". *Earth Surface Processes and Landforms*. 5: 25-36.
- Goldie, H.S. (2009) Kluffkarren or grikes as fundamental karstic phenomena. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Gomez-Heras, M., Smith, B.J., Viles, H.A., Meneely, J., McCabe, S. (2008) "HD laser scanning for the evaluation of salt decay laboratory simulations of building Stone". SWBSS. Copenhagen 2008.
- Gómez-Pujol, L. and Fornos, J.L. (2009) Microrills. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Gómez-Pujol, L., Fornós, J.J., Swantesson, O.H. (2006) "Rock surface millimetre-scale roughness and weathering of supratidal Mallorcan carbonate coasts (Balearic Islands)" *Earth Surface Processes and Landforms*. 31: 1792-1801.
- Goudie, A.S. (1999) "A comparison of the relative resistance of limestones to frost and salt weathering". *Permafrost and Periglacial Processes*. 10: 309-316.
- Goudie, A.S., Viles, H.A., Parker, A.G. (1997) "Monitoring of rapid salt weathering in the central Namib Desert using limestone blocks". *Journal of Arid Environments*. 37: 581-598.
- Grimes, K.G. (2007) "Microkarren in Australia – a request for information". *Helictite*. 40 (1): 21-23.
- Gutiérrez, F., Galve, J.P., Lucha, P., Castañeda, C., Bonachea, J., Guerrero, J. (2011) "Integrating geomorphological mapping, trenching, InSAR and GPR for the identification and characterization of sinkholes: A review and application in the

mantled evaporite karst of the Ebro Valley (NE Spain)". *Geomorphology*. 134: 144-156.

Holder, J.M., Wynn-Williams, D.D., Rull Perez, F., Edwards, H.G.M. 2000. "Raman spectroscopy of pigments and oxalates *in situ* within epilithic lichens: *Acarospora* from the Antarctic and Mediterranean". *New Phytologist*. 145: 271-280.

Hutchinson, D.W. (1996) Rinnels, rinnenkarren and mäanderkarren: form, classification and relationships. In: Fornós, J.J. and Ginés, A. (Eds.) *Karren Landforms*. Universitat de les Illes Balears, Palma de Mallorca. 209-223.

Jennings, J.N. (1987) *Karst Geomorphology*. Oxford: Blackwell.

Jennings, J.N. and Sweeting, M.M. (1963) "The limestone ranges of the Fitzroy Basin, Western Australia". *Bonner Geographische Abhandlungen*. 32: 60pp.

Knez, M., Otoničar, B., Slabe, T. (2003) "Subcutaneous Stone forest (Trebnje, central Slovenia)". *Acta Carsologica*. 32 (1): 29-38.

Krabill, W.B., Thomas, R.H., Martin, C.F., Swift, R.N., Frederick, E.B. (1995) "Accuracy of airborne laser altimetry over the Greenland ice sheet" *International Journal of Remote Sensing*. 16: 1211-1222.

Kruse, S., Grasmueck, M., Weiss, M., Viggiano, D. (2006) "Sinkhole structure imaging in covered karst terrain". *Geophysical Research Letters*. 33. L16405, doi:10.1029/2006GL026975.

Kunaver, J. (2009) Corrosion terraces, megaausgleichsfläche or a specific landform of bare glaciokarst. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.

Laudermilk, J.D. and Woodford, A.O. (1932) "Concerning Rillensteine". *American Journal of Science*. 23 (134): 135-154.

Lundberg, J. and Ginés, A. (2009) Rillenkarren. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.

Lyew-Ayee, P., Viles, H.A., Tucker, G.E. (2007) "The use of GIS-based digital morphometric techniques in the study of cockpit karst". *Earth Surface Processes and Landforms*. 32: 165-179.

McIlroy de la Rosa, J.P., Warke, P.A., Smith, B.J. (2012) "Microscale biopitting by the endolithic lichen *Verrucaria baldensis* and its proposed role in mesoscale solution basin development on limestone". *Earth Surface Processes and Landforms*. 37: 374-384.

Meneely, J.D., Smith, B.J., Curran, J., Ruffell, A. (2009) "Developing a 'non-destructive scientific toolkit' to monitor monuments and sites". ICOMOS Scientific Symposium. Malta 2009.

Montané, J.M. (2001) *Geophysical analysis of a central Florida karst terrain using light detection and ranging (LIDAR) and ground penetrating radar (GPR)*. FIU Electronic Thesis and Dissertations. Florida International University.

Monte, M. 2003. "Oxalate film formation on marble specimens caused by fungus". *Journal of Cultural Heritage*. 4: 255-258.

Moses, C.A., Spate, A.P., Smith, D.I., Greenaway, M.A. (1995) "Limestone weathering in eastern Australia. Part 2: Surface micromorphology study". *Earth Surface Processes and Landforms*. 20: 501-514.

Moses, C.A. and Viles, H.A. (1996) Nanoscale morphologies and their role in the development of karren. In: Fornós, J.J. and Ginés, A. (Eds.) *Karren Landforms*. Universitat de les Illes Balears, Palma de Mallorca. 225-238.

Mottershead, D.N., Moses, C.A., Lucas, G.R. (2000) "Lithological control of solution flute form: A comparative study". *Zeitschrift für Geomorphologie*. 44 (4): 491-512.

Mottershead, D.N., Wright, J.S., Inkpen, R.J., Duane, W. (2007) "Bedrock slope evolution in saltrock terrain" *Zeitschrift für Geomorphologie Supplementary Issues*. 51 (1): 81-102.

- Mottershead, D.N., Duane, W.J., Inkpen, R.J., Wright, J.S. (2008) "An investigation of the geometric controls on the morphological evolution of small-scale salt terrains, Cardona, Spain". *Environmental Geology*. 53 (5): 1091-1098.
- Perna, G. and Sauro, U. (1978) *Atlante delle microforme di dissoluzione carsica superficiale del Trentino e del Veneto*, Museo Tridentino, Trento. pp. 176.
- Pueyo-Anchuela, Ó., Pocoví Juan, A., Soriano, M.A., Casas-Sainz, A.M. (2009) "Characterization of karst hazards from the perspective of the doline triangle using GPR- Examples from the Central Ebro Basin (Spain)". *Engineering Geology*. 108: 225-236
- Russ, J., Palma, R.L., Loyd, D.H., Boutton, T.W., Coy, M.A. 1996. "Origin of the whewellite-rich rock crust in the lower Pecos region of southwest Texas and its significance to palaeoclimate reconstructions". *Quaternary Research*. 46: 27-36.
- Rüther, H., Chazan, M., Schroeder, R., Neeser, R., Held, C., Walker, S.J., Matmon, A., Horwitz, L.K. (2009) "Laser scanning for conservation and research of African cultural heritage sites: the case study of Wonderwerk Cave, South Africa". *Journal of Archaeological Science*. 36 (9): 1847-1856.
- Sanders, W.B., Ascaso, C., Wierzchos, J. (1994) "Physical interactions of two rhizomorph-forming lichens with their rock substrate". *Botanica Acta*. 107: 432-439.
- Sauro, U. (2003) "Dolines and sinkholes: Aspects of evolution and problems of classification". *Acta Carsologica*. 32 (2): 41-52.
- Siart, C., Forbriger, M., Ghilardi, M. (2011) "Fusing surface and subsurface geodata: a case study on Cretan karst landforms". *Geophysical Research Abstracts*. Vol. 13.
- Smith, B.J., Warke, P.A. and Moses, C.A. (2000) "Limestone weathering in contemporary arid environments: A case study from southern Tunisia". *Earth Surface Processes and Landforms*. 25: 1343-1354.
- Spatte, A.P., Jennings, J.N., Smith, D.I., Greenaway, M.A. (1985) "The micro-erosion meter: Use and limitations". *Earth Surface Processes and Landforms*. 10 (5): 427-440.
- Stephenson, W.J., Kirk, R.M., Hemmingsen, S.A., Hemmingsen, M.A. (2010) "Decadal scale micro erosion rates on shore platforms". *Geomorphology*. 114: 22-29.
- Stephenson, W.J. and Finlayson, B.L. (2009) "Measuring erosion with the micro-erosion meter- Contributions to understanding landform evolution". *Earth-Science Reviews*. 95: 53-62.
- Sweeting, M.M. (1995) *Karst in China: its Geomorphology and Environment*. Berlin: Springer-Verlag.
- Thornbush, M. and Viles, H. (2006) "Changing patterns of soiling and microbial growth on building stone in Oxford, England after implementation of a major traffic scheme". *Science of the Total Environment*. 367: 203-211.
- Thornbush, M.J. and Viles, H.A. (2007) "Simulation of the dissolution of weathered versus unweathered limestone in carbonic acid solutions of varying strength". *Earth Surface Processes and Landforms*. 32: 841-852.
- Tóth, G. (2009) Some methodologies on karren research. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Trudgill, S., High, C.J., Hanna, F.K. (1981) "Improvements to the micro-erosion meter". *British Geomorphological Research Group Technical Bulletin*. 29: 3-17.
- Urushibara-Yoshino, K. Miotke, F.D., Kashima, N. (1999) "Solution rate of limestone in Japan". *Physics and Chemistry of the Earth, Series A*. 24 (10): 899-903.
- Veress, M. (2000) "The morphogenetics of karren meander and its main types". *Karsztfejlődés*. 4: 41-76.
- Veress, M. (2009a) Rinnenkarren. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.)

- (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Veress, M. (2009b) Meanderkarren. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Veress, M. (2009c) Trittkarren. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.
- Veress, M., Tóth, G., Zentai, Z., Kovács, G. (2001) "Study of a new method for characterising karren surfaces based on alpine researches". *Revue de Géographie Alpine*. 89: 49-62.
- Viles, H.A. and Moses, C.A. (1998) "Experimental production of weathering nanomorphologies on carbonate stone". *Quarterly Journal of Engineering Geology*. 31: 347-357.
- Viles, H.A., Taylor, M.P., Yates, T.J.S., Massey, S.W. (2002) "Soiling and decay of N.M.E.P. limestone tablets". *The Science of the Total Environment*. 292: 215-229.
- Vincent, P.J. (1983) "The morphology and morphometry of some arctic Trittkarren". *Zeitschrift für Geomorphologie*. 27 (2): 205-222.
- Vleugels, G. (1992) Weathering of bare and treated limestone under field-exposure conditions in Belgium: study of the runoff water from micro-catchment units. Universiteit Antwerpen, Departement Scheikunde.
- Wadsten, T. and Moberg, R. 1985. Calcium oxalate hydrates on the surface of lichens". *Lichenologist*. 17 (3): 239-245.
- White, B.W. (1988) *Geomorphology and Hydrology of Karst Terrains*. New York: Oxford University Press.
- Williams, P.W. (1966) "Limestone Pavements with Special Reference to Western Ireland". *Transactions of the Institute of British Geographers*. 40: 155-172.
- Xiong, K. (1992) "Morphometry and evolution of fenglin karst in the Shuicheng area, western Guizhou, China". *Zeitschrift für Geomorphologie*. 36 (2): 278-248.
- Zeller, J. (1967) Meandering channels in Switzerland. International Union of Geodesy and Geophysics, International Association of Scientific Hydrology (IUGG/IASH), Symposium on River Morphology, Bern, 174-186.
- Zseni, A., Goldie, H., Bárány-Kevei, I. (2003) "Limestone pavements in Great Britain and the role of soil cover in their evolution". *Acta Carsologica*. 32 (5): 57-67.
- Zseni, A. (2009) Subsoil shaping. In: Ginés, A., Knez, M., Slabe, T., Dreybrodt, W. (eds.) (2009) *Karst Rock Features: Karren Sculpturing*. Ljubljana: ZRC Publishing.

3.7.1. A review of *in situ* measurement techniques for investigating suspended sediment dynamics in lakes

Daniel N. Schillereff¹

¹ School of Environmental Sciences, University of Liverpool (dns@liv.ac.uk)



ABSTRACT: Lakes can function as important sinks of catchment-derived material, offering a valuable data resource for geomorphologists investigating soil erosion, sediment yields or particulate contamination, for example. The complex internal dynamics of lakes necessitates careful consideration of data collection techniques, however. This contribution outlines the four classes of device used for *in situ* measurements of suspended sediment concentration (SSC) and potentially tracking of river-borne sediment plume dispersal (pressure differential, bulk optic, acoustic backscatter and laser optic) and highlights potential pitfalls and limitations of their use in lacustrine settings. Considerations when designing and installing sediment traps are provided in detail, with a particular focus on appropriate trap dimensions, mooring setup and deployment protocol. The relative merits of employing SSC sensors versus sediment traps are considered and, in general, method selection will be guided by the research question(s). The contribution to process understanding that can be made by comparing *in situ* measurements with lake sediment cores is also highlighted.

KEYWORDS: Limnology, suspended sediment, sediment traps, thermal regime, river plume

Introduction

Lakes represent a critical component of watersheds, acting as effective sinks for particles eroded from catchment slopes and floodplains and subsequently transported through the fluvial system (Mackereth, 1966). As a result, exploring the internal dynamics of lakes, especially the controls on fluxes and fates of catchment-derived particles, may be of interest to geomorphologists investigating catchment erosion rates (Dearing, 1991), sediment yields (Foster *et al.*, 2011), responses to soil-vegetation destabilisation (Foster *et al.*, 2003) or potential hydrometeorological drivers of extreme sediment delivery events (Cockburn and Lamoureux, 2008) or those evaluating risks to aquatic ecosystems from excessive fine particulates (Wood and Armitage, 1997) or trace metal contamination (Douglas and Rippey, 2000).

However, the internal complexity of lakes (see Lerman *et al.*, 1995) means adequately

capturing spatial and temporal variability within direct measurements of water and particulates is challenging. This chapter aims to present an overview of *in situ* measurement techniques for some principle physical limnological processes that influence suspended sediment dynamics, including the dispersion of riverine plumes and thermal stratification regimes, as well as reviewing procedures for acquiring high-quality data from sediment traps. The rationale for selecting certain sensors is subsequently considered and, lastly, the value of integrating present-day measurements with sediment cores is emphasised.

Suspended sediment monitoring

Measurement techniques

Traditional methods for monitoring suspended sediment concentrations (SSC) required bottle sampling at discrete intervals; these have been widely superseded by devices capable of near-real time *in situ*

measurements (Felix *et al.*, 2013). Instruments developed in fluvial (Gray and Gartner, 2009) or marine (Agrawal and Pottsmith, 2000) settings are also deployable in lakes and tend to be classified based on their underpinning physical principle: pressure differential, bulk optic, acoustic backscatter or laser optic (Figure 1; Wren *et al.*, 2000; Gray and Gartner, 2009; Hubbart *et al.*, 2014). Measuring SSC from a fluvial perspective is covered by Perks (2014) so only brief technical descriptions will be provided here; their respective suitability and methodological limitations when deployed in lakes are emphasised.

- i) The pressure differential technique infers SSC from estimations of water density that are based on paired, simultaneous measurements of water pressure from transducers fixed at two known depths and corrected for water temperature (Gray and Gartner, 2009; Lewis and Rasmussen, 1999).
- ii) Bulk optic turbidimeters employ either transmissometry (a measure of the visible light fraction reaching a detector from a directly-aligned source) or nephelometry (the optical backscatter of incident light at $\geq 90^\circ$ within a chamber). In both cases, the amount of light reaching the detector will reflect the clarity or turbidity of the water. These values are calibrated against

empirical measurements that should be performed on a site-specific basis (Wren *et al.*, 2000; Gray and Gartner, 2009).

- iii) Acoustic backscatter techniques provide non-intrusive insight into the movement of suspended material within a lake across considerable vertical distances, typically using a Doppler current profiler (Guerrero *et al.*, 2011, 2012). Acoustic backscatter and Doppler shift mathematical principles are described in Kostaschuk *et al.* (2005) and Sassi *et al.* (2012); essentially, high frequency sound pulses emitted sequentially from instrument transducers rebound at strengths that reflect the concentration of suspended particles in the water column (Wren *et al.*, 2000).
- iv) Laser optical devices (often labelled 'laser *in situ* scattering and transmissiometry' (LISST); Agrawal and Pottsmith, 2000; Agrawal *et al.*, 2008) are underpinned by laser diffraction theory, whereby the angle of scatter (diffraction) of a laser beam upon hitting a spherical particle is a function of its diameter (Loizeau *et al.*, 1994). Suspended sediment concentrations and the spectrum of particle sizes are calculated from the scattering pattern detected by a ring sensor as a laser beam is emitted into a mass of particles held in suspension (Mikkelsen and Pejrup, 2001).

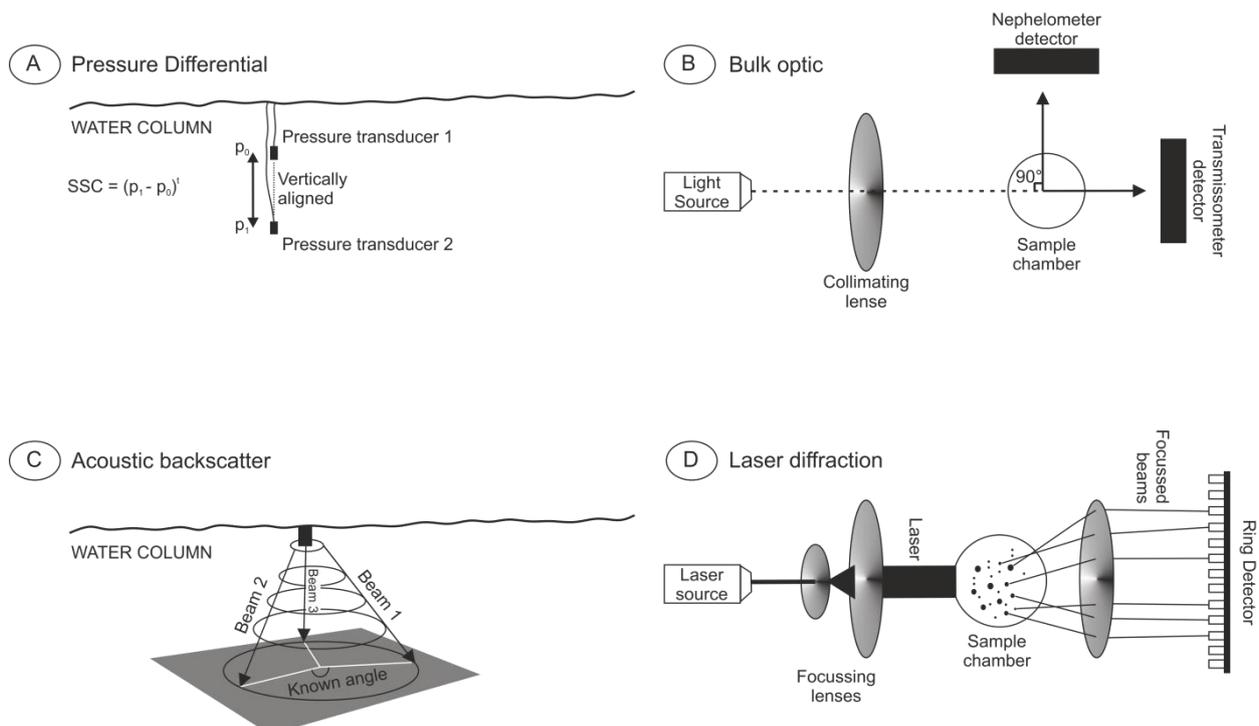


Figure 1. Schematic diagrams of the a) pressure differential; b) bulk optic; c) acoustic backscatter (re-drawn from Kostachuk *et al.* 2005) and d) laser diffraction techniques.

Suitability and limitations in lakes

Attempting to account for internal movements of suspended material, variable SSC as well as calm or turbulent subsurface flows may prevent certain SSC sensors from operating at maximum capability.

Concurrent temperature measurements at both depths are critical when employing the pressure differential technique (Gray and Gartner, 2009), rendering it potentially unsuitable for thermally stratified lakes. Likewise, tracking the dispersion of a sediment-rich plume as an interflow confined at the thermocline (see Figure 2) will require the precise suspension of both transducers within the plume itself.

The performance of optic instruments can be affected at certain SSC: transmissometers are more sensitive at low SSC, nephelometric sensors in turbid waters (Gray and Gartner, 2009), but both are hampered by exceptionally high particle concentrations (Wren *et al.*, 2000), potentially limiting their effectiveness when measuring infrequent, high-magnitude sediment delivery events.

Optic and acoustic techniques struggle to reconcile accurate SSC readings in the presence of variable particle sizes (Agrawal and Pottsmith, 2000), especially across the silt domain (Ludwig and Hanes, 1990) commonly deposited in lakes. Laser diffraction instruments (LISST-type) are capable of estimating concentrations and size distribution of sediments loads composed of widely distributed particle sizes (Agrawal and

Pottsmith, 2000), although a high proportion of particles finer than instrument operating limits (listed in Table 1) may skew SSC values (Andrews *et al.*, 2011).

Catchment geology or the lake-watershed configuration can influence LISST readings. Sediment loads dominated by freshly eroded, angular and/or flaky particles may cause LISSTs to overestimate SSC by 1.5 and 8 fold, respectively (Felix *et al.*, 2013). Similarly in granitic regions where discoid mica particles are a substantial contributor to sediment load, the assumption that particles are spherical when using a LISST-type device may not hold (Agrawal *et al.*, 2008). Highly turbid water (optical transmission $\tau < 0.3$; Felix *et al.*, 2013) may also preclude the use of LISST devices due to excessive lens obscuration (Cockburn and Lamoureux, 2008) or the possibility of multiple scattering (Agrawal and Pottsmith, 2000).

Lamoureux (2005) provides schematics for a LISST-type device deployed on a permanent mooring that contains sensors that detect light emitted from a column of LEDs on the opposite side of the unit. Sediment entering the central tube of the device will systematically impede the emitted light from reaching the adjacent sensors (Lamoureux, 2005). When used in conjunction with a data logger, the sensor can be deployed for multiple months with data acquired at minutely intervals.

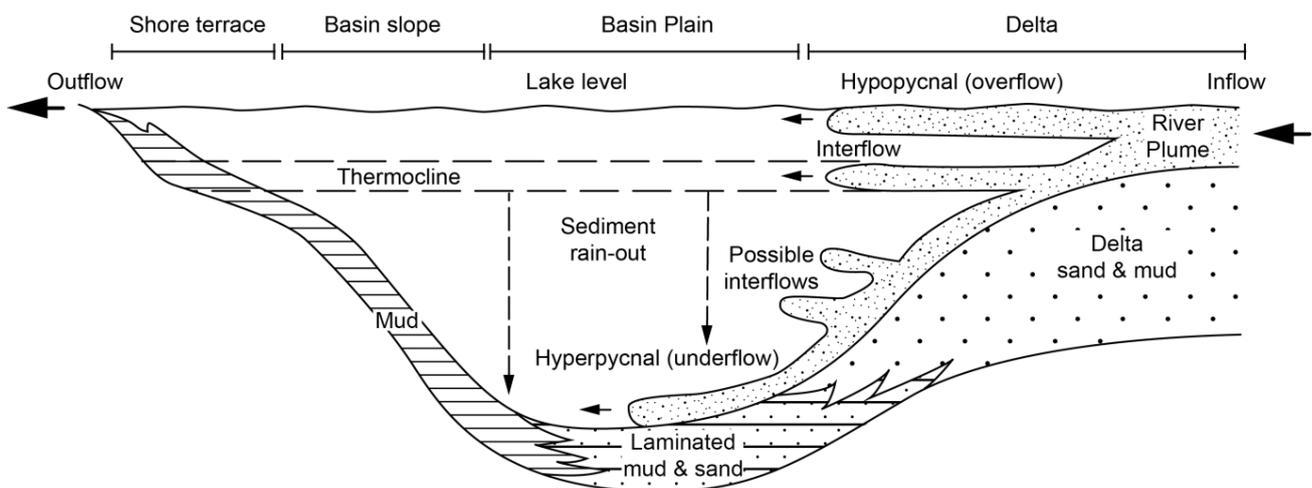


Figure 2: Processes of clastic sediment dispersal within a lake basin. Lake dimensions and sediment thicknesses are not to scale. From Schillereff *et al.* (2014).

Table 1. Technical details for different suspended sediment concentration sensors discussed in the text.

Operating principle (device name)	Minimum SSC (mg L ⁻¹)	Maximum SSC (mg L ⁻¹)	Particle size (µm)	References
Pressure differential	10000-20000	-----	-----	Lewis and Rasmussen, 1999 Gray and Gartner, 2009
Nephelometry (Optical Backscatterance Sensor; OBS)	-----	2000 ¹ 10000 ²	Inaccurate SSC values if <100, ideally 200-400	Ludwig and Hanes, 1990 Gray and Gartner, 2009
Transmissometry	-----	50	-----	Gray and Gartner, 2009
Laser diffraction (Sequoia LISST-100X Type C)	1	800 25 000 ³	2.5 – 500 ⁴ 1.9 – 381 ⁵	Gray and Gartner, 2009 Felix <i>et al.</i> , 2013
Acoustic backscatter (aDCP)	100	10000		Gray and Gartner, 2009

¹if sediment load is predominantly clay and silt

²if sediment load is predominantly sand-sized grains

³if optical pathway is narrowed

⁴if predominantly spherical grains are being measured

⁵if random-shaped grains are being measured

Information on both particulate concentration and size distribution can be gathered from acoustic Doppler current profiler (aDCP) measurements, although variable particle sizes pose problems when estimating total SSC from the backscatter profile (Reichel, 1998). Dual or multi-frequency deployments help address this issue (1500 and 500 kHz are common; Kostaschuk *et al.*, 2005). Optimum sensitivity lies in the coarse sand range that is rarely delivered to lakes, whereas fine-grained, cohesive material that dominates lacustrine sediment loads can hamper acoustic backscatter measurements, and higher frequency beams are required to detect silt-sized material (Kostaschuk *et al.*, 2005).

DCP surveys can be a powerful tool for investigating the internal behaviour and short-term pulses of turbulent flows, such as sediment-laden underflow plumes (Figure 2; Best *et al.*, 2005; Kostaschuk *et al.*, 2005). Surveying from both a moored position and along a delta-proximal to distal transect can

determine the flow regime of individual flood-generated underflows (Best *et al.*, 2005).

Laboratory testing should be used to develop calibration curves for calculating more accurate SSC values (Gray and Gartner, 2009; Felix *et al.*, 2013). Alternatively, employing multiple techniques in parallel may overcome issues posed by variable sediment load characteristics, but will introduce logistical and financial burdens. A thorough understanding of the lake-catchment system and clear research objectives will also help guide SSC instrument selection.

Sediment traps

Sediment traps have been used extensively in limnological research for a wide range of purposes, and detailed reviews of sediment trap theory and applications have been conducted by Bloesch and Burns (1980), Blomqvist and Håkanson (1981), Håkanson *et al.* (1989) and Bloesch (1996).

Sediment trap design

Sediment traps are simple instruments constructed from cylindrical, funnel or bottle-shaped tubes with a removable vessel attached at the base for capturing sediment (Figure 3). Traps with upward-facing openings installed at water depths of choice intercept the rainout of particles in suspension (Eadie, 1997) that settle vertically at rates determined by their diameters (Stokes, 1851). The often-stronger horizontal forces exerted within lakes (up to ten times the vertical component; Håkanson and Jansson, 1983) diminish as the particle enters the sediment trap opening, enabling it to sink vertically into the capturing vessel.

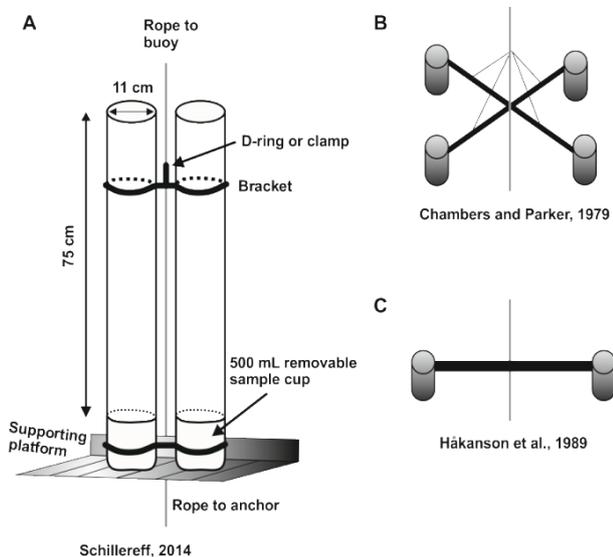


Figure 3. Illustrations of the sediment trap arrays used by a) Schillereff, 2015; b) Chambers and Parker, 1979 and c) Håkanson et al., 1989.

Trap shape determines whether its capture potential is representative of the actual vertical sediment flux within the lake and may vary under calm or turbulent flow conditions (See Figure 3.17, Håkanson and Jansson, 1983). Laboratory and field investigations conclude that cylinders are the preferable shape to avoid over or under-trapping of suspended sediment flux (95-100% of real sedimentation rate; Bloesch and Burns, 1980; Gardner, 1980). This applies in both calm and turbulent flow conditions for particles most likely to be trapped in lake: particles with diameters $<500 \mu\text{m}$ and low Reynolds number ($Re < 0.5$) or organic-rich particles (low density).

The cylinder height to diameter ratio (the aspect ratio) is also critical in trap design (Bloesch and Burns, 1980). Openings with a diameter $<40 \text{ mm}$ should be avoided to minimise under-sampling and aspect ratios greater than 15:1 may induce the development of anoxic conditions within the vessel (Håkanson and Jansson, 1983). The general recommendation to ensure particles are held at the base of the vessel is a minimum aspect ratio of 5:1 in calm conditions and 10:1 where more turbulent flow may occur (Hargrave and Burns, 1979; Blomqvist and Kofoed, 1981; Bloesch and Burns, 1980; Bloesch, 1996). The dimensions of some published trap designs are summarised in Table 2.

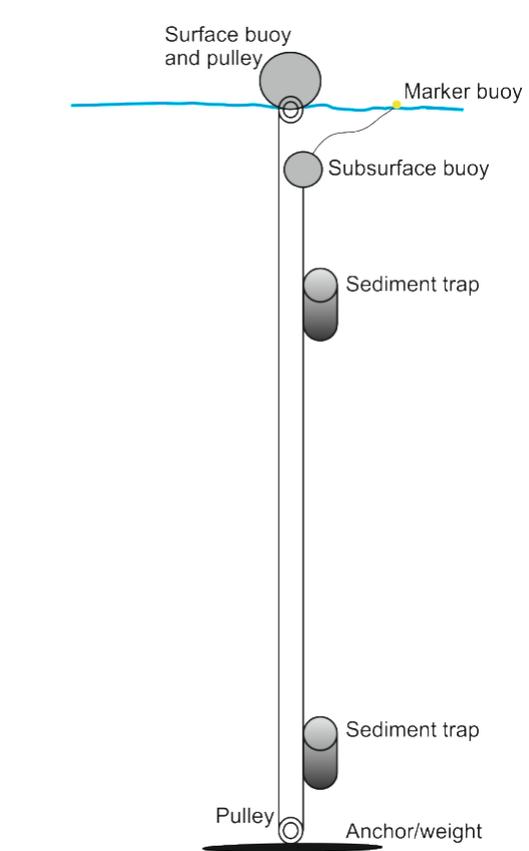


Figure 4. Idealised mooring setup for a sediment trapping programme. Modified from Ohlendorf and Sturm (2001)

The cylindrical body is most commonly constructed from transparent or non-transparent PVC or plexi-glass tubes. Field tests indicate that attaching collars or mesh at the opening should be avoided and lids that close during recovery are not required (Bloesch and Burns, 1980). Ohlendorf and Sturm (2001) installed a stopper at the base of the cylinder to release supernatant water

without disturbing the accumulated sediments but this is not required if the basal, removable vessel is sufficiently deep to prevent re-suspension when detaching from the tube during collection (Schillereff, 2015). Attach traps to a mooring cable or pre-stretched plastic rope held taught between an anchor at the lake bed and a surface buoy (Figure 4). Installing the upper buoy below the water surface (Håkanson *et al.* (1989) recommend 0.5 m) ensures the cabling and traps are kept

vertical during turbulent wave action or lake level fall (Douglas and Rippey, 2000). Ohlendorf and Sturm (2001) recommend the use of a pulley system between anchor and buoy to minimise re-suspension and a D-ring clasp (Schillereff, 2015) or clamp (Ohlendorf and Sturm, 2001) helps avoid trap tilt (Gardner, 1985). For further details, Sturm (2001) and Schillereff (2015) describe the deployment procedures at Lake Baikal and Brotherswater (Cumbria), respectively.

Table 2. Design parameters and deployment intervals used in published sediment trapping research.

Height (cm)	Diameter (cm)	Aspect Ratio	# of joined tubes (distance apart)	Sampling Interval	Reference
75	11	6.8:1	2	Variable	Schillereff, 2015
160	20	8:1	1	1.5-3.5 d	Waples and Klump, 2013
45.6	7.6	6		Weekly	Gelda, 2012
65	6	10.8:1	4	Monthly ¹	Tylmann <i>et al.</i> , 2011
50.8	10.1	5:1		Annual	McDonald <i>et al.</i> , 2010
91.4	8.9	10.3:1	4	7-10 d	Fortino <i>et al.</i> , 2009
45.6	7.6	6:1	3	Weekly	Effler <i>et al.</i> , 2006
36.4	5.2	7:1	3	Annual	Rose and Monteith, 2005
25.5	5.08	5:1	1	Before and after runoff events	Matisoff <i>et al.</i> , 2005
32.4	5.4	6:1	5	14 d	Horppila and Nurminen, 2005
78	15	5.2:1	3	Monthly during ice-free conditions	Chu <i>et al.</i> , 2005
80	11	7.3	6	20-41 d	Foster <i>et al.</i> , 2003
100	9.1	11:1	2	11 d ¹	Ohlendorf and Sturm, 2001
100	10	10:1	2 (40 cm)		Douglas and Rippey, 2000
100	10	10:1	2	14 d	Raubitschek <i>et al.</i> , 1999
50	10	5:1	1		Eadie, 1997
160	20	8:1			Eadie, 1997
60	5.1	11.8:1	4	One year	Flower, 1990
15	9 (12 at collar)		4 (30 cm)	4 m	Chambers and Parker, 1979

¹During ice-free conditions. d = day; m = month

Sediment trap deployment

Depending on the research question being asked and the lake configuration, traps may be installed at one or multiple water depths and at one or more mooring sites. For example, river-borne sediment plumes diffusing across the lake as either over-, inter- or underflows (Figure 2) may be detected by installing arrays at the sediment surface, near the thermocline and in the upper water column. Paired, replicate traps (Figure 3a) yield insignificant variability in terms of dry weight of collected sediment ($\pm 10\%$; Bloesch and Burns, 1980); their deployment in small lakes or zones of low sedimentation is useful to ensure adequate sample size for further laboratory-based analyses (Chambers and Parker, 1979). Sediment trap arrays can be deployed to great depth (e.g., ~1390 m in Lake Baikal; Sturm, 2001; Ryves *et al.*, 2003).

The research question will most likely guide the frequency of trap collection. Palaeoecological studies may only require annual sampling (e.g., Woodbridge and Roberts, 2010). Systems dominated by low-frequency, intense sediment delivery events (e.g., Arctic lakes with nival hydrological regimes) may require daily (Cockburn and Lamoureux, 2008) or near-daily (Dugan *et al.*, 2009) trap replacement. Investigating internal sedimentation processes will necessitate multiple moorings positioned with regard to the inflow position and basin configuration. Douglas and Rippey (2000) installed paired traps at five locations in Lough Neagh while Lewis *et al.* (2002) employed 33 traps at twelve mooring locations in Bear Lake (arctic Canada).

Other considerations include the risk of contamination from algae growing on external trap casings or the mooring rope (Chu *et al.*, 2005) and the addition of a preserving agent. Bloesch and Burns (1980) and Håkanson *et al.* (1989) concluded that the potential biological effects of added chemicals are significantly negative, yet some researchers continue to add chloroform (Eadie, 1997) or sodium azide (NaN_3 ; Chang *et al.*, 2013) to avoid sediment decomposition. Injecting oasis foam or a similar reagent to the capturing vessel upon retrieval may enable the stratigraphy to be preserved (Woodbridge and Roberts, 2010).

Mass accumulation rates (MAR; $\text{mg cm}^{-2} \text{d}^{-1}$) per sampling interval can be calculated from Equation 1 using dried sediment weights, where m is the total dry mass (mg) in each container, d is the number of days during the sampling interval and r is the radius of the trapping vessel (cm^2).

$$MAR = \frac{m}{d} / \pi r^2 \quad (\text{Equation 1})$$

Sequencing sediment traps

Computer-controlled traps offer increased temporal resolution and can shed light on short-lived hydrometeorological controls on sedimentation behaviour as multiple bottles housed within an internal carousel rotate under the cylindrical tube at pre-programmed intervals (Eadie, 1997; Muzzi and Eadie, 2002). Trap and circuit board schematics are published from Baker and Milburn (1983) and Muzzi and Eadie (2002).

Summary and recommendations

Method selection

Choosing to deploy SSC sensors or sediment traps will largely be guided by the research question(s). Sediment traps provide a time-averaged signal, whereas SSC devices can measure at sufficiently high temporal resolution to investigate sub-daily interactions between hydrometeorological events, such as wind, and SS dynamics (e.g., Gilbert and Lamoureux, 2004). Sediment traps are appropriate for monitoring nival flood events (Gilbert and Butler, 2004) that may be logistically challenging to monitor in real-time and may offer opportunities to inspect the internal sedimentology of a flood deposit if sufficient material accumulates in a trapping vessel. Furthermore, captured material can be returned to a laboratory where desk-based instruments with more sensitive detection limits may be available. Conversely, sediment traps do not provide information on the relationship between turbidity and suspended sediment concentrations and SSC sensors alongside current meters should be used instead (e.g., Schiefer and Gilbert, 2008). SSC sensors may also be able to detect particles flocculating in the water column (e.g., Droppo *et al.*, 1997) and may be more appropriate when investigating the influence of short-lived events such as internal waves or turbidity currents on SS dynamics.

Sediment traps stationed at different heights may detect differences in the sediment regime under thermal stratification or evidence that a turbidity current has occurred (e.g., Schillereff, 2015) but SSC techniques will be more effective at discerning their duration, spatial extent across the basin floor and the critical SS threshold at which they are triggered in individual lakes (Gilbert and Butler, 2004; Schiefer and Gilbert, 2008).

Long-term process monitoring

Although case studies are rare, a detailed perspective on depositional mechanisms and the implications for the accumulated basal sediments can be acquired through integrating *in situ* measurements of SSC, turbidity, thermal stratification and other lake processes with sediment trapping and comparisons to adjacent sediment core records (e.g., Gilbert and Butler, 2004; Gilbert *et al.*, 2006). Such an approach can yield unprecedented information on sediment dynamics during discrete floods and the potential for a palaeoflood sedimentary signature to be preserved (Gilbert *et al.*, 2006) and foster greater confidence in palaeoenvironmental reconstructions produced from long lake sediment cores.

Conclusions

A broad range of equipment is available to researchers investigating limnological processes and technological advances are continually being made to improve the temporal and spatial resolution of data collection. However, the complexity of internal lake dynamics necessitates careful consideration of appropriate methods when investigating suspended sediment dynamics. Each type of suspended sediment monitoring device has drawbacks, in particular related to variable particle sizes, and choice should be guided by knowledge of the field site. Implementing a programme of analysis that combines multiple methods of SSC monitoring is ideal but financially and logistically more demanding. Sediment trapping is a time and labour-intensive undertaking that potentially offers valuable insight into hydrometeorological controls on sediment flux, provided the design and deployment protocols are appropriate. The coupling of catchment sediment budget

calculations with lake-based sediment trapping could be a useful pathway for future geomorphological research.

Acknowledgements

The author would like to thank Dr Jennifer Clear and Dr Neil Macdonald for commenting on an earlier draft. The constructive comments of two anonymous reviewers are gratefully acknowledged. The author was awarded funding from the School of Environmental Sciences, University of Liverpool to construct and deploy the sediment traps in Brotherswater.

References

- Agrawal YC, Pottsmith HC. 2000. Instruments for particle size and settling velocity observations in sediment transport. *Marine Geology*, **168**: 89–114. DOI:10.1016/S0025-3227(00)00044-X
- Agrawal YC, Whitmire A, Mikkelsen O A., Pottsmith HC. 2008. Light scattering by random shaped particles and consequences on measuring suspended sediments by laser diffraction. *Journal of Geophysical Research* **113**: C04023. DOI:10.1029/2007JC004403
- Andrews SW, Nover DM, Reuter JE, Schladow SG. 2011. Limitations of laser diffraction for measuring fine particles in oligotrophic systems: Pitfalls and potential solutions. *Water Resources Research* **47**: W05523. DOI:10.1029/2010WR009837
- Baker ET, Milburn HBI. 1983. An instrument system for the investigation of particle fluxes. *Continental Shelf Research* **1**: 425–435.
- Best JL, Kostaschuk R A., Peakall J, Villard P V., Franklin M. 2005. Whole flow field dynamics and velocity pulsing within natural sediment-laden underflows. *Geology* **33**: 765–768. DOI:10.1130/G21516.1
- Bloesch J. 1996. Towards a new generation of sediment traps and a better measurement/understanding of settling particle flux in lakes and oceans: A hydrodynamical protocol. *Aquatic Sciences* **58**: 283–296.

- Bloesch J, Burns N. 1980. A critical review of sedimentation trap technique. *Schweizerische Zeitschrift für Hydrologie* **42**: 15–55.
- Blomqvist S, Håkanson L. 1981. A review on sediment traps in aquatic environments. *Archiv für Hydrobiologie* **91**: 101–132.
- Blomqvist S, Kofoed C. 1981. Sediment trapping - a subaquatic in situ experiment. *Limnology and Oceanography* **26**: 585–590.
- Chambers K, Parker A. 1979. A modified design for lake-section traps and a simple method for relocating them. *Earth Surface Processes and Landforms* **4**: 73–76.
- Chang AS, Bertram M A., Ivanochko T, Calvert SE, Dallimore A, Thomson RE. 2013. Annual record of particle fluxes, geochemistry and diatoms in Effingham Inlet, British Columbia, Canada, and the impact of the 1999 La Niña event. *Marine Geology* **337**: 20–34. DOI:10.1016/j.margeo.2013.01.003
- Chu G, Liu J, Schettler G, Li J, Sun Q, Gu Z, Lu H, Liu Q, Liu T. 2005. Sediment Fluxes and Varve Formation in Sihailongwan, a Maar Lake from Northeastern China. *Journal of Paleolimnology* **34**: 311–324. DOI:10.1007/s10933-005-4694-0
- Cockburn JMH, Lamoureux SF. 2008. Inflow and lake controls on short-term mass accumulation and sedimentary particle size in a High Arctic lake: implications for interpreting varved lacustrine sedimentary records. *Journal of Paleolimnology* **40**: 923–942. DOI:10.1007/s10933-008-9207-5
- Dearing JA. 1991. Lake sediment records of erosional processes. *Hydrobiologia* **214**: 99–106.
- Douglas RW, Rippey B. 2000. The random redistribution of sediment by wind in a lake. *Limnology and Oceanography* **45**: 686–694.
- Droppo I, Leppard G, Flannigan D, Liss S. 1997. The freshwater floc: a functional relationship of water and organic and inorganic floc constituents affecting suspended sediment properties. *Water, Air and Soil Pollution* **99**, 43–54.
- Dugan HA, Lamoureux SF, Lafrenière MJ, Lewis T. 2009. Hydrological and sediment yield response to summer rainfall in a small high Arctic watershed. *Hydrological Processes* **1526**: 1514–1526.
- Eadie BJ. 1997. Probing particle processes in Lake Michigan using sediment traps. *Water, Air and Soil Pollution* **99**: 133–139.
- Effler SW, Matthews DA, Kaser JW, Prestigiacomo AR, Smith DG, Prestigiacomo R. 2006. Runoff event impacts on a water supply reservoir: suspended sediment loading, turbid plume behavior, and sediment deposition. *Journal of the American Water Resources Association* **10595**: 1697–1710.
- Felix D, Albayrak I, Boes RM. 2013. Laboratory investigation on measuring suspended sediment by portable laser diffractometer (LISST) focusing on particle shape. *Geo-Marine Letters* **33**: 485–498. DOI: 10.1007/s00367-013-0343-1
- Flower RJ. 1990. Seasonal changes in sedimenting material collected by high aspect ratio sediment traps operated in a holomictic eutrophic lake. *Hydrobiologia* **214**: 311–316.
- Fortino K, Hershey A, Keyse MD, Whalen S. 2009. Summer sedimentation in six shallow arctic lakes. *Hydrobiologia* **621**: 75–84
- Foster GC, Dearing J, Jones R, Crook D, Siddle D, Harvey A, James P, Appleby P, Thompson R, Nicholson J, Loizeau J-L. 2003. Meteorological and land use controls on past and present hydro-geomorphic processes in the pre-alpine environment: an integrated lake-catchment study at the Petit Lac d'Annecy, France. *Hydrological Processes* **17**: 3287–3305. DOI: 10.1002/hyp.1387
- Foster IDL, Collins A. L, Naden PS, Sear D A., Jones JI, Zhang Y. 2011. The potential for paleolimnology to determine historic sediment delivery to rivers. *Journal of Paleolimnology* **45**: 287–306. DOI:10.1007/s10933-011-9498-9
- Gardner WD. 1980. Sediment trap dynamics and calibration: a laboratory evaluation. *Journal of Marine Research* **38**: 17–39.
- Gardner WD. 1985. The effect of tilt on sediment trap efficiency. *Deep-Sea Research* **32**: 349–361.

- Gelda R. 2012. Calibration and application of a sediment accumulation rate model – a case study. *Inland Waters* **2**: 23–36. DOI: 10.5268/IW-2.1.454
- Gilbert R, Butler R. 2004. The Physical Limnology and Sedimentology of Meziadin Lake, Northern British Columbia, Canada. *Arctic, Antarctic, and Alpine Research* **36**, 33–41.
- Gilbert R, Lamoureux S. 2004. Processes affecting deposition of sediment in a small, morphologically complex lake. *Journal of Paleolimnology* **31**, 37–48.
- Gilbert R, Crookshanks S, Hodder K, Spagnol J, Stull R. 2006. The Record of an Extreme Flood in the Sediments of Montane Lillooet Lake, British Columbia: Implications for Paleoenvironmental Assessment. *Journal of Paleolimnology* **35**, 737–745.
- Gray JR, Gartner JW. 2009. Technological advances in suspended-sediment surrogate monitoring. *Water Resources Research* **45**: W00D29. DOI:10.1029/2008WR007063
- Guerrero M, Rütther N, Szupiany RN. 2012. Laboratory validation of acoustic Doppler current profiler (ADCP) techniques for suspended sediment investigations. *Flow Measurement and Instrumentation* **23**: 40–48. DOI:10.1016/j.flowmeasinst.2011.10.003
- Guerrero M, Szupiany RN, Amsler M. 2011. Comparison of acoustic backscattering techniques for suspended sediments investigation. *Flow Measurement and Instrumentation* **22**: 392–401. DOI:10.1016/j.flowmeasinst.2011.06.003
- Håkanson L, Floderus S, Wallin M. 1989. Sediment trap assemblages - a methodological description. *Hydrobiologia* **176/177**: 481–490.
- Håkanson L, Jansson M. 1983. *Principles of Lake Sedimentology*. Blackburn Press, USA.
- Hargrave BT, Burns NM. 1979. Assessment of sediment trap collection efficiency. *Limnology and Oceanography* **24**: 1124–1136.
- Horppila J, Nurminen L. 2005. Effects of calculation procedure and sampling site on trap method estimates of sediment resuspension in a shallow lake. *Sedimentology* **52**: 903–913. DOI:10.1111/j.1365-3091.2005.00726.x
- Hubbart J A., Kellner E, Freeman G. 2014. A case study considering the comparability of mass and volumetric suspended sediment data. *Environmental Earth Sciences* **71**: 4051–4060. DOI:10.1007/s12665-013-2788-y
- Kostaschuk R, Best J, Villard P, Peakall J, Franklin M. 2005. Measuring flow velocity and sediment transport with an acoustic Doppler current profiler. *Geomorphology* **68**: 25–37. DOI:10.1016/j.geomorph.2004.07.012
- Lamoureux SF. 2005. A sediment accumulation sensor for use in lacustrine and marine sedimentation studies. *Geomorphology* **68**: 17–23. DOI:10.1016/j.geomorph.2004.02.014
- Lerman A, Imboden DM, Gat J. 1995. *Physics and Chemistry of Lakes*. Springer-Verlag, Berlin.
- Lewis AJ, Rasmussen TC. 1999. Determination of suspended sediment concentrations and particle size distributions using pressure measurements. *Journal of Environmental Quality* **28**: 1490–1496.
- Lewis T, Gilbert R, Lamoureux SF. 2002. Spatial and temporal changes in sedimentary processes at proglacial Bear Lake, Devon Island, Nunavut, Canada. *Arctic, Antarctic and Alpine Research* **34**: 119–129.
- Loizeau J-L, Arbouille D, Santiago S, Vernet J-P. 1994. Evaluation of a wide range laser diffraction grain size analyser for use with sediments. *Sedimentology* **41**: 353–361.
- Ludwig KA, Hanes DM. 1990. A laboratory evaluation of optical backscatterance suspended solids sensors exposed to sand-mud mixtures. *Marine Geology* **94**: 173–179.
- McDonald CP, Urban NR, Barkach JH, McCauley D. 2010. Copper profiles in the sediments of a mining-impacted lake. *Journal of Soils and Sediments* **10**: 343–348. DOI:10.1007/s11368-009-0171-0.
- Mackereth FJH. 1966. Some chemical observations on post-glacial lake sediments. *Philosophical Transactions of The Royal Society B: Biological Sciences* **250**: 165–213.

- Matisoff G, Wilson C, Whiting P. 2005. The $^{7}\text{Be}/^{210}\text{Pb}$ ratio as an indicator of suspended sediment age or fraction new sediment in suspension. *Earth Surface Processes and Landforms* **30**: 1191-1201.
- Mikkelsen O, Pejrup M. 2001. The use of a LISST-100 laser particle sizer for in-situ estimates of floc size, density and settling velocity. *Geo-Marine Letters* **20**: 187–195. DOI:10.1007/s003670100064
- Muzzi RW, Eadie BJ. 2002. The Design and Performance of a Sequencing Sediment Trap for Lake Research. *Marine Technology Society Journal* **36**: 23–28. DOI:10.4031/002533202787914025
- Ohlendorf C, Sturm M. 2001. Precipitation and dissolution of calcite in a Swiss high alpine lake. *Arctic, Antarctic and Alpine Research* **33**: 410-417.
- Onset. 2013. HOBO Water Temperature Pro v2 Logger. Specification Sheet [online] Available from: [http://www.onsetcomp.com/files/data-sheet/Onset HOBO Water Temp Pro v2 Data Logger.pdf](http://www.onsetcomp.com/files/data-sheet/Onset_HOBO_Water_Temp_Pro_v2_Data_Logger.pdf) (Accessed 11 November 2014)
- Perks MT. 2014. Section 3.3.6: Suspended sediment sampling. In: Clarke LE, Nield JM (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology: London, UK.
- Raubitschek S, Lücke A, Schleser GH. 1999. Sedimentation patterns of diatoms in Lake Holzmaar, Germany - (on the transfer of climate signals to biogenic silica oxygen isotope proxies). *Journal of Paleolimnology* **21**: 437–448.
- Reichel G. 1998. Suspended sediment monitoring: use of acoustic Doppler current profiler. In: Herschy RW and Fairbridge RW (Eds.) *Encyclopedia of Hydrology and Water Resources*, Kluwer Academic Publishers, The Netherlands: Dordrecht; 638–644.
- Rose NL, Monteith DT. 2005. Temporal trends in spheroidal carbonaceous particle deposition derived from annual sediment traps and lake sediment cores and their relationship with non-marine sulphate. *Environmental Pollution* **137**: 151–163. DOI:10.1016/j.envpol.2004.12.022
- Ryves DB, Jewson DH, Sturm M, Battarbee RW, Flower RJ, Mackay AW, Granin NG. 2003. Quantitative and qualitative relationships between planktonic diatom communities and diatom assemblages in sedimenting material and surface sediments in Lake Baikal, Siberia. *Limnology and Oceanography* **48**: 1643–1661. DOI:10.4319/lo.2003.48.4.1643
- Sassi MG, Hoitink A. JF, Vermeulen B. 2012. Impact of sound attenuation by suspended sediment on ADCP backscatter calibrations. *Water Resources Research* **48**: W09520. DOI:10.1029/2012WR012008
- Schiefer E, Gilbert R. 2008. Proglacial sediment trapping in recently formed Silt Lake, Upper Lillooet Valley, Coast Mountains, British Columbia. *Earth Surface Processes and Landforms* **33**, 1542-1556.
- Schillereff DN. 2015. *Lake sediment records of flood frequency and magnitude*. Unpublished PhD Thesis, University of Liverpool, 320 pp.
- Schillereff DN, Chiverrell RC, Macdonald N, Hooke JM. 2014. Flood stratigraphies in lake sediments: A review. *Earth-Science Reviews* **135**: 17–37. DOI:10.1016/j.earscirev.2014.03.011
- Stokes GG. 1851. On the effect of the internal friction of fluids on the motion of pendulums. *Transactions of the Cambridge Philosophical Society* **IX**: 8.
- Sturm M. 2001. Sediment traps for evaluation of particle dynamics in Lake Baikal. CONTINENT Protocol #1. EAWAG, Dübendorf, Switzerland.
- Tylmann W, Szpakowska K, Ohlendorf C, Woszczyk M, Zolitschka B. 2011. Conditions for deposition of annually laminated sediments in small meromictic lakes: a case study of Lake Suminko (northern Poland). *Journal of Paleolimnology* **47**: 55–70. DOI:10.1007/s10933-011-9548-3
- Waples JT, Klump JV. 2013. Vertical and horizontal particle transport in the coastal waters of a large lake: An assessment by sediment trap and thorium-234 measurements. *Journal of Geophysical*

Research: Oceans **118**: 5376–5397.

DOI:10.1002/jgrc.20394

Wood PJ, Armitage PD. 1997. Biological Effects of Fine Sediment in the Lotic Environment. *Environmental Management* **21**: 203–217.

Woodbridge J, Roberts N. 2010. Linking neo- and palaeolimnology: a case study using crater lake diatoms from central Turkey. *Journal of Paleolimnology* **44**: 855–871.
DOI:10.1007/s10933-010-9458-9

Wren DG, Barkdoll BD, Kuhnle RA, Darrow RW. 2000. Field techniques for suspended-sediment measurement. *Journal of Hydraulic Engineering* **126**: 97–104.

3.11.1. Hyporheic Zone: *In Situ* Sampling

Matilda Biddulph¹

¹ Department of Environmental and Geographical Sciences, School of Science and Technology, University of Northampton (matilda.biddulph@northampton.ac.uk)



ABSTRACT: The hyporheic zone (HZ) is a fundamental part of a river ecosystem. It is the zone of interaction between groundwater and channel water. Movement of channel water and groundwater occurs through changes in hydraulic pressure and occurs at multiple scales. The HZ generally contains higher chemical concentrations and lower dissolved oxygen levels than the channel due to lower rates of exchange, lack of sunlight and longer residence times. Organisms residing in this zone are known as the hyporheos, and can also include benthic invertebrates dwelling there temporarily for refugia, or as part of a life cycle. Research in this area may involve sampling invertebrate distribution and abundance, flow dynamics between and within the HZ and channel, hydrochemistry of the HZ and characterisation of the substrate. Common sampling techniques for these areas of research involve coring, tracing and *in situ* extraction. This article briefly introduces these techniques, with primary focus on *in situ* extraction, in particular pump sampling, which is an effective method for monitoring hydrochemistry.

KEYWORDS: pump sampling; freeze coring; hyporheos; Bou-Rouch; hydrochemistry; interstitial water

Introduction

Defining the Hyporheic Zone

The hyporheic zone (HZ) is a highly dynamic region, loosely defined as the saturated interstitial sediment below the streambed and adjacent riverbanks where exchange of channel water and groundwater occurs (White, 1993; Figure 1). This definition reflects three decades of a gradual inclusion of the HZ into fluvial research, since early recognition of its importance (e.g. Orghidan, 1959; Schwoerbel, 1961; Williams and Hynes, 1974).

Until the late 1980s, there was a focus on defining the extent of the HZ, rather than looking at its functional significance or characteristics. Definition was based mainly on invertebrate sampling (e.g. Schwoerbel, 1961; Williams and Hynes, 1974; Williams, 1989), by the distribution of surface and subsurface organisms and their lateral and vertical extents (Boulton *et al.*, 2010). It was expressed in the late 1980s that there are a large number of factors that control the HZ,

as well as the distribution of invertebrates themselves, rendering this sampling strategy insufficient (e.g. Danielopol, 1989; White, 1993).

This recognition led to a number of investigations into the physicochemical properties of the HZ (e.g. Triska *et al.*, 1989), attempting to distinguish groundwater and channel water contributions to subsurface regions. Gibert *et al.* (1990) further developed our understanding of the HZ, describing its role as a “dynamic ecotone”. The boundaries of this ecotone are likely to change temporally and spatially with response to hydrological behaviour and characteristics of the sediment (Gibert *et al.*, 1990; Boulton *et al.*, 2010; Williams *et al.*, 2010). Vervier *et al.* (1992) concluded that a HZ depended on the elasticity, permeability, biodiversity and connectivity of the ecotone. This pushed focus towards its functional significance, both ecologically and hydrologically, and its influence on the surface stream (Boulton *et al.*, 2010).

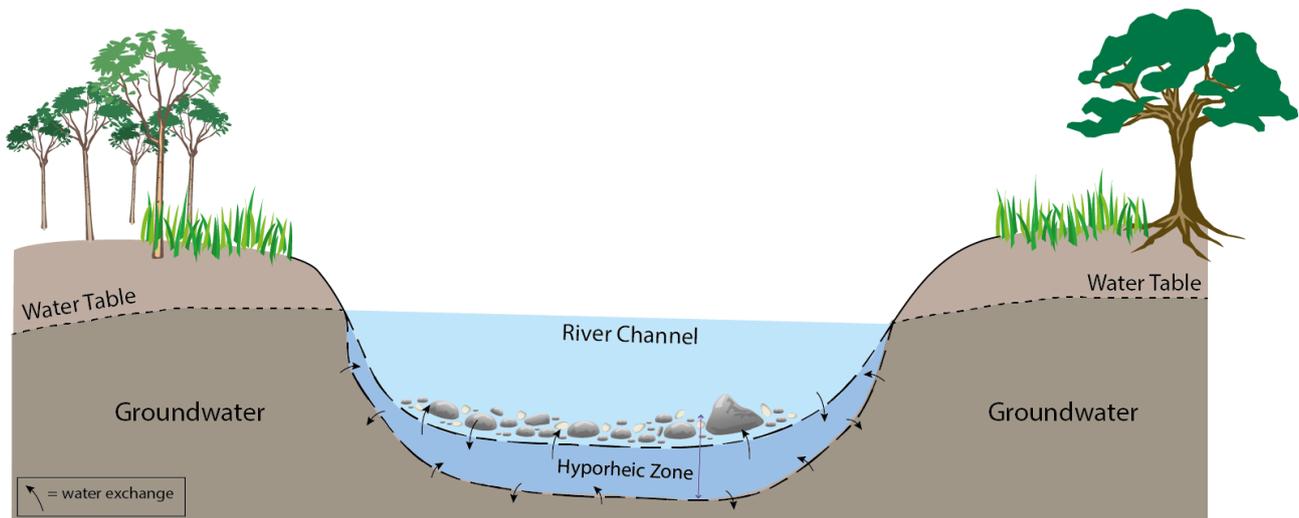


Figure 1: The approximate position of the hyporheic zone from a cross-sectional view of a river catchment. Typically considered to extend between a few centimetres to a few metres in size (Stanford and Ward, 1988).

Characteristics of the Hyporheic Zone

The exchange of water between a surface stream and the HZ is governed by pressure differences at multiple scales (Wondzell, 2006; Boulton, 2007; Boulton *et al.*, 2010). At the catchment scale, exchange occurs when there are differences in stream and groundwater levels. At a smaller, reach scale, interactions between channel flow and geomorphological features create pressure variations, such as slope, riffle-pool sequences (Elliott and Brooks, 1997) and step sequences (Kasahara and Wondzell, 2003). A common pressure difference can be found along a riffle (Figure 2); shallowing of channel water at the head of a riffle creates higher pressures, causing downwelling of surface waters into the interstitial areas. Deepening of the water at the tail end of the riffle produces lower pressures, causing upwelling of groundwater into the surface stream (Wondzell, 2006; Boulton, 2007).

The channel water and groundwater that combine to make up the interstitial water of the HZ have characteristic differences. In a healthy channel the water is typically clean, oxygenated, has variable discharge, short residence times and changing physicochemical conditions (Boulton *et al.*, 1998). In contrast, groundwater is often characterised by long residence times in the surrounding catchment, with chemically reducing conditions that can decrease dissolved oxygen levels, bring high chemical concentrations and lower the temperature (Soulsby *et al.*, 2001; Malcolm *et al.*, 2004).

The combination of these waters vary temporally, due to changes in water depth and catchment water levels, and spatially, due to changes in channel morphology (Malcolm *et al.*, 2004).

Invertebrate communities living in the HZ are known as the hyporheos, and include permanent and semi-permanent inhabitants (Williams and Hynes, 1974; Stubbington, 2012). The differences in hydrochemistry between the HZ and channel, and within the HZ itself, have a broader significance with regard to the organisms that dwell in this zone temporarily. The HZ has been recognised as a place of refugia, for example during times of drought, extreme flows, or high levels of pollution (Williams and Hynes, 1974; Dole-Olivier *et al.*, 1997; Wood *et al.*, 2010; Stubbington *et al.*, 2011; Stubbington, 2012; Crossman *et al.*, 2012). Organisms may also use the HZ as a refuge during vulnerable stages of their life cycle, a prime example are salmonid eggs. Much hyporheic research has focused on the survival of salmonid eggs in redds constructed in gravel-bed rivers (e.g. Soulsby *et al.*, 2001; Geist *et al.*, 2002; Malcolm *et al.*, 2004, 2010).

A number of sampling techniques have been developed, using a variety of methods, to try to better understand the interactions and habitat preferences of the hyporheos, the dynamics of hydrochemistry, and the level of influence that the HZ has on stream health.

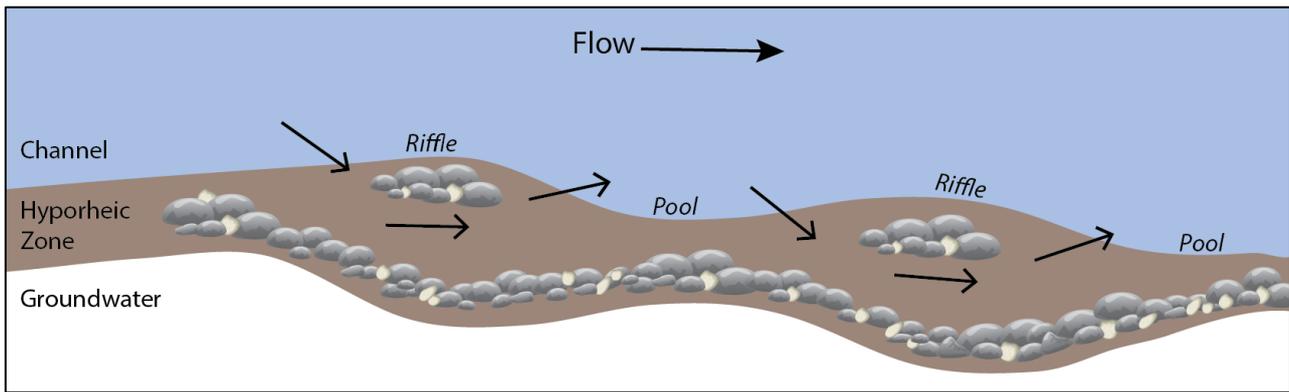


Figure 2: Directional flow of water through riffle-pool sequences in the hyporheic zone. Arrows indicate regions of downwelling and upwelling due to changes in hydraulic pressure.

Sampling the Hyporheic Zone

Overview of Sampling Techniques

Sampling the HZ is difficult, because it is relatively inaccessible and a fragile system (Palmer, 1993). During sampling, the natural conditions of the HZ may be altered and therefore may not provide a true representation (Palmer, 1993). The three main techniques followed for sampling the HZ are coring, tracing, and *in situ* extraction; for each of these techniques there are a variety of methods that can be used. The methods chosen will depend on the research that is being undertaken.

It was suggested by Palmer (1993) that it is best to compare a number of methods to provide better accuracy and precision. Useful comparisons of sampling procedures can be found in Fraser and Williams (1997), Hunt and Stanley (2000), Scarsbrook and Halliday (2002), Dole-Olivier *et al.* (2014) and Tanaka *et al.* (2014). Table 1 summarises relevant literature outlining the HZ methods and associated research themes.

Coring

If examining invertebrate distribution or abundance, it is necessary to sample the natural, undisturbed conditions beneath the bed. For this, a freeze core is often used (e.g. Stocker and Williams, 1972; Pugsley and Hynes, 1983; Adkins and Winterbourn, 1999; Scarsbrook and Halliday, 2002; Nogaro *et al.*, 2010; Toran *et al.*, 2013), which freezes the sediment and water within the cylinder prior to extraction, including any organism within it. However, there are a number of complications with this method, such as

organism escape during disturbance (Bretschko, 1985), or ineffective coring due to the presence of large boulders (Fraser and Williams, 1997; Toran *et al.*, 2013). Another limitation to this method is that it is expensive and labour-intensive (Hunt and Stanley, 2000). An alternative to freeze coring is the use of a standpipe corer (Williams and Hynes, 1974; Franken *et al.*, 2001; Storey and Williams, 2004). This involves extracting a core of the bed and immediately storing the sample in sealed containers (Storey and Williams, 2004).

Another coring method that has been developed for research into the hyporheos is the use of colonisation pots (Nelson and Roline, 2003; Crossman *et al.*, 2012). These generally consist of cages filled with substrate that attempts to mimic the characteristics of that riverbed. These are submerged into the HZ and left to colonise for a number of weeks (Crossman *et al.*, 2012). Upon collection, the baskets are removed and the hyporheos can be analysed. This can be a highly destructive sampling method, as sections of the riverbed must be disturbed during insertion and removal of the cages (Fraser and Williams, 1997; Scarsbrook and Halliday, 2002). There is also no guarantee that the substrate will replicate that of the HZ surrounding it (Fraser and Williams, 1997).

Tracing

If investigating rates of hyporheic water exchange and nutrient processes, tracer experiments are the favoured technique (e.g. Jonsson *et al.*, 2003; Wondzell, 2006; Toran *et al.*, 2013). In general, these methods involve the injection of one or more non-

Table 1: Examples of studies that have employed the different sampling methods for the main research areas of hyporheic zone science.

	Hydrochemistry	Flow Dynamics	Invertebrate Sampling	Substrate characterisation
Freeze Coring	-	Toran <i>et al.</i> (2013)	Pugsley and Hynes (1983); Adkins and Winterbourn (1999); Scarsbrook and Halliday (2002)	Olsen and Townsend (2005); Nogaro <i>et al.</i> (2010); Toran <i>et al.</i> (2013)
Standpipe Coring	Franken <i>et al.</i> (2001)	-	Williams and Hynes (1974); Storey and Williams (2004)	-
Colonisation Pots	-	-	Scarsbrook and Halliday (2002); Nelson and Roline (2003); Crossman <i>et al.</i> (2012)	-
Tracers	Jonsson <i>et al.</i> (2003); Van Stempvoort <i>et al.</i> (2011)	Wondzell (2006); Pinay <i>et al.</i> (2009); Birkinshaw and Webb (2010); McCallum <i>et al.</i> (2012); Toran <i>et al.</i> (2013)	-	-
Pump Sampling	Findlay <i>et al.</i> (1993); Riss <i>et al.</i> (2008); Cornut <i>et al.</i> (2012)	-	Hunt and Stanley (2000); Scarsbrook and Halliday (2002)	-
Piezometers	Soulsby <i>et al.</i> (2001); Malcolm <i>et al.</i> (2004); Hlavacova <i>et al.</i> (2005); Lewandowski <i>et al.</i> (2011)	Grimaldi and Chaplot (2000); Wondzell (2006); Ibrahim <i>et al.</i> (2010)	-	-

reactive tracer into the channel. Wells at specified locations in the riverbed then monitor the travel times or chemical transformations in the HZ (Triska *et al.*, 1993; Harvey and Wagner, 2000; Jonsson *et al.*, 2003; Pinay *et al.*, 2009). Tracers used are often naturally available environmental substances, generally considered to be conservative (Engelhardt *et al.*, 2011). These tracers include stable isotopes, chloride (Pinay *et al.*, 2009), pharmaceutical compounds (Van Stempvoort *et al.*, 2011), and temperature (Birkinshaw and Webb, 2010). Tracer experiments are advantageous against other techniques, in that they can be carried out at larger scales (Harvey *et al.*, 1996). However, they are found to be less

reliable at flows higher than base levels (Harvey *et al.*, 1996).

In situ

In situ extraction is a widely used technique, appropriate for sampling the spatial and/or temporal variation in hyporheic hydrochemistry and invertebrates. *In situ* extraction involves the installation of a permanent or semi-permanent well into the riverbed, either submerged as a piezometer (Hlavacova *et al.*, 2005; Ibrahim *et al.*, 2010; Lewandowski *et al.*, 2011), or reaching the surface of the channel through a rigid tube containing small perforations (Soulsby *et al.*, 2001; Cornut *et al.*, 2012). These wells then provide an opening from which samples can

be extracted. The remainder of the article will focus on an explanation of *in situ* extraction sampling. For detailed explanations of coring and tracer techniques see *Geomorphological Techniques* Sections 3.11.2 and 3.11.3, respectively.

Common methods for *in situ* sampling the hyporheos follow those developed by Bou and Rouch (1967) and Boulton *et al.* (1992) (e.g. Marmonier *et al.*, 1992; Boulton *et al.*, 1997, 2003; Hunt and Stanley, 2000; Wood *et al.*, 2010; Stubbington *et al.*, 2011). This method has also been adapted for sampling the hydrochemistry of the HZ (Findlay *et al.*, 1993; Soulsby *et al.*, 2001; Youngson *et al.*, 2004; Cornut *et al.*, 2012). Water and invertebrates are pumped out of the wells by creating a vacuum, and this can then be analysed for hydrochemical properties, invertebrate type, abundance, and diversity. A detailed example of this adapted method, known as pump sampling, is presented in the following case study.

Case study: Water quality of the hyporheic zone in relation to the survival of Freshwater Pearl Mussels, River Esk, North Yorkshire.

The following section describes an example of the pump sampling method, used in a small study of the HZ of the River Esk, North Yorkshire. The aim of the study was to determine reasons for lack of freshwater pearl mussel (*Margaritifera margaritifera*) survival, by looking at comparisons between the hydrochemistry within the HZ and in the channel itself, in varying rates of flow. This can be found in greater detail in Biddulph (2012).

Freshwater pearl mussels are slow-growing molluscs with a complex life cycle (Bauer, 1992). Juvenile mussels spend five years buried beneath the substrate of a riverbed before emerging as mature adults (Geist, 2010). It is at this juvenile stage that they are most vulnerable, as they are very sensitive to changes in water quality (Skinner *et al.*, 2003). In the last century the previously large populations of mussels have declined dramatically, with very few signs of recovery (Reid *et al.*, 2012).

Freshwater pearl mussels require clean rivers that are low in nutrients (Geist *et al.*, 2006),

and depend on substratum that is relatively stable and well aerated (Geist and Auerwald, 2007). The main concern around the lack of revival of these mussel populations are the lack of juvenile recruitment (Geist, 2010), so research needs to focus on their habitats during the juvenile stage. A preliminary study was carried out to determine the hyporheic water quality along the mussel-dwelling reach of the River Esk. Sites were selected based on where pearl mussels were known to live, as well as locations where they theoretically could live. Comparable locations were also selected in areas of unsuitable habitat.

A number of pump samplers were constructed (Figure 3), based loosely on those used by Soulsby *et al.* (2001). Hard plastic tubing was used to create wells that were 1 m in length and 2 cm in diameter. Small holes were drilled around the circumference over a length of 20 cm, so that once placed in the riverbed the wells sampled a depth from 10-30 cm below the bed. Augers were used to bore into the riverbed, so that the tubes could be inserted, with a bung in the bottom so that they did not infill with sediment.

When the wells were not in use, a bung was placed in the top (Figure 3b), so that channel water could not enter the tube in periods of high flow. During sampling, a different bung was inserted, containing a narrower, flexible tube running through it and down to the bottom of the well. This tube extended to a one litre bottle, which in turn was connected to a manual vacuum pump (Figure 3a). The vacuum draws water through the perforations in the well, up the narrow tube and into the bottle. One litre of water was drawn up from the HZ, which was then immediately tested for pH, dissolved oxygen (DO), conductivity, temperature and redox potential using a multi-parameter probe. This water was then taken to the laboratory for anion and cation analysis. Sampling took place once a fortnight for six months.

The locations chosen for well installation are important, as hyporheic water quality varies greatly at different points on a riverbed. For this project, the focus was on freshwater pearl mussels, so samples were taken in areas where they are likely to inhabit. A suitable habitat would be at the tail end of a

riffle, where flows are moderate enough to re-oxygenate hyporheic water and reduce sedimentation. This is in contrast to the adjacent deeper pools, where residence time

of water in the HZ would be longer, often characterised by lower oxygen levels and higher rates of chemical transformation.

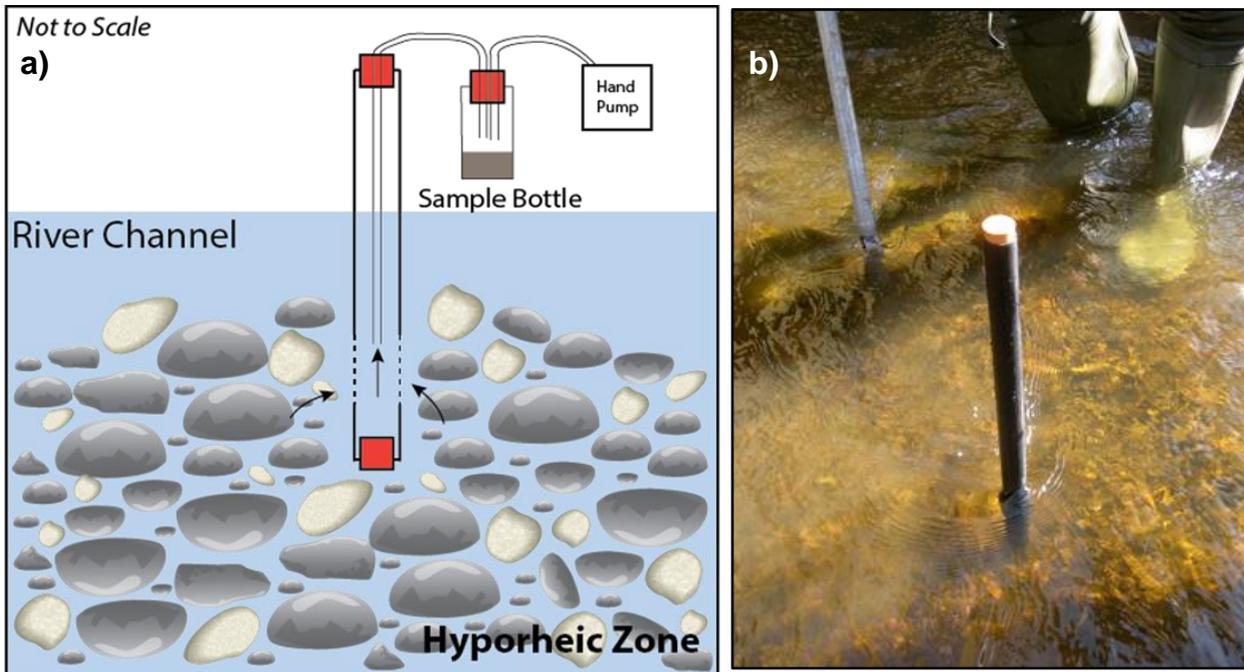


Figure 3: a) Sampling well and pump designed for this study (sediment is not to scale and not of a representative size) b) Sampling well in place on the Risk Esk, North Yorkshire.

Advantages and limitations of pump sampling

Pump sampling is an effective and convenient method. The wells cause minimal disturbance or compaction to the substrate upon insertion, and can therefore be used immediately (Wood *et al.*, 2010). They can then be sampled from at regular intervals, over long time periods (Fraser and Williams, 1997; Wood *et al.*, 2010). Pump samplers are also affordable, with minimal labour required to construct or use them (Hunt and Stanley, 2000). Another advantage of this method is that perforations in the rigid tubes can be concentrated at particular depths to create multi-level samplers (Riss *et al.*, 2008; Rivett *et al.*, 2008). These can give a more detailed picture of the changing water quality with distances between groundwater and channel water.

Despite the wide use and advantages of this method, sampling should be approached with care, as there are many factors that may lead to inaccuracy or inconsistency (Palmer, 1993; Scarsbrook and Halliday, 2002). The

presence of the wells in the riverbed may have created a new, altered habitat, with different micro-scale flow patterns and substrate characteristics (Palmer, 1993; Hunt and Stanley, 2000). It is also noted that pump sampling is biased towards smaller and less tenacious organisms (Fraser and Williams, 1997), therefore may not give a true representation of the hyporheos.

Advice for pump sampling

- It is important that pH, DO, conductivity, temperature and redox potential are tested immediately after extraction, to reduce inaccuracy with changes in environmental conditions. Riss *et al.* (2008) give a detailed account of a more sophisticated method of *in situ* DO testing, where it is measured within the well itself before extraction.
- To minimise equipment loss, it is important to lodge the wells firmly into the bed. It is also sensible to keep the well height above the HZ as short as is feasible to increase stability.

- Carefully note locations of installed wells, as they may be hard to spot when only just above water level. Black tubes were used in this study to reduce aesthetic impact, but in a more remote location white tubes may be more appropriate.
- This technique could be used to study water quality at different hyporheic depths (Riss *et al.*, 2008; Rivett *et al.*, 2008), by changing the location of the perforations or the depth of tube beneath the bed. It could also be undertaken at different temporal frequencies.

Acknowledgements

The author would like to thank the two anonymous reviewers for their invaluable contribution to the improvement of this paper. Thanks must also go to the North York Moors National Park Authority for funding the research in the River Esk. Fieldwork assistance was provided unfailingly by Stuart Crow. Thank you finally to Dr Naomi Holmes and Dr Lucy Clarke for constructive feedback and advice.

References

- Adkins SC, Winterbourn MJ. 1999. Vertical distribution and abundance of invertebrates in two New Zealand stream beds: a freeze coring study. *Hydrobiologia* **400**: 55-62.
- Bauer G. 1992. Variation in the life span and size of the freshwater pearl mussel. *Journal of Animal Ecology* **61(2)**: 425-436.
- Biddulph MF. 2012. *In-stream and hyporheic water quality of the River Esk, North Yorkshire: implications for Freshwater Pearl Mussel habitats*. M.Sc. Thesis, Durham University: UK.
- Birkinshaw SJ, Webb B. 2010. Flow pathways in the Slapton Wood catchment using temperature as a tracer. *Journal of Hydrology* **383**: 269-279.
- Bou C, Rouch R. 1967. Un nouveau champ de recherches sur la faune aquatique souterraine. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences Serie D*. **265(4)**: 369-370.
- Boulton AJ. 2007. Hyporheic rehabilitation in rivers: restoring vertical connectivity. *Freshwater Biology* **52(4)**: 632-650.
- Boulton AJ, Datry T, Kasahara T, Mutz M, Stanford JA. 2010. Ecology and management of the hyporheic zone: stream-groundwater interactions of running waters and their floodplains. *Journal of North American Benthological Society* **29**: 26-40.
- Boulton AJ, Dole-Olivier M-J, Marmonier P. 2003. Optimizing a sampling strategy for assessing hyporheic invertebrate biodiversity using the Bou-Rouch method: within-site replication and sample volume. *Hydrobiologia* **156(4)**: 431-456.
- Boulton AJ, Findlay S, Marmonier P, Stanley EH, Maurice Valett H. 1998. The functional significance of the hyporheic zone in streams and rivers. *Annual Review of Ecological Systems* **29**: 59-81.
- Boulton AJ, Scarsbrook MR, Quinn JM, Burrell GP. 1997. Land-use effects on the hyporheic ecology of five small streams near Hamilton, New Zealand. *New Zealand Journal of Marine and Freshwater Research* **31(5)**: 609-622.
- Boulton AJ, Vallett HM, Fisher SG. 1992. Spatial distribution and taxonomic composition of the hyporheos of several Sonoran Desert streams. *Archiv für Hydrobiologie* **125(1)**: 37-61.
- Bretschko G. 1985. Quantitative sampling of the fauna of gravel streams. *Verhandlungen der Internationale Vereinigung für theoretische und angewandte Limnologie* **22**: 2049-2052.
- Cornut J, Elger A, Greugny A, Bonnet M, Chauvet E. 2012. Coarse particulate organic matter in the interstitial zone of three French headwater streams. *International Journal of Limnology* **48**: 303-313.
- Crossman J, Bradley C, Milner A, Pinay G. 2012. Influence of environmental instability of groundwater-fed streams on hyporheic fauna, on a glacial floodplain, Denali National Park, Alaska. *River Research and Applications* **29(5)**: 548-559.
- Danielopol DL. 1989. Groundwater fauna associated with riverine aquifers. *Journal of North American Benthological Society* **8(1)**: 18-35.
- Dole-Olivier M-J, Maazouzi C, Cellot B, Fiers F, Galassi DMP, Claret C, Martin D, Merigoux S, Marmonier P. 2014. Assessing invertebrate assemblages in the subsurface zone of stream sediments (0-15 cm deep)

- using a hyporheic sampler. *Water Resources Research* **50**: 453-465.
- Dole-Olivier M-J, Marmonier P, Befly J-L. 1997. Response of invertebrates to lotic disturbance: is the hyporheic zone a patchy refugium? *Freshwater Biology* **37**: 257-276.
- Elliott AH, Brooks NH. 1997. Transfer of nonsorbing solutes to a streambed with forms: Theory. *Water Resources Research* **33**:1: 123-136.
- Engelhardt I, Piepenbrink M, Trauth N, Stadler S, Kludt C, Schulz M, Schüth C, Ternes TA. 2011. Comparison of tracer methods to quantify hydrodynamic exchange within the hyporheic zone. *Journal of Hydrology* **400**: 255-266.
- Findlay S, Strayer D, Goumbala C, Gould K. 1993. Metabolism of streamwater dissolved organic carbon in the shallow hyporheic zone. *Limnology and Oceanography* **38**(7): 1493-1499.
- Franken RJM, Storey RG, Williams DD. 2001. Biological, chemical and physical characteristics of downwelling and upwelling zones in the hyporheic zone of a north-temperate stream. *Hydrobiologia* **444**: 183-195.
- Fraser BG, Williams DD. 1997. Accuracy and precision in sampling hyporheic fauna. *Canadian Journal of Fisheries and Aquatic Sciences* **54**: 1135-1141.
- Geist J. 2010. Strategies for the conservation of endangered freshwater pearl mussels (*Margaritifera margaritifera* L.): a synthesis of Conservation Genetics and Ecology. *Hydrobiologia* **644**: 69-88.
- Geist J, Auerswald K. 2007. Physiochemical stream bed characteristics and recruitment of the freshwater pearl mussel (*Margaritifera margaritifera*). *Freshwater Biology* **52**: 2299-2316.
- Geist J, Porkka M, Kuehn R. 2006. The status of host fish populations and fish species richness in European freshwater pearl mussel (*Margaritifera margaritifera*) streams. *Aquatic Conservation: Marine and Freshwater Ecosystems* **16**: 251-266.
- Geist DR, Hanrahan TP, Arntzen EV, McMichael GA, Murray CJ, Chien Y-J. 2002. Physicochemical characteristics of the hyporheic zone affect redd site selection by Chum Salmon and fall Chinook Salmon in the Columbia River. *North American Journal of Fisheries Management* **22**(4): 1077-1085.
- Gibert J, Dole-Olivier M-J, Marmonier P, Vervier P. 1990. Surface water-groundwater ecotones. In: Naiman RJ, Decamps H. (eds) *The Ecology and Management of Aquatic-Terrestrial Ecotones*. UNESCO/Parthenon Publishers: UK, 199-226.
- Grimaldi C, Chaplot V. 2000. Nitrate depletion during within-stream transport: effects of exchange processes between streamwater, the hyporheic and riparian zones. *Water, Air and Soil Pollution* **124**: 95-112.
- Harvey JW, Wagner BJ. 2000. Quantifying hydrologic interactions between streams and their subsurface hyporheic zones. In *Streams and Ground Waters*, Jones JB, Mulholland PJ (eds). Academic Press: San Diego; 3-344.
- Harvey JW, Wagner BJ, Bencala KE. 1996. Evaluating the reliability of the stream tracer approach to characterise stream-subsurface water exchange. *Water Resources Research* **32**(8): 2441-2451.
- Hlavacova E, Rulik M, Cap L. 2005. Anaerobic microbial metabolism in hyporheic sediment of a gravel bar in a small lowland stream. *River Research and Applications* **21**: 1003-1011.
- Hunt GW, Stanley EH. 2000. An evaluation of alternative procedures using the Bou-Rouch method for sampling hyporheic invertebrates. *Canadian Journal of Fisheries and Aquatic Sciences* **57**: 1545-1550.
- Ibrahim TG, Thornton SF, Wainwright J. 2010. Interplay of geomorphic and hydrogeologic features at reach- and channel unit-scales on riverbed hydrology and hydrochemistry: a conceptual model in the Lower Coal Measures, South Yorkshire, UK. *Hydrogeology Journal* **18**: 1391-1411.
- Jonsson K, Johansson H, Worman A. 2003. Hyporheic exchange of reactive and conservative solutes in streams- tracer methodology and model interpretation. *Journal of Hydrology* **278**: 153-171.
- Kasahara T, Wondzell SM. 2003. Geomorphic controls on hyporheic exchange flow in mountain streams. *Water Resources Research* **39**(1): SBH 3 1-14.
- Lewandowski J, Putschew A, Schwesig D, Neumann C, Radke M. 2011. Fate of organic

- micropollutants in the hyporheic zone of a eutrophic lowland stream: results of a preliminary field study. *Science of the Total Environment* **409**: 1824-1835.
- Malcolm IA, Middlemas CA, Soulsby C, Middlemas SJ, Youngson AF. 2010. Hyporheic zone processes in a canalised agricultural stream: implications for salmonid embryo survival. *Fundamental and Applied Limnology* **176(4)**: 319-336.
- Malcolm IA, Soulsby C, Youngson AF, Hannah DM, McLaren IS, Thorne A. 2004. Hydrological influences on hyporheic water quality: implications for salmon egg survival. *Hydrological Processes* **18**: 1543-1560.
- Marmonier P, Dole-Olivier M-J, Des Creuz Châtelliers MÉ. 1992. Spatial distribution of interstitial assemblages in the floodplain of the Rhone River. *Regulated Rivers: Research and Management* **7(1)**: 75-82.
- McCallum JL, Cook PG, Berhane D, Rumpf C, McMahon GA. 2012. Quantifying groundwater flows to streams using differential flow gaugings and water chemistry. *Journal of Hydrology* **416-417**: 118-132.
- Nelson SM, Roline RA. 2003. Effects of multiple stressors on hyporheic invertebrates in a lotic system. *Ecological Indicators* **3**: 65-79.
- Nogaro G, Datry T, Mermillod-Blondin F, Descloux S, Montuelle B. 2010. Influence of streambed sediment clogging on microbial processes in the hyporheic zone. *Freshwater Biology* **55(6)**: 1288-1302.
- Olsen DA, Townsend CR. 2005. Flood effects on invertebrates, sediments and particulate organic matter in the hyporheic zone of a gravel-bed stream. *Freshwater Biology* **50(5)**: 839-853.
- Orghidan T. 1959. Ein neuer Lebensraum des unterirdischen Wassers, der hyporheische Biotop. *Archiv für Hydrobiologie* **55**: 392-414.
- Palmer MA. 1993. Experimentation in the hyporheic zone: challenges and prospects. *Journal of the North American Benthological Society* **12(1)**: 84-93.
- Pinay G, O'Keefe TC, Edwards RT, Naiman RJ. 2009. Nitrate removal in the hyporheic zone of a salmon river in Alaska. *River Research and Applications* **25**: 367-375.
- Pugsley CW, Hynes HBN. 1983. A modified freeze-core technique to quantify the depth distribution of fauna in stone streambeds. *Canadian Journal of Fisheries and Aquatic Sciences* **43(9)**: 1812-1817.
- Reid N, Keys A, Preston JS, Moorkens E, Roberts D, Wilson CD. 2012. Conservation status and reproduction of the critically endangered freshwater pearl mussel (*Margaritifera margaritifera*) in Northern Ireland. *Aquatic Conservation: Marine and Freshwater Ecosystems* **23(4)**: 571-581.
- Riss HW, Meyer EL, Niepagenkemper O. 2008. A novel and robust device for repeated small-scale oxygen measurement in riverine sediments- implications for advanced environmental surveys. *Limnology and Oceanography: Methods* **6**: 200-207.
- Rivett MO, Ellis PA, Greswell RB, Ward RS, Roche RS, Cleverly MG, Walker C, Conran D, Fitzgerald PJ, Willcox T, Dowle J. 2008. Cost-effective mini drive-point piezometers and multilevel samplers for monitoring the hyporheic zone. *Engineering Geology and Hydrogeology* **41**: 49-60.
- Scarsbrook MR, Halliday J. 2002. Detecting patterns in hyporheic community structure: does sampling method alter the story? *New Zealand Journal of Marine and Freshwater Research* **36(2)**: 443-453.
- Schwoerbel J. 1961. Über die lebensbedingungen und die Besiedlung des hyporheischen Lebensraumes. *Archiv für Hydrobiologie* **25**: 182-214.
- Skinner A, Young M, Hastie L. 2003. *Ecology of the Freshwater Pearl Mussel*. Conserving Natura 2000 Rivers Ecology Series 2. English Nature, Peterborough.
- Soulsby C, Malcolm IA, Youngson AF. 2001. Hydrochemistry of the hyporheic zone in salmon spawning gravels: a preliminary assessment in a degraded agricultural stream. *Regulated Rivers: Research and Management* **17**: 651-665.
- Stanford JA, Ward JV. 1988. The hyporheic habitat of river ecosystems. *Nature* **335**: 64-66.
- Stocker ZSJ, Williams DD. 1972. A freezing core method for describing the vertical distribution of sediments in a streambed. *Limnology and Oceanography* **17**: 136-138.

- Storey RG, Williams DD. 2004. Spatial responses of hyporheic invertebrates to seasonal changes in environmental parameters. *Freshwater Biology* **49(11)**: 1468-1486.
- Stubbington R. 2012. The hyporheic zone as an invertebrate refuge: a review of variability in space, time, taxa and behaviour. *Marine and Freshwater Research* **63(4)**: 293-311.
- Stubbington R, Wood PJ, Reid I. 2011. Spatial variability in the hyporheic zone refugium of temporary streams. *Aquatic Sciences* **73**: 499-511.
- Tanaka A, Namba T, Tanida K, Takemon Y. 2014. Evaluation of a pump method for unbiased sampling of stream hyporheos. *Hydrobiologia* **730**: 29-43.
- Toran L, Hughes B, Nyquist J, Ryan R. 2013. Freeze core sampling to validate time-lapse resistivity monitoring of the hyporheic zone. *Groundwater* **51(4)**: 635-640.
- Triska FJ, Duff JH, Avanzino J. 1993. The role of water exchange between a stream channel and its hyporheic zone in nitrogen cycling at the terrestrial-aquatic interface. *Hydrobiologia* **251**: 167-184.
- Triska FJ, Kennedy VC, Avanzino RJ. 1989. Retention and transport of nutrients in a third-order stream in northwestern California: hyporheic processes. *Ecology* **70**: 1893-1905.
- Van Stempvoort DR, Roy JW, Brown SJ, Bickerton G. 2011. Artificial sweeteners as potential tracers in groundwater in urban environments. *Journal of Hydrology* **401**: 126-133.
- Vervier P, Gibert J, Marmonier P, Dole-Olivier M-J. 1992. A perspective on the permeability of the surface freshwater-groundwater ecotone. *Journal of the North American Benthological Society* **11**: 93-102.
- White DS. 1993. Perspectives on defining and delineating hyporheic zones. *Journal of North American Benthological Society* **12**: 61-69.
- Williams DD. 1989. Towards a biological and chemical definition of the hyporheic zone in two Canadian rivers. *Freshwater Biology* **22**: 189-208.
- Williams DD, Hynes HBN. 1974. The occurrence of benthos deep in the substratum of a stream. *Freshwater Biology* **4**: 233-256.
- Williams DD, Febria CM, Wong JCY. 2010. Ecotonal and other properties of the hyporheic zone. *Fundamental and Applied Limnology* **176(4)**: 349-364.
- Wondzell SM. 2006. Effect of morphology and discharge on hyporheic exchange flows in two small streams in the Cascade Mountains of Oregon, USA. *Hydrological Processes* **20**: 267-287.
- Wood PJ, Boulton AJ, Little S, Stubbington R. 2010. Is the hyporheic zone a refugium for aquatic macroinvertebrates during severe low flow conditions? *Fundamental and Applied Limnology* **176(4)**: 377-390.
- Youngson AF, Malcolm IA, Thorley JL, Bacon PJ, Soulsby C. 2004. Long-residence groundwater effects on incubating salmonid eggs: low hyporheic oxygen impairs embryo development. *Canadian Journal of Fisheries and Aquatic Sciences* **61(12)**: 2278-2287.

4.1.1. Coring Methods

Craig Frew¹

¹ School of Geosciences, University of Aberdeen (CraigRFrew@gmail.com)



ABSTRACT: Coring operations are an essential element in studies seeking to complement surface landform evidence with a more continuous record of landscape change. Site assessment and the choice of coring location within a basin is an important first step in any coring task. A set of established tools and techniques are available for coring operations that seek to minimise the disturbance to material during extraction and maximise the potential for long and continuous records. The choice of tools will depend largely on the coring environment and remoteness of the site. Importantly, there is no universal tool available for any coring task and, commonly, multiple pieces of equipment are required to extract a continuous record. This contribution seeks to outline the process of planning a coring operation, detail the advantages/disadvantages of the various piston, gouge and Russian style corers available and highlight the various conventions used for the reporting of results.

KEYWORDS: coring, sediment Sampling, palaeo-environmental, reconstruction

Introduction

The discipline of geomorphology is largely concerned with the understanding of earth surface process and the origin/interpretation of landform assemblages. Despite this, it is often the case that a more continuous record of landscape development is required. The accumulation of sediments and organic materials over time forms multiproxy archives of palaeoecological, sedimentological and/or hydrological change within a catchment or landscape. Additionally, detailed investigation of subsurface stratigraphic horizons can be used to understand the depositional settings and interrelationships of surface landforms. In the absence of exposed sedimentary cross sections, for example in river terraces or during construction works, suitable material must be retrieved by means of coring. A typical coring task involves careful planning, with the diverse range of environments suitable for coring activities each presenting a unique challenge in terms of operation and equipment selection. For example, coring in more remote locations is often limited in terms of achievable depth by the amount of equipment required and the feasibility of effectively transporting samples from the site,

as is the case in many arctic or high altitude settings.

This paper aims to highlight the various challenges, and potential pitfalls of any coring task. The key stages of planning, equipment operation and transport of material will be addressed. The focus for this paper will be on coring in terrestrial and lacustrine environments. For marine, deep sea and ice coring, the reader is referred elsewhere.

Site Assessment

Geomorphological Assessment

In many studies, an important first step is to determine the potential sources of material accumulating within the depositional setting e.g. fluvial systems, talus cones, dirty snow avalanches, debris flows etc. Therefore, geomorphological mapping (See Otto and Smith 2014) of the immediate catchment is an important first task in any coring activity. To avoid disturbed records, sites close to steep slopes or obvious mass movements, avalanche tracks and across floodplains where evidence of meandering has taken place in the past should be avoided. Following this, the preservation potential of

sediments can be estimated at sites within the chosen accumulation basin.

Another consideration for site selection is to ensure that access is granted from the relevant land owner and permission sought from the relevant agency if the site is designated as protected. It is vital to ensure that there is no risk of damaging any subsurface utilities (gas, electricity etc.). This can be done by consulting the relevant up-to-date local utilities map.

Lacustrine Sediment Preservation

Preservation in lacustrine environments is governed by both biotic and abiotic processes. For reconstructing palaeoenvironments from organic sediments, it is important to consider changes to the primary signal through degradation of material in the water column as well as post depositional diagenesis. In the water column, organic material is subject to degradation through oxidisation. Lehmann *et al.* (2002) show how the carbon and nitrogen isotopic composition of lake sediments, typically measured as palaeo-climate indicator, varies with oxygenation level whereby preservation is enhanced under anoxic conditions. Similarly, lipid biomarkers, such as alkenones often used to infer a wide range of palaeoclimatic conditions, can be affected by prolonged oxygen exposure and microbial degradation (Rontani *et al.* 2013). As such, investigations using such methods can benefit by targeting coring in certain depositional environments. For example, shallower settings with shorter residence times have greater potential for the preservation of organic deposits (Meyers and Ishiwatari 1993). As discussed below however, such shallow settings can be prone to sediment re-suspension, encouraging post-depositional oxidisation of material and further degradation. On the lake bed, sediments can be subject to bioturbation which involves the disturbance through the movement and burrowing of benthic organisms. Typically, this occurs within the top 10cm of deposits and can result in a temporal smoothing of the sequence as stratigraphic layers are overturned. Benthic organisms require dissolved oxygen to survive, meaning that anoxic conditions, largely found in eutrophic environments, significantly reduce the risk of bioturbation in sediments (Meyers and Ishiwatari 1993).

In terms of abiotic factors, wave action and wind driven turbulence within the water column play a key role in the redistribution of fine grained sediments on the lake floor. The significance of such processes depends on both wind speed and fetch (Meyers and Ishiwatari 1993). Basins protected by surrounding topography and those that are deep with small surface areas offer the greatest protection from wave action. Re-suspension also depends on the critical shear stress of the sediment bed in question, with recently deposited material being re-suspended more easily than that which is more compacted. Such processes have been observed to produce highly variable sediment distribution in lakes. During storms, wind driven turbulence can be responsible for eroding near shore material and transporting it towards deeper parts of the lake (Bengtsson *et al.* 1990). Careful consideration of wind driven redistribution is therefore necessary on exposed lakes where sedimentary chronologies may be disrupted, even in accumulation zones. Colluvial processes such as slumping and turbidity currents also have the potential to disrupt lacustrine sedimentary sequences through the erosion and re-deposition of thick sedimentary units. Turbidity flows are benthic currents induced by dense inflows laden with high suspended sediment loads (Meiburg and Kneller 2010). The density difference between inflow and ambient lake waters can cause the inflow to continue as an erosive current along the lake bed. The current may subsequently separate from the lake bed at depth if the density of the ambient water becomes greater than that of the current. Shallower settings with steep sided delta faces are more prone to erosion through turbidity currents and are more common in settings that are fed by fluvial networks with a high suspended sediment load, such as proglacial environments.

The difficulties outlined above can be overcome in some cases by sampling at multiple locations to bridge erosional gaps in the sequence and gain a fuller understanding of sediment distribution throughout the basin. The number of coring locations targeted will be dependent on the logistical constraints of the site. However a single core approach is rarely sufficient to identify a representative stratigraphy. For larger lakes, sub-basins should be considered as part of a sampling

strategy that targets wide, flat and deep areas that are distant to the lakeshore. Other processes such as the occurrence of dropstones or thick volcanic ash layers have the potential to introduce errors in the age depth relationship of a sequence, although these processes tend to occur more widely throughout the basin and are largely unavoidable. Ground penetrating radar (GPR; see Robinson *et al.* 2013) or seismic investigations make it possible to map the distribution of sediment within a basin, helping to determine the most suitable coring site. Analysis using radioisotopic profiles from short cores can provide an opportunity to assess the disturbance of recent sediments before a more comprehensive coring operation takes place, however this is not always practical due to time constraints.

Preservation of Bog and Wetland deposits

Paus (2013) identify soil erosion as a major potential source of error in palaeoenvironmental investigations. In bog or wetland settings, soil erosion can be a consequence of deforestation, agriculture and peat cutting. The erosion, transportation and deposition of old sediments may introduce errors in the age depth relationship and as well as foreign or reworked proxy material. Soil erosion is substantially increased in areas with human presence, and as such, the study of sediments deposited since the mid-Holocene can be particularly susceptible.

Borehole transects using basic gouge auger tools (outlined in the next section) are often conducted to develop a three dimensional understanding of the subsurface stratigraphy prior to recovery of sequences for detailed analysis. This step is crucial for determining the most complete and representative sequence of deposits for analysis.

Equipment Selection

Equipment selection is determined in large part by the coring environment (bog, floodplain, lake etc.), the nature of the material being extruded and site accessibility. A wide range of equipment available, each offering advantages in terms of transportability, operating depth and sample length (Table 1). Typical requirements for coring in remote locations include reliability, modular designs and air/sea freight capability (Kelts *et al.* 1986). The advantages and disadvantages of a range of equipment will now be described. For more detailed instructions regarding the operation of the equipment described here, the reader is referred to the references provided.

Gouge Auger

Gouge augers are easily transportable tools that permit relatively quick survey of subsurface sediments in terrestrial environments. Sampling is particularly rudimentary and involves thrusting a semi cylindrical chamber into deposits and twisting the device using a handle at the surface to capture the sample. Consecutive drives are enabled by the addition of extension rods.

Table 1: Summary of equipment available for coring tasks

Device	Number of operators	Max depth (m)	Section length (m)	Reference(s)
Gouge Auger & 'Russian' Corer	2	10-15	0.5, 1	Franzén and Ljung (2009)
Rod operated Piston Corer	3	<30	1-5	Livingstone (1955); Kelts <i>et al.</i> (1986); Mingram <i>et al.</i> (2007)
Wire operated Piston Corer	3	>100	1-6	Nesje (1992); Chambers and Cameron (2001)
Mackereth	2	>100	2, 3, 6	Mackereth (1958); Barton and Burden (1979)
Surface samplers	1	>100	~1	Wright (1980); Blomqvist (1985); Glew <i>et al.</i> (2001); Larsen <i>et al.</i> (2011)

The retrieved sample is subject to significant disturbance as the open chamber is prone to resampling of material from depths above those required especially where sands underlie the softer organic material above. Additionally, more-consolidated material can force its way upwards over less consolidated horizons within the chamber. For these reasons, it is not recommended that the gouge auger is used to retrieve samples for analysis. However, such equipment is suitable for developing a first order approximation of the three dimensional stratigraphy of a sedimentary basin, bog or mire by conducting multiple boreholes in transect. This can be done rapidly over a wide surface area with a minimum of two operators.

'Russian' Peat Corer

The Gouge auger is often used in tandem with a 'Russian' peat corer which is capable of retrieving largely undisturbed samples from organic material or fine-medium grained clastic sediment. The corer consists of a half barrel which pivots around a flat 'fin' plate

(Franzén and Ljung 2009). The chamber is inserted in an open position, relative to the 'fin' plate (Figure 1a), and subsequently closed at the required depth to capture the sample (Figure 1b). Care must be taken to ensure that the chamber is fully closed to avoid contamination.

Russian corers can be used in both unsaturated and semi-saturated environments, however overly saturated material is difficult to sample as the core is not retained in its own sample tube and instead must be transferred from the fin plate to a suitable container for transportation. Russian corers have been extensively used in terrestrial studies in peat bog, marshland and coastal environments (e.g. Dawson *et al.* 2004; Jordan *et al.* 2010; Long *et al.* 2011; Schofield and Edwards 2011; Long *et al.* 2012). The base 15-20cm of sedimentary basins can be difficult to sample due to the 'nosecone' of the Russian corer. Also, the system can be heavy making transport to remote sites difficult.



Figure 1: Russian Peat Corer. (a) The corer in an 'open' position with detachable extension rods and sections of plastic guttering for transporting samples, (b) A sample collected with the chamber now in a 'closed' position. Photographs courtesy of S. Dawson

Piston Corer

'Piston' type corers are specifically designed for the extraction of saturated sediments without significant disturbance to the sample. Such corers are largely based on the Kullenberg design (Emery and Broussard 1956) used in marine coring. This involves

driving a metal or plastic sample tube into sediment over a fixed-position 'piston' which acts to maintain a vacuum below it, reducing deformation and holding the sample in place as it is extracted (Figure 2). Several modifications to the Kullenberg design have

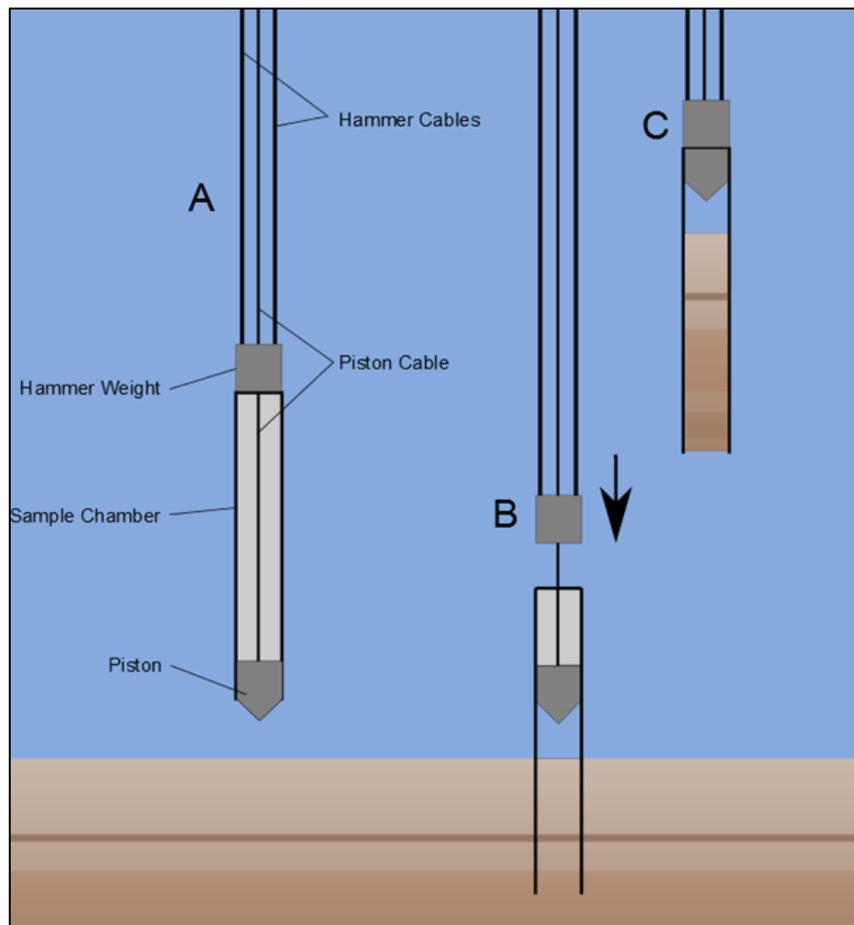


Figure 2: Wireline piston corer operation. A) The device is submerged and lowered to a short distance above the sediment water interface. B) The piston wire is secured at the surface before the sample chamber is hammered into the sediment below the fixed position piston. C) The system is raised to the surface with the sample retained in a vacuum below the piston.

been proposed over the years. Originally Livingstone (1955) adapted the marine corer for use in shallow lacustrine environments, using extension rods to manually drive the coring chamber below the fixed-piston.

Despite the success of this device, the use of extension rods increases the total weight of the system, making coring in remote locations difficult and limiting the reasonable operating depth to 20-30m. To overcome this, several wireline systems have been developed which use a weighted hammer that can be raised and dropped to drive the coring chamber into the sediment (e.g. Kelts *et al.* 1986; Fisher *et al.* 1992; Nesje *et al.* 2008) Such wireline systems are capable of retrieving samples in excess of 100m water depth, however this is generally limited to a single coring drive due to difficulties associated with penetrating the same bore hole at great depth multiple times. The length of sample retrieved using a single drive can

be increased by lengthening the coring chamber, however this increases flexibility of the chamber, especially of those manufactured from PVC, meaning that much of the driving force is lost through flexing (Nesje 1992). The typical maximum coring drive for such systems is approximately 6m.

Piston coring systems have typically been used in lacustrine environments where they are capable of retrieving samples that maintain the integrity of fragile sedimentary structures, such as varves (e.g. (Moore *et al.* 2001; Snowball and Sandgren 2002), as well as penetrating through thick sequences of glacio-fluvial diamicton and clay. Additionally, they have been used in terrestrial settings where the addition of a sharp cutting mouth at the end of the core barrel helps to penetrate through layers of wood and fibrous peat (Wright Jr *et al.* 1984; Mingram *et al.* 2007). Difficulties using piston corers typically involve the upward

displacement of the piston during retrieval, causing the sample to be drawn upwards inside the core barrel resulting in sediment disturbance. Disturbed or compressed sequences are evident based on comparison between the length of sample retrieved and the recorded depth of penetration during the coring task. Several recovery systems that minimise the potential disturbance to the sample have been proposed (Chambers and Cameron 2001; Mingram *et al.* 2007). In deep-water environments, ridged casing can be used to prevent excessive bending of extension rods and to guide the corer to the borehole location for consecutive drives, however such casing is heavy, expensive and of limited use in remote locations (Livingstone 1955; Mingram *et al.* 2007).

A regularly cited advantage of piston corers is that they are capable of retrieving the sediment-water interface intact. However in most cases the excessive force required to sample longer sequences disturbs the uppermost sections of the sample. As such, piston corers are typically used alongside surface samplers, dredges and box corers that have proven to be more effective tools for this purpose.

Mackereth Corer

Mackereth type corers use a Kullenberg style fixed-piston in a similar manner to the generic piston corers discussed above (Figure 2). However, the method of driving the coring chamber below the fixed-piston is fundamentally different. The Mackereth corer relies on a large cylindrical anchor drum to stabilise the system on the lake bed. The anchor drum is coupled to a main coring barrel, inside which, a retracted coring chamber rests (Figure 3). Compressed air is fed by the operator at the surface through high pressure tubes into the main barrel above the coring chamber. This increases pressure above the coring chamber, forcing the coring chamber below the fixed piston and into the sediment. Once the chamber is fully extended, the compressed air is then redirected into the anchor drum, creating buoyancy that lifts the entire system towards the surface. A key advantage of the Mackereth corer is that it can be used with relative ease in deep-water, limited largely by the hydrostatic pressure at the lake bed that must be overcome to return the equipment to the surface.

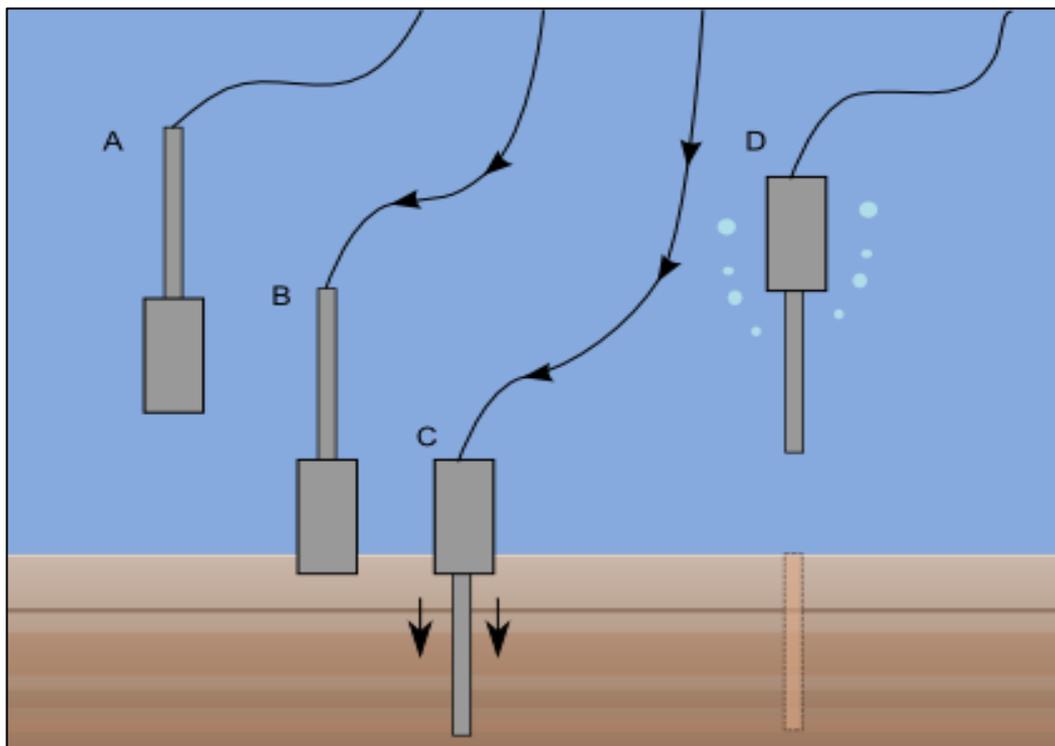


Figure 3: Mackereth corer operation. A) The device is submerged and falls under its own weight to the sediment water interface B) The system anchors itself into place as the anchor drum penetrates the sediment. C) Compressed air used to drive the coring chamber below a fixed position piston. D) Once the chamber has fully extended, air is diverted into the anchor drum, creating buoyancy that lifts the system to the surface.

Notable disadvantages are that the device is not always capable of capturing full sequences. This occurs when the frictional force of sediment on the coring barrel exceeds the force required to displace the anchor drum from the bed. This results in the full extension of the coring chamber from the main barrel, returning the system to the surface. The equipment is generally heavy, requiring multiple containers of compressed air and is therefore not suitable for extremely remote locations. However smaller, more portable versions are available.

Surface Samplers

For sampling of the sediment-water interface, several tools are available. Nearly all surface samplers rely on gravity for penetration into sediment. Sampling in this way can be susceptible to fine grained sediment washout caused by the influence of bow waves as the device approaches the sediment water interface. For this reason, it is recommended that gravity driven surface samplers are lowered to a short distance from the sediment water interface before being allowed to fall under their own weight. However, some tools include diode type features that allow for water to pass through during descent with rubber flaps that act to maintain the sediment sample during ascent.

Grab samplers typically consist of a set of jaws that can be closed by triggering a release mechanism with a messenger weight to capture sediment at the surface. Similarly, Box corers are open sampling boxes that cut into the sediment and are closed from below as the sample is lifted (De Groot *et al.* 1982). Such tools are principally used for collecting large volumes of sediment over a wide surface area. Box corers have the advantage of retrieving samples in a relatively undisturbed state. However in particularly consolidated sediment, the locking mechanism on box corers may fail to close fully resulting in the loss of sample. Grab samplers are more suited to this purpose, with the downside however of significant sample disturbance. Due to the heavy combined weight of sample and device, large volume grab samplers require a substantial raft platform with suitable counter-balance for operation. Such samplers are only suitable for a general assessment of the sediments and do not preserve lithological boundaries.

Freeze coring offers an opportunity to preserve delicate sedimentary structures at the sediment water interface (Kulbe and Niederreiter 2003). Such corers commonly use a mixture of dry ice and ethanol that is circulated around the inside face of a stainless steel wedge to freeze the sediments *in-situ*. Freeze coring devices require a power source and a supply of cooling agents in the field as well as a cooling box for the transport of the frozen samples from the field to the laboratory. As such they are not suitable for remote sites.

Other surface sediment corers include designs that maintain the sample in a coring chamber through the messenger triggered deployment of a suction cup at the head of the device. This creates a vacuum, capable of holding fine grained, consolidated sediments. The waterlogged sediments can then be subsampled for transport using an extruder device (Figure 4). Such systems suffer similar problems in terms of the bow wave induced washout of surface fines, however they tend to be lightweight, modular and therefore more suitable for remote field sites.



Figure 4: Subsampling of a surface core at Hämelsee, Germany, using an extruder device. Photography courtesy of W. van der Bilt.

Additional Equipment

In addition to the tools described above, it is advisable to be equipped with a series of cutting tools, tape measures, waterproof writing utensils and safety equipment. Sharp knives or cheese wire can be used to separate the core from the sampling plate when using the Russian Corer. A hacksaw can be used to cut lengths of PVC core liner pipe for the piston style corers. Waterproof gloves should be used to protect the hands from sharp objects and from getting cold. A back brace may also be necessary for heavy lifting.

When coring in lacustrine or shallow marine environments a raft or boat is often necessary, for which several designs are available. Many raft designs are also modular, meaning that they can be transported with relative ease to remote locations and assembled on site. Some rafts make use of foam blocks for buoyancy (e.g. Figure 5), while others are simply metal, wooden or plastic platforms strapped to a Zodiac-style inflatable boat. It is important to ensure that the chosen raft has sufficient buoyancy for the extraction of deep sediment cores, as the core can only be extruded when the force of the raft buoyancy exceeds that of the friction and vacuum-suction acting on the coring chamber. Many rafts are mounted with a tripod and winch system which can make the extraction of heavy samples easier.



Figure 5: Coring from an isopore block raft platform, SE Greenland.

Critically, there is no universally applicable tool that is suitable for any environment or coring task. As described above, the nature of the sedimentary material and accessibility

of the site play a large part in the selection of equipment. More often than not, a combination of equipment is required to retrieve a full, undisturbed sequence.

Coring Procedure

The coring procedure should follow a systematic approach. In lacustrine settings, the water depth should be measured using an echo-sounding device or weighted rope measure. In settings with a particularly diffuse sediment water interface, the true water depth can be difficult to judge. When coring, a dual or triplicate core approach should be used to cover boundaries between core sections. Care should be taken to accurately record the penetration depth which should be compared against the depth of material sampled to identify any compression/ loss of material. The location of boreholes should be measured with GPS and levelled according to a known datum if necessary.

It is good practice to photograph and record basic characteristics of key sedimentary horizons in the field while the sediment is fresh. This way any major disturbances during transport can be identified. There are several classification schemes for doing this, (e.g. Troels-Smith 1955; Barnston and Livezey 1987; Schnurrenberger *et al.* 2003). Often however, cores are maintained in their PVC casings until they are opened in the laboratory. Finally, it is important to know if the samples retrieved represent a full record of deposition at the site. Therefore it is important to record, and report in any publication, whether or not the base of the depositional basin is believed to have been reached.

For transport, cores collected in full PVC liners should be plugged at either end using 'oasis' floral foam or 'sodium polyacrylate' gel which will fill any excess water or air space in the core tube. The cores should then be sealed using a tight fitting cap and secured with waterproof tape. Importantly, core samples intended for further analysis, such as luminescence dating, should be collected in opaque core liners and not opened until in a light secure environment. Lacustrine sediments should be maintained in an upright position for as long as possible during transportation to minimise sediment disturbance (Figure 6). It is possible to freeze

cores for transportation, however ice crystal growth can potentially disrupt fragile sedimentary features. For cores collected using a 'Russian' peat corer, the sample can be transferred into half sections of PVC guttering and secured with appropriate packing material (paper or plastic wrap if the samples are to be radiocarbon dated) to soak up excess water before being wrapped in plastic wrap for transportation.



Figure 6: Storage of cores in the field can be difficult. Here, cores are strapped to a boulder in an upright position in order to allow the sample to settle.

Conclusion

In conclusion, sediment coring represents an essential and useful technique to complement surface geomorphological investigations. A wide variety of tools have been developed to allow coring in increasingly diverse and remote locations. The choice of equipment will be governed largely by the nature of sediments to be cored and the accessibility of the site. The methods and tools available for coring will continue to evolve as we seek to develop longer and better preserved sedimentary records.

References

- Barnston AG, Livezey RE. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, **115**: 1083-1126.
- Barton C, Burden F. 1979. Modifications to the Mackereth corer. *Limnology and Oceanography*, **24**: 977-983.
- Bengtsson L, Hellström T, Rakoczi L. 1990. Redistribution of sediments in three Swedish lakes. *Hydrobiologia*, **192**: 167-181.
- Blomqvist S. 1985. Reliability of core sampling of soft bottom sediment—an in situ study. *Sedimentology*, **32**: 605-612.
- Chambers J, Cameron N. 2001. A rod-less piston corer for lake sediments; an improved, rope-operated percussion corer. *Journal of Paleolimnology*, **25**: 117-122.
- Dawson S, Smith DE, Jordan J, Dawson AG. 2004. Late Holocene coastal sand movements in the Outer Hebrides, N.W. Scotland. *Marine Geology*, **210**: 281-306.
- De Groot A, Zschuppel K, Salomons W. 1982. Standardization of methods of analysis for heavy metals in sediments. *Hydrobiologia*, **91**: 689-695.
- Emery GR, Broussard DE. 1956. A modified Kullenberg piston corer. *Journal of Sedimentary Research* **24**: 3.
- Fisher MM, Brenner M, Reddy KR. 1992. A simple, inexpensive piston corer for collecting undisturbed sediment/water interface profiles. *Journal of Paleolimnology*, **7**: 157-161.
- Franzén L, Ljung T. 2009. A carbon fibre composite (CFC) Byelorussian peat corer. *Mires and Peat*, **5**: 1-9.
- Glew JR, Smol JP, Last WM. 2001. Sediment core collection and extrusion. In: Last WM, Smol JPS (Eds.) *Tracking Environmental Change Using Lake Sediments. Volume 1: Basin Analysis, Coring and Chronological Techniques*. Kluwer Academic Publishers: The Netherlands: 73-105.
- Jordan JT, Smith DE, Dawson S, Dawson AG. 2010. Holocene relative sea-level changes in Harris, Outer Hebrides, Scotland, UK. *Journal of Quaternary Science*, **25**: 115-134.
- Kelts K, Briegel U, Ghilardi K, Hsu K. 1986. The limnogeology - ETH coring system. *Swiss Journal of Hydrology*, **48**: 104-115.
- Kulbe T, Niederreiter R. 2003. Freeze coring of soft surface sediments at a water depth of several hundred meters. *Journal of Paleolimnology*, **29**: 257-263.
- Larsen NK, Kjær KH, Olsen J, Funder S, Kjeldsen KK, Nørgaard-Pedersen N. 2011. Restricted impact of Holocene climate variations on the southern Greenland Ice

- Sheet. *Quaternary Science Reviews*, **30**: 3171-3180.
- Lehmann MF, Bernasconi SM, Barbieri A, McKenzie JA. 2002. Preservation of organic matter and alteration of its carbon and nitrogen isotope composition during simulated and in situ early sedimentary diagenesis. *Geochimica et Cosmochimica Acta*, **66**: 3573-3584.
- Livingstone D. 1955. A lightweight piston sampler for lake deposits. *Ecology*, **36**: 137-139.
- Long AJ, Woodroffe SA, Milne GA, Bryant CL, Simpson MJR, Wake LM. 2012. Relative sea-level change in Greenland during the last 700yrs and ice sheet response to the Little Ice Age. *Earth and Planetary Science Letters*, **315-316**: 76-85.
- Long AJ, Woodroffe SA, Roberts DH, Dawson S. 2011. Isolation basins, sea-level changes and the Holocene history of the Greenland Ice Sheet. *Quaternary Science Reviews*, **30**: 3748-3768.
- Mackereth F. 1958. A portable core sampler for lake deposits. *Limnology and Oceanography*, **3**: 181-191.
- Meiburg E, Kneller B. 2010. Turbidity currents and their deposits. *Annual Review of Fluid Mechanics*, **42**: 135-156.
- Meyers PA, Ishiwatari R. 1993. Lacustrine organic geochemistry—an overview of indicators of organic matter sources and diagenesis in lake sediments. *Organic Geochemistry*, **20**: 867-900.
- Mingram J, Negendank JF, Brauer A, Berger D, Hendrich A, Köhler M, Usinger H. 2007. Long cores from small lakes—recovering up to 100 m-long lake sediment sequences with a high-precision rod-operated piston corer (Usinger-corer). *Journal of Paleolimnology*, **37**: 517-528.
- Moore J, Hughen K, Miller G, Overpeck J. 2001. Little Ice Age recorded in summer temperature reconstruction from varved sediments of Donard Lake, Baffin Island, Canada. *Journal of Paleolimnology*, **25**: 503-517.
- Nesje A. 1992. A piston corer for lacustrine and marine sediments. *Arctic and Alpine Research*, **24(3)**: 257-259.
- Nesje A, Dahl SO, Thun T, Nordli Ø. 2008. The 'Little Ice Age' glacial expansion in western Scandinavia: Summer temperature or winter precipitation? *Climate Dynamics*, **30**: 789-801.
- Otto JC, Smith MJ. 2014. Section 2.6: Geomorphological Mapping. In: Clarke LE, Nield JM (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology, London. ISSN: 2047-0371.
- Paus A. 2013. Human impact, soil erosion and vegetation response lags to climate change: challenges for the mid-Scandinavian pollen-based transfer-function temperature reconstructions. *Vegetation History and Archaeobotany*, **22(3)**: 269-284.
- Robinson M, Bristow C, McKinley J, Ruffell A. 2013. Section 1.5.5: GPR. In: Clarke LE, Nield JM (Eds.) *Geomorphological Techniques (Online Edition)*. British Society for Geomorphology, London. ISSN: 2047-0371.
- Rontani J, Volkman JK, Prah FG, Wakeham SG. 2013. Biotic and abiotic degradation of alkenones and implications for paleoproxy applications: A review. *Organic Geochemistry*, **59**: 95-113.
- Schnurrenberger D, Russell J, Kelts K. 2003. Classification of lacustrine sediments based on sedimentary components. *Journal of Paleolimnology*, **29**: 141-154.
- Schofield JE, Edwards KJ. 2011. Grazing impacts and woodland management in Eriksfjord: *Betula*, coprophilous fungi and the Norse settlement of Greenland. *Vegetation History and Archaeobotany*, **20**: 181-197.
- Snowball I, Sandgrena P. 2002. Geomagnetic field variations in northern Sweden during the Holocene quantified from varved lake sediments and their implications for cosmogenic nuclide production rates. *The Holocene*, **12**: 517-530.
- Troels-Smith J. 1955. *Karakterisering af løse jordarter. Characterization of unconsolidated sediments*. Geological Survey of Denmark Series Volume 3 (10).
- Wright Jr H, Mann D, Glaser P. 1984. Piston corers for peat and lake sediments. *Ecology*, **65**: 657-659.
- Wright HE. 1980. Cores of soft lake sediments. *Boreas*, **9**: 107-114.

Chironomid analysis: background, methods and geomorphological applications

Naomi Holmes¹

¹Department of Geographical and Environmental Sciences, School of Science and Technology, University of Northampton, St. George's Avenue, Northampton, NN2 6JD (naomi.holmes@northampton.ac.uk)



ABSTRACT: Subfossil chironomids can provide both qualitative and quantitative reconstructions of past climatic and environmental conditions. Chironomid data are frequently used in palaeoecological and palaeoclimate studies, however geomorphological studies using chironomid data are currently few. Chironomid records are produced from lacustrine and alluvial sediments; these terrestrially based reconstructions can be of benefit to geomorphologists. In order to increase geomorphologists' awareness of the potential of chironomid data this paper discusses the technique, with a particular focus on geomorphological applications. Chironomids and chironomid analysis are briefly introduced, and the methods used, both in the field and the laboratory, are outlined. Geomorphological studies which have utilised existing chironomid data, including glacial geomorphological, sea level and palaeohydrological studies, are discussed, as well as a chironomid study which produced data specifically aimed at investigating changing sea level. It is hoped that further applications for chironomid data in geomorphological studies will emerge.

KEYWORDS: chironomid – palaeolimnology - sea level - glacial geomorphology - floodplain – palaeohydrology

Introduction

Chironomids (Insecta; Diptera; Chironomidae), also known as non-biting midges, are a family of two-winged flies with an estimated 15000 species globally (Cranston, 1995). Chironomids are often the most abundant insect group found in freshwaters (Pinder, 1983; Cranston, 1995) and are the most ubiquitous of all aquatic species (Oliver and Roussel, 1983). Chironomids are holometabolous insects, with four life-stages (egg, larva, pupa, adult). It is the larval stage which is of interest to palaeoecologists. There are four instars of larval chironomids; each larval stage undergoes ecdysis (Oliver, 1971). The early instars do not preserve very well, but while the soft parts of the third and fourth instars tend to decay, the chitinous head capsule will preserve, recording the presence of this chironomid. The chitinous head capsules of the larvae are often well preserved and

abundant in lake sediments (Walker, 2001). It is possible to obtain sediment cores from lacustrine environments and to extract and identify the chironomid head capsules preserved within them, to provide information about which chironomids used to live in an environment.

Due to the fact that chironomid species have different environmental preferences and tolerances it is possible to reconstruct past environmental conditions from the head capsule data. Key to these reconstructions is an understanding of modern ecological processes (Eggermont and Heiri, 2012; Velle *et al.*, 2012; Juggins, 2013), and a number of studies have been carried out to better understand the relationships between chironomid species and their environments (e.g., Marziali and Rossaro, 2013; Rae, 2013).

If these relationships are found to be significant statistically it is possible to produce a model which allows quantitative reconstruction of past conditions (e.g., Birks, 1995; 1998). In 2000 Battarbee stated 'The most promising biological transfer function approach for direct temperature reconstruction is for chironomids' (p112). Since then a large amount of work has been carried out and numerous chironomid-inferred temperature reconstructions produced from many parts of the world (e.g., Woodward and Shulmeister, 2007; Eggermont *et al.*, 2010; Bunbury and Gajewski, 2012; Axford *et al.*, 2013; Massaferrro and Larocque-Tobler, 2013; Nazarova *et al.*, 2013; Berntsson *et al.*, 2014), with transfer functions constantly being developed and improved (e.g., Heiri *et al.*, 2011; Holmes *et al.*, 2011; Engels *et al.*, 2014; Luoto *et al.*, 2014). Other parameters, such as pH (e.g., Rees and Cwynar, 2010), salinity (e.g., Eggermont *et al.*, 2006), hypolimnetic oxygen (e.g., Summers *et al.*, 2012) and water depth (e.g., Engels *et al.*, 2012) can also be quantitatively reconstructed.

Quantitative chironomid data are frequently used in palaeoecological and palaeoclimate studies, however geomorphological studies using chironomid data are currently few. This article aims to introduce geomorphologists to chironomid analysis, outline the methods used, highlight some of the geomorphological studies which have used chironomid data and discuss how chironomid data could be further used to support geomorphological studies.

Methods

Field sampling

Key to obtaining a useful chironomid data set is obtaining reliable sediment samples to work on.

Corers

Depending on the nature of the sediments to be studied, and the focus of the study, different coring methods are used. In order to obtain a core from a lake it is nearly always necessary to have a well anchored boat or coring platform to work from. Long sequences can be obtained using piston corers (e.g., Livingstone, 1955, see Figure 1a), Mackereth corers (Mackereth, 1958),

percussion, hammer or tapper corers (Reasoner, 1986; Satake, 1988; Nesje, 1992, see Figure 1b), and Russian corers (Jowsey, 1966). Shorter sequences, or the sediment-water interface required for training sets, can be taken using gravity corers (Blomqvist and Abrahamsson, 1985; Glew, 1991; Renberg, 1991). Freeze corers can also be used to obtain sediments, and are particularly useful if laminated sequences are present (Renberg and Hansson, 2010). See *Geomorphological Techniques* Section 4.1.1 for more detailed information on coring methods.



Figure 1: (a) Obtaining a core from Cregganmore, County Mayo, Ireland, using a Livingstone corer (Photo: Steve Davis) (b) Returning to shore with a tapper core from Baulárvallavatn, Iceland.

Coring locations

Due to the time consuming nature of chironomid analysis it is common for only one sediment core from a site to be analysed for subfossil chironomids; this is true for both time-stratigraphical studies and for development of a training set. It is therefore important to select the most suitable location from which to obtain a core.

A number of studies have investigated the intra-lake variability of subfossil chironomid surface-samples and whether taking one core sample from a lake is representative of the whole lake (Heiri, 2004; Eggermont *et al.*, 2007; Bunbury and Gajewski, 2008; Holmes *et al.*, 2009; Kurek and Cwynar, 2009; Luoto, 2010; 2012; van Hardenbroek *et al.*, 2011; Heggen *et al.*, 2012; Zhang *et al.*, 2013). Results differ in different types of lakes, but in general it is accepted that obtaining a sample from the centre of the lake basin will allow the whole lake assemblage to be represented (van Hardenbroek *et al.*, 2011, Heggen *et al.*, 2012).

Laboratory methods

Chironomid analysis is a time consuming process (Brooks *et al.*, 2007); it can frequently take one to two days to fully analyse a single sample. Standard chironomid laboratory methods are discussed in full in Brooks *et al.* (2007), however the methods are briefly detailed here.

Sample storage

Samples do not need to be worked on immediately but must be stored appropriately. Intact cores and wet sediment samples should be stored wrapped at approximately 4°C. Samples can be dried (air, oven or freeze dried) and should then be stored in a cool dry environment.

Sample preparation

A known amount of sediment (measured as volume of wet sediment or weight of dry sediment) should be deflocculated in 10% potassium hydroxide. This should be warmed to approximately 75°C for roughly 15 minutes; the sample should be frequently stirred and should not be allowed to boil as this may cause damage to head capsules. The sample should then be sieved, and can be passed through one sieve (90 µm mesh) or, in order to allow easier sorting, through two sieves (90 µm mesh and 180 or 212 µm mesh). The sievings are washed into small pots/vials for storage, keeping the two size fractions separate if two sieves were used.

If samples are particularly high in carbonate or clay it is suggested that, following sieving, the samples are ultrasounded before being re-sieved and then washed into small

pots/vials (Lang *et al.*, 2003). This has been found to significantly increase head capsule yields.

Sample sorting

The sievings from a sample are pipetted into a grooved perspex (Bogorov) sorter (Gannon, 1971). If this is not available, a scored petri dish may be used, though care must be taken to ensure all the sample is sorted through. The sorter is placed on a stereo microscope (under x20-25 magnification) with a light source and head capsules are picked out using fine forceps. The head capsules are placed into a small container filled with water (if using Hydro Matrix® mountant; other mountants may require the use of different procedures). It is good practice for all the head capsules in a sample to be picked out for identification. Ideally a sample will contain a minimum of 50 head capsules, although a sample size of 100-150 head capsules is considered to give more reliable results (Heiri and Lotter, 2001; Larocque, 2001; Quinlan and Smol, 2001). If less than 50 head capsules are found it may be necessary to prepare more sediment.

Microscope slide mounting

Hydro-Matrix® is a popular mountant used for preparing chironomid head capsule slides. It is a water-soluble, non-toxic, non-darkening, non-polluting, fast solidifying and drying permanent mounting compound (Imscope, n.d.). A drop of Hydro-Matrix® is placed onto a microscope slide. Between two to four head capsules are placed, ventral side up, into the Hydro-Matrix® drop. A 6 mm or 10 mm diameter cover slip is carefully placed on to the drop ensuring the head capsules are covered. It is possible to place six to eight drops on each microscope slide. Hydro-Matrix® takes about 20 minutes to solidify at a temperature of approximately 25°C so slides should be stored horizontally immediately after they are prepared.

Chironomid identification

A biological microscope is used for identification (x40 magnification). Chironomid head capsules (see Figure 2) are identified by using a range of keys and identification guides (e.g., Cranston, 1982; Oliver and Roussel, 1983; Wiederholm, 1983; Schmid, 1993; Rieradevall and Brooks, 2001; Brooks *et al.*, 2007; Andersen *et al.*, 2013; Larocque-

Tobler, 2014). Key characteristics, such as head capsule shape, ventromental plates, shape and number of teeth on the mentum (see Figure 2), and, if present, mandibles, are used to first categorise which subfamily a head capsule belongs to (Tanypodinae, Chironominae, Orthoclaadiinae, Diamesinae, Prodiamesinae and Podonominae). Following this, a key can be used to identify a specimen to as high a taxonomic level as possible. A complete head capsule is recorded as one, while an incomplete head capsule with over 50% of it present is recorded as half. Remains which comprise <50% of a head capsule are not counted.

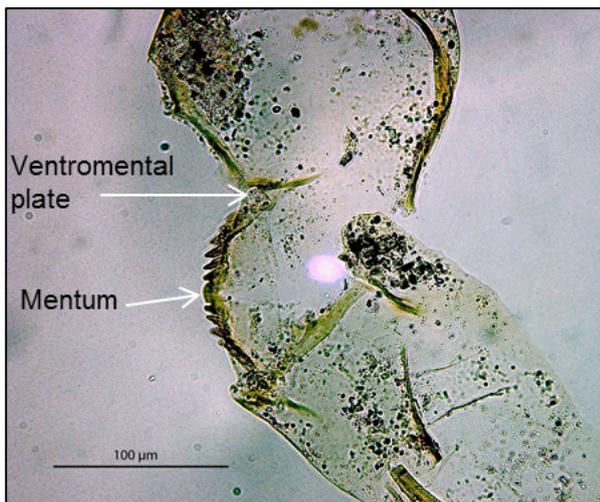


Figure 2: A chironomid head capsule.

Data presentation and analysis

Discussion of chironomid data presentation and analysis is beyond the scope of this paper; for information on these topics see Birks *et al.* (2012).

Geomorphological Applications

Although chironomid analysis is not frequently integrated into geomorphological studies, chironomid data have been used by some geomorphologists. The quantitative reconstructions produced by chironomid studies provide an independent record which can be used to help support geomorphological studies (e.g., Benn and Ballantyne, 2005; Giguet-Covex *et al.*, 2012; Doughty *et al.*, 2013). Chironomid data are also produced with the primary aim of informing geomorphological investigations (e.g., Dickson *et al.*, 2014).

Glacial geomorphology and palaeoclimate

Geomorphological mapping of past (Younger Dryas) glacier extent has been carried out in many parts of Scotland (e.g., Cornish, 1981; Ballantyne, 1989; 2006; Finlayson, 2006). Past Equilibrium Line Altitudes (ELAs) have been estimated; along with a past temperature estimate these can be used to calculate palaeoprecipitation using modern relationships between temperature, precipitation and ELA. Ballantyne (2002) suggested that an independent proxy palaeotemperature record could provide the estimate of past temperature. However it was thought that the Whitrig Bog chironomid record (Brooks and Birks, 2000), the only independent proxy temperature record for Scotland, was located too far away from the study location to be representative (Ballantyne, 2002). Further work used the Whitrig Bog chironomid temperature reconstruction, along with unpublished data from Abernethy forest, to calculate palaeotemperature for past regional ELAs (Benn and Ballantyne, 2005). Using these data it was possible to derive past precipitation estimates and to investigate past precipitation gradients (Benn and Ballantyne, 2005). A number of similar studies have been carried out and these have allowed for a more detailed understanding of palaeoclimate, particularly precipitation, and glacier land system behaviour in Scotland during the Younger Dryas (Benn and Lukas, 2006; Ballantyne, 2007; Lukas and Bradwell, 2010).

Doughty *et al.* (2013) used a combination of geomorphological mapping, a coupled energy-balance and ice-flow model and chironomid-inferred temperature data to evaluate late glacial temperatures and investigate the Antarctic Cold Reversal (ACR) in New Zealand. Chironomid-inferred temperature reconstructions produced by Vandergoes *et al.* (2008) were used to force the model (see Figure 3). Doughty *et al.* (2013) compared the glacier simulations produced to the mapped and dated glacial geomorphological (moraine) record of Kaplan *et al.* (2010). The simulated glacier length and the geomorphological record were found to provide consistent information about climate, and suggest that during the ACR temperatures in New Zealand were 2-3°C

cooler than present day (Doughty *et al.*, 2013).

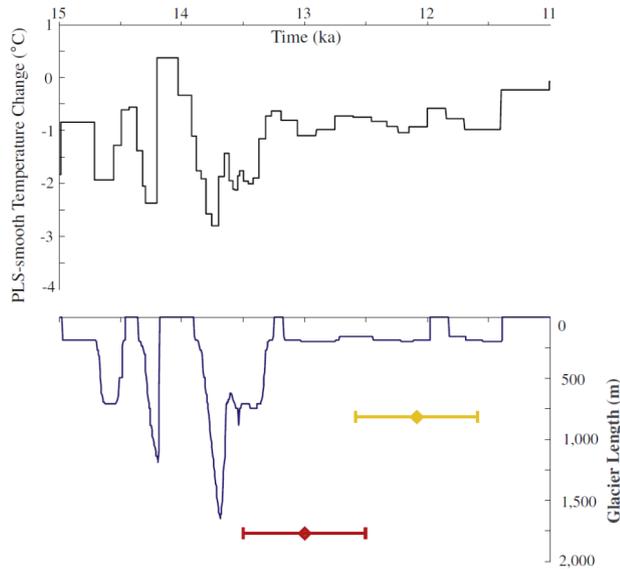


Figure 3: Chironomid-inferred temperature reconstruction (top) and modelled glacier length (bottom). Orange and yellow diamonds represent the moraine positions dated to 13.0 ± 0.5 and 12.1 ± 0.5 ka respectively. Source: Adapted from Figure 5 in Doughty *et al.* (2013).

Sea-level change and isostatic uplift

Subfossil chironomids can be used as indicators of salinity (Heinrichs and Walker, 2006). It is possible to both qualitatively and quantitatively reconstruct past salinity and these records can provide information about past sea levels (Heinrichs and Walker, 2006).

Solem *et al.* (1997) analysed sediment cores from coastal lakes in Norway in order to investigate invertebrate colonisation following isolation and uplift. It was possible to estimate the timing of isolation and the timescale over which the saline water was replaced by freshwater. At one of the sites it was estimated that this occurred within a 40 year time period (Solem *et al.*, 1997). Hofmann and Winn (2000) analysed a number of sites in the Western Baltic Sea in order to study the *Littorina* marine transgression. Before the transgression, the sites studied were separate freshwater bodies, which supported a diverse chironomid fauna (see Figure 4) (Hofmann and Winn, 2000). Following the transgression the sites supported three chironomid fauna (*Clunio marinus*, *Chironomus salinarius* and *Cricotopus/Halocladius*), all representative of brackish environments (Hofmann and Winn, 2000).

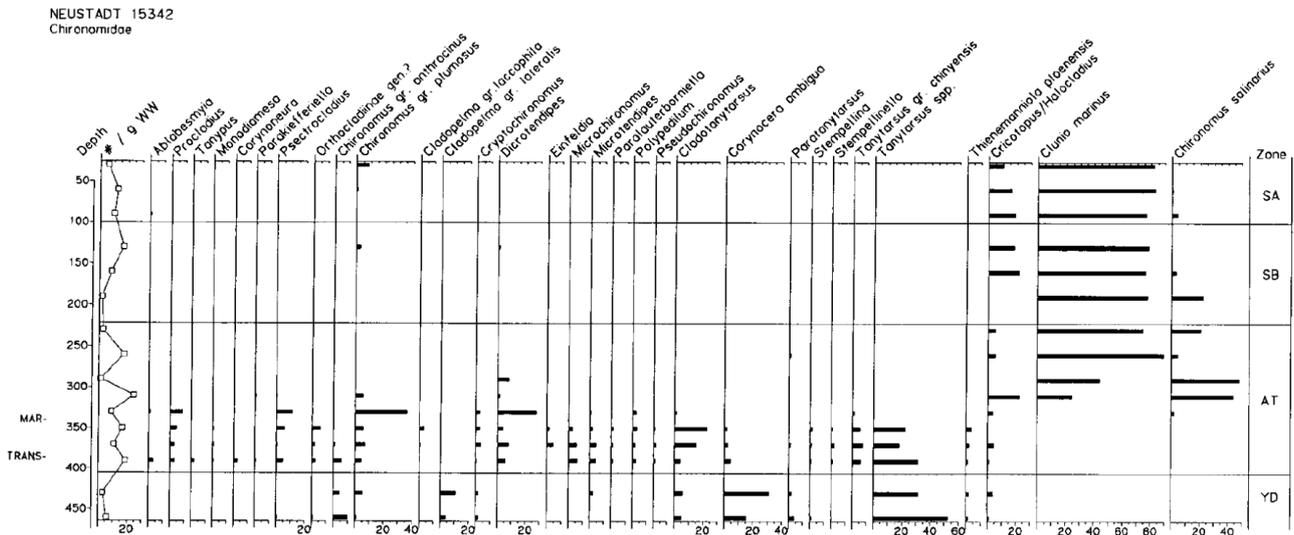


Figure 4: The Neustadt Bay (core 15342) chironomid percentage diagram showing changes in chironomid assemblages through the core. Freshwater assemblages are present in the lower samples, before an assemblage indicative of brackish water (*Clunio marinus*, *Chironomus salinarius* and *Cricotopus/Halocladius*) appears. TRANS indicates the position of the transgression horizon. MAR indicates the onset of marine conditions. Zone reflects pollen zones (YD – Younger Dryas; AT – Atlantic; SB – Subboreal; SA – Subatlantic). Source: Figure 2 in Hofmann and Winn (2000).

Rosenberg *et al.* (2005) produced a quantitative chironomid-inferred salinity reconstruction and discussed this in the context of past coastal emergence and submergence, suggesting that the record provided useful information about basin isolation from the sea. It is thought that a short-lived peak in salinity may be the result of marine incursion (by a storm surge or tsunami) (Rosenberg *et al.*, 2005), suggesting palaeo-tsunamis could be investigated using chironomid records from coastal locations. Cranston (2007) studied the impact of the 2004 Indian Ocean tsunami on aquatic habitats in coastal south west Thailand on chironomids. Studies such as this could help inform palaeo-tsunami studies.

Dickson *et al.* (2014) developed a chironomid-salinity transfer function with the specific purpose of investigating sea-level change due to isostatic uplift. The next steps in their research will apply this transfer function to four long cores in order to produce reconstructions of sea level in the Hudson Bay lowlands, an area which has experienced relatively recent isostatic rebound which continues today (Dickson *et al.*, 2014).

Palaeohydrology

Most subfossil chironomid studies are carried out on lacustrine deposits, however many chironomids are found in rivers (Pinder, 1995) and alluvial deposits offer an opportunity for the study of subfossil chironomids. Gandouin *et al.* (2005) undertook a study of chironomids in floodplain deposits. Chironomid taxa were classified into three categories: lentic habitat taxa (associated with oxbows or temporarily connected side arms); lotic habitat taxa (associated with the main channel or connected side arms); and other chironomids (includes taxa which could not be fully identified and taxa whose distribution is unknown or not related to current flow rate). Gandouin *et al.* (2006) introduced a fourth category 'ubiquitous taxa'. Using these classifications it was possible to qualitatively reconstruct past hydrological conditions and river morphology, including past connectivity (Gandouin *et al.*, 2005, 2006, 2007). In a study of the lateglacial-Holocene transition

Gandouin *et al.* (2007) found a shift in chironomid taxa from cold-water adapted and rheophilous (lotic) taxa to warm-water adapted and limnophilous (lentic) taxa, as the river switched from a braided to a meandering system (see Figure 5). Chironomids can be a useful technique for reconstructing palaeoenvironmental change in river floodplains due to the fact that they are influenced by hydrological changes, which are often induced by climatic changes (Gandouin *et al.*, 2007).

Existing chironomid data have also been used to inform palaeo-hydrological studies. Giguët-Covex *et al.* (2012) compared a palaeo-flood reconstruction with a chironomid-inferred temperature reconstruction (Millet *et al.*, 2009) from the same core. This enabled relationships between flood intensity and temperature to be investigated (Giguët-Covex *et al.*, 2012); with the finding that floods of higher energy generally occurred during periods of warmer temperatures.

Potential applications

The ability of subfossil chironomids to reconstruct a number of variables such as air temperature, lake water depth, and salinity suggests that chironomids can provide useful palaeoenvironmental data which can be used in geomorphological studies. As well as providing long timescale data, chironomids may also be able to provide information about short-lived events, such as past storm surges and tsunamis (Heinrichs and Walker, 2006), and palaeofloods (Gandouin *et al.*, 2006).

Recently, humans have become important agents of geomorphological change (Lóczy and Sütő, 2011). Chironomids have been found to respond to changing amounts of fine sediment (Angradi, 1999). As land use changes and erosion increase, sediment and organic matter inputs into streams and rivers will vary, potentially causing a change in the chironomid record (Angradi, 1999). Chironomid analysis has the potential to identify human activity and its impacts on lacustrine and fluvial sites (e.g., Cohen *et al.*, 2005; Cao *et al.*, 2013; Zhang *et al.*, 2013; Frossard *et al.*, 2014).

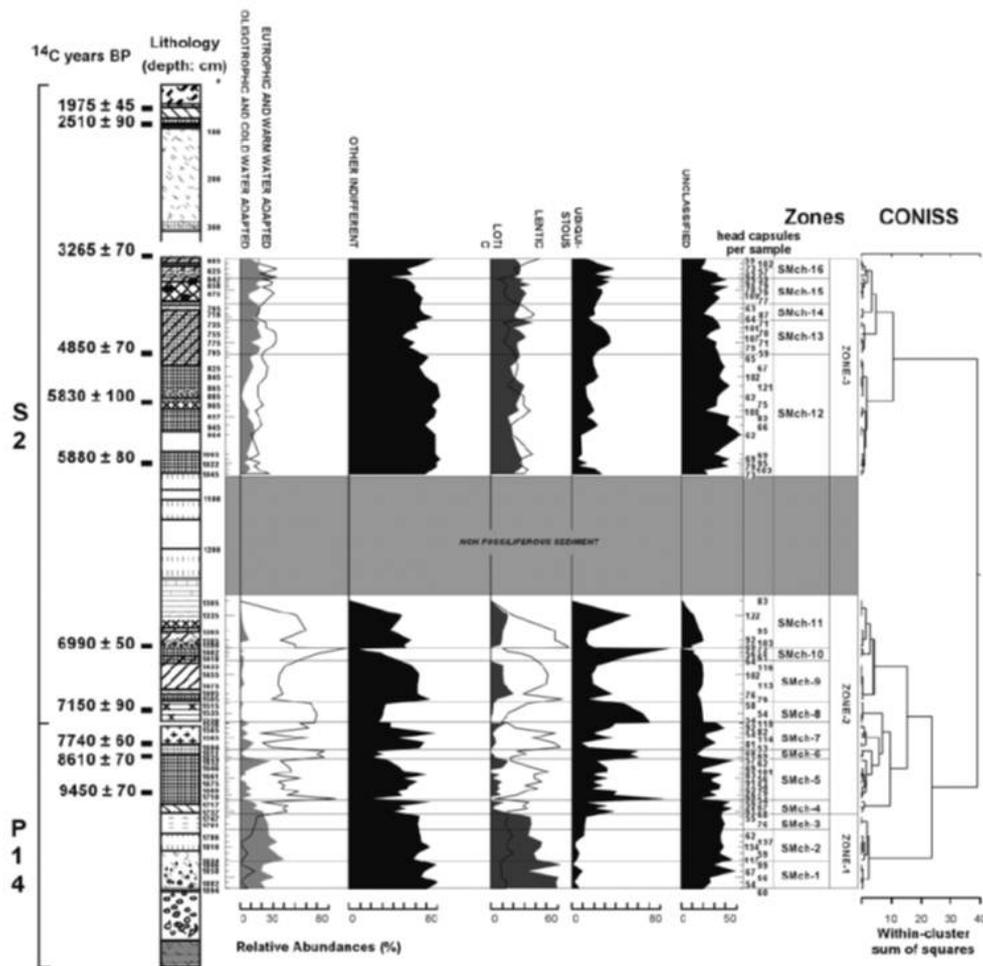


Figure 5: Synthesis diagram showing relative abundances (%) of cold-water adapted, warm-water adapted, eurythermous, lotic, lentic, ubiquitous and other unclassified taxa vs. depth. There is a shift from cold-water adapted lotic taxa to warm-water adapted lentic taxa as the river system changed from braided to meandering. Source: Figure 5 in Gandouin et al., (2007).

It is hoped that, as more geomorphologists become aware of the potential of chironomid and other palaeoecological data, both qualitative and quantitative, more interdisciplinary studies will be undertaken, particularly contributing to the improved understanding of geomorphological processes in the Holocene, a time period for which limited environmental data are available.

References

- Andersen, T., Cranston, P.S. and Epler, J.H. (Eds.). 2013. The larvae of Chironomidae (Diptera) of the Holarctic Region — Keys and diagnoses. *Insect Systematics & Evolution*. Supplement **66**: 1- 571.
- Angradi, T.R. 1999. Fine Sediment and Macroinvertebrate Assemblages in Appalachian Streams: A Field Experiment with Biomonitoring Applications. *Journal of the North American Benthological Society*. **18**: 49-66.
- Axford, Y., Losee, S., Briner, J.P., Francis, D.R., Langdon, P.G., and Walker, I.R. 2013. Holocene temperature history at the western Greenland Ice Sheet margin reconstructed from lake sediments. *Quaternary Science Reviews*. **59**: 87-100.
- Ballantyne, C.K. 1989. The Loch Lomond Advance on the Isle of Skye, Scotland: glacier reconstruction and palaeoclimatic implications. *Journal of Quaternary Science*. **4**: 95-108.

- Ballantyne, C.K. 2002. The Loch Lomond Readvance on the Isle of Mull, Scotland: glacier reconstruction and palaeoclimatic implications. *Journal of Quaternary Science*. **17**: 759-771.
- Ballantyne, C.K. 2006. Loch Lomond Stadial glaciers in the Uig Hills, western Lewis, Scotland. *Scottish Geographical Journal*. **11**: 256-273.
- Ballantyne, C.K. 2007. The Loch Lomond readvance on North Arran, Scotland: Glacier reconstruction and palaeoclimatic implications. *Journal of Quaternary Science*. **22**: 343-359.
- Battarbee, R.W. 2000. Palaeolimnological approaches to climate change, with special regard to the biological record. *Quaternary Science Reviews*. **19**: 107-124.
- Benn, D.I. and Ballantyne, C.K. 2005. Palaeoclimatic reconstruction from Loch Lomond Readvance glaciers in the West Drumochter Hills, Scotland. *Journal of Quaternary Science*. **20**: 577-592.
- Benn, D.I. and Lukas, S. 2006. Younger Dryas glacial landsystems in North West Scotland: an assessment of modern analogues and palaeoclimatic implications. *Quaternary Science Reviews* **25**: 2390-2408
- Berntsson, A., Rosqvist, G.C. and Velle, G. 2014. Late-Holocene temperature and precipitation changes in Vindelfallen, mid-western Swedish Lapland, inferred from chironomid and geochemical data. *Holocene*. **24**: 78-92.
- Birks, H.J.B. 1995. Quantitative Palaeoenvironmental Reconstructions. In: Maddy, D. and Brew, J.S. (Eds.) *Statistical Modelling of Quaternary Science Data*. Quaternary Research Association, Cambridge. pp161-254.
- Birks, H.J.B. 1998. Numerical tools in palaeolimnology – Progress, potentialities, and problems. *Journal of Paleolimnology*. **20**: 307-332.
- Birks, H.J.B., Lotter, A.F., Juggins, S. and Smol, J.P. (Eds.) 2012. *Tracking Environmental Change using Lake Sediments Volume 5. Data Handling and Numerical Techniques*. Springer, Dordrecht. pp745.
- Blomqvist, S. and Abrahamsson, B. 1985. An improved Kajak-type gravity core sampler for soft bottom sediments. *Swiss Journal of Hydrology*. **47**: 81-84.
- Brooks, S.J. and Birks, H.J.B. 2000. Chironomid-inferred Late-glacial air temperatures at Whitrig Bog, Southeast Scotland. *Journal of Quaternary Science*. **15**: 759-764.
- Brooks, S.J., Langdon, P.G., and Heiri, O. 2007. *The Identification and Use of Palaeartic Chironomidae Larvae in Palaeoecology*. Quaternary Research Association Technical Guide No. 10. Quaternary Research Association, London. pp276.
- Bunbury, J. and Gajewski, K. 2008. Does a one point sample adequately characterize the lake environment for paleoenvironmental calibration studies. *Journal of Paleolimnology*. **39**: 511-531.
- Bunbury, J. and Gajewski, K. 2012. Temperatures of the past 2000 years inferred from lake sediments, southwest Yukon Territory, Canada. *Quaternary Research*. **77**: 355-367.
- Cao, Y., Zhang, E., Langdon, P., Liu, E. and Shen, J. 2013. Spatially different nutrient histories recorded by multiple cores and implications for management in Taihu lake, eastern China. *Chinese Geographical Science*. **23**: 537-549.
- Cohen, A.S., Palacios-Fest, M.R., Msaky, E.S., Alin, S.R., McKee, B., O'Reilly, C.M., Dettman, D.L., Nkotagu, H. and Lezzar, K.E. 2005. Paleolimnological investigations of anthropogenic environmental change in Lake Tanganyika: IX. Summary of paleorecords of environmental change and catchment deforestation at Lake Tanganyika and impacts on the Lake Tanganyika ecosystem. *Journal of Paleolimnology*. **34**: 125-145.
- Cornish, R. 1981. Glaciers of the Loch Lomond Stadial in the western Southern Uplands of Scotland. *Proceedings of the Geologists' Association*. **92**: 105-114.

- Cranston, P.S. 1982. *A key to the larvae of the British Orthocladiinae (Chironomidae)*. Freshwater Biological Association, Kendal. pp152.
- Cranston, P.S. 1995. Introduction. In: Armitage, P.D., Cranston, P.S. and Pinder, L.C.V. (Eds.) *The Chironomidae: Biology and ecology of non-biting midges*. Chapman and Hall, London. pp1-10.
- Cranston, P.S. 2007. The Chironomidae larvae associated with the tsunami-impacted waterbodies of the coastal plain of southwestern Thailand. *The Raffles Bulletin of Zoology*. **55**: 231-144.
- Dickson, T.R., Bos, D.G., Pellatt, M.G. and Walker, I.R. 2014. A midge-salinity transfer function for inferring sea level change and landscape evolution in the Hudson Bay Lowlands, Manitoba, Canada. *Journal of Paleolimnology*. **51**: 325-341.
- Doughty, A.M., Anderson, B.M., Mackintosh, A.N., Kaplan, M.R., Vandergoes, M.J., Barrell, D.J.A., Denton, G.H., Schaefer, J.M., Chinn, T.J.H. and Putnam, A.E. 2013. Evaluation of Lateglacial temperatures in the Southern Alps of New Zealand based on glacier modelling at Irishman Stream, Ben Ohau Range. *Quaternary Science Reviews*. **74**: 160-169.
- Eggermont, H. and Heiri, O. 2012. The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. *Biological Reviews*. **87**: 430-456.
- Eggermont H., Heiri, O. and Verschuren, D. 2006. Subfossil Chironomidae (Insecta: Diptera) as quantitative indicators for past salinity variation in African lakes. *Quaternary Science Reviews*. **25**: 1966-1994.
- Eggermont, H., De Deyne, P. and Verschuren, D. 2007. Spatial variability of chironomid death assemblages in the surface sediments of a fluctuating tropical lake (Lake Naivasha, Kenya). *Journal of Paleolimnology*. **38**: 309-328.
- Eggermont, H., Heiri, O., Russell, J., Vuille, M., Audenaert, L. and Verschuren, D. 2010. Paleotemperature reconstruction in tropical Africa using fossil Chironomidae (Insecta: Diptera). *Journal of Paleolimnology*. **43**: 413-435.
- Engels, S., Cwynar, L.C., Rees, A.B.H. and Shurman, B.N. 2012. Chironomid-based water depth reconstructions: an independent evaluation of site-specific and local inference models. *Journal of Paleolimnology*. **48**: 693-709.
- Engels, S., Self, A.E., Luoto, T.P., Brooks, S.J. and Helmens, K.F. 2014. A comparison of three Eurasian chironomid-climate calibration datasets on a W-E continentality gradient and the implications for quantitative temperature reconstructions. *Journal of Paleolimnology*. **51**: 529-547.
- Finlayson, A.G. 2006. Glacial geomorphology of the Creag Meagaidh Massif, Western Grampian Highlands: implications for local glaciation and palaeoclimate during the Loch Lomond Stade. *Scottish Geographical Journal*. **122**: 293-307.
- Frossard, V., Millet, L., Verneaux, V., Jenny, J.P., Arnaud, F., Magny, M. and Perga, M.E. 2014. Depth-specific responses of a chironomid assemblage to contrasting anthropogenic pressures: a palaeolimnological perspective from the last 150 years. *Freshwater Biology*. **59**: 26-40.
- Gandouin, E., Franquet, E. and Van Vliet-Lanoë, B. 2005. Chironomids (Diptera) in river floodplains: their status and potential use for palaeoenvironmental reconstruction purposes. *Archive fur Hydrobiologie*. **162**: 511-534.
- Gandouin, E., Maasri, A., Van Vliet-Lanoë, B. and Franquet, E. 2006. Assemblages from a Gradient of Lotic and Lentic Waterbodies in River Floodplains of France: A Methodological Tool for Paleoeological Applications. *Journal of Paleolimnology* **35**: 149-166.
- Gandouin, E., Ponel, P., Franquet, E., Van Vliet-Lanoë, B., Andrieu-Ponel, V., Keen, D.H., Brulhet, J. and Brocandel, M. 2007. Chironomid responses (Insect: Diptera) to Younger Dryas and Holocene environmental changes in a river floodplain from northern France (St-Momelin, St-Omer basin). *Holocene*. **17**: 331-347.

- Gannon, J.E. 1971. Two counting cells for the enumeration of zooplankton, microcrustacea. *Transactions of American Microscopical Society*. **90**: 486-490.
- Giguet-Covex, C., Arnaud, F., Enters, D., Poulencard, J., Millet, L., Francus, P., David, F., Rey, P.J., Wilhelm, B. and Delannoy, J.J. 2012. Frequency and intensity of high-altitude floods over the last 3.5 ka in northwestern French Alps (Lake Anterne). *Quaternary Research*. **77**: 12-22.
- Glew, J.R. 1991. Miniature gravity corer for recovering short sediment cores. *Journal of Paleolimnology*. **5**: 285-287.
- Heggen, M.P., Birks, H.H., Heiri, O., Grytnes, J.A. and Birks, H.J.B. 2012. Are fossil assemblages in a single sediment core from a small lake representative of total deposition of mite, chironomid, and plant macrofossil remains? *Journal of Paleolimnology*. **48**: 669-691.
- Heinrichs, M.L., Walker, I.R. 2006. Fossil midges and palaeosalinity: potential as indicators of hydrological balance and sea-level change. *Quaternary Science Reviews*. **25**: 1948-1965.
- Heiri, O. 2004. Within-lake variability of subfossil chironomid assemblages in shallow Norwegian lakes. *Journal of Paleolimnology*. **32**: 67-84.
- Heiri, O. and Lotter, A.F. 2001. Effect of low count sums on quantitative environmental reconstructions: an example using subfossil chironomids. *Journal of Paleolimnology*. **26**: 343-350.
- Heiri, O., Brooks, S.J., Birks, H.J.B. and Lotter, A.F. 2011. A 274-lake calibration data-set and inference model for chironomid-based summer air temperature reconstruction in Europe. *Quaternary Science Reviews*. **30**: 3445-3456.
- Hofmann, W., & Winn, K. 2000. The Littorina Transgression in the Western Baltic Sea as Indicated by Subfossil Chironomidae (Diptera) and Cladocera (Crustacea). *International Review of Hydrobiology*. **85**: 267-291
- Holmes, N., Langdon, P.G. and Caseldine, C.J. 2009. Subfossil chironomid variability in surface samples from Icelandic lakes: implications for the production and use of training sets. *Journal of Paleolimnology*. **42**: 281-295.
- Holmes, N., Langdon, P.G., Brooks, S.J., Caseldine, C.J. and Birks, H.J.B. 2011. Merging chironomid training sets: implications for palaeoclimate reconstructions. *Quaternary Science Reviews*. **30**: 2793-2804.
- Jowsey, P.C. 1966. An improved peat sampler. *New Phytologist*. **65**: 245-248.
- Juggins S. 2013. Quantitative reconstructions in palaeolimnology: new paradigm or sick science? *Quaternary Science Reviews*. **64**: 20-32.
- Kaplan, M.R., Schaefer, J.M., Denton, G.H., Barrell, D.J.A., Chinn, T.J.H., Putnam, A.E., Anderson, B.G., Finkel, R.C., Schwartz, R. and Doughty, A.M. 2010. Glacier retreat in New Zealand during the Younger Dryas stadial. *Nature*. **467**: 194-197.
- Kurek, J. and Cwynar, L.C. 2009. Effects of within-lake gradients on the distribution of fossil chironomid from maar lakes in western Alaska: implications for environmental reconstructions. *Hydrobiologia*. **623**: 37-52.
- Lang, B., Bedford, A.P., Richardson, N. and Brooks, S.J. 2003. The use of ultra-sound in the preparation of carbonate and clay sediments for chironomid analysis. *Journal of Paleolimnology*. **30**: 451-460.
- Larocque, I. 2001. How many chironomid head capsules are enough? A statistical approach to determine sample size for palaeoclimatic reconstructions. *Palaeogeography Palaeoclimatology Palaeoecology*. **172**: 133-142.
- Larocque-Tobler, I. 2014. The Polish sub-fossil chironomids. *Palaeontologia Electronica*. **17**: 1:3A.
- Livingstone, D.A. 1955. A lightweight piston sampler for lake deposits. *Ecology*. **36**: 137-139.
- Imscope. no date. *Technical Data Hydro-Matrix®*. Available: <http://www.imscope.com/produkt22/HYDRO>

[MATRIX Daten en.shtml](#) [Last accessed 24/04/14]

- Lóczy, D. and Sütő, L. 2011. Human Activity and Geomorphology. In: Gregory, K.J. and Goudie, A.S. (Eds.) *The SAGE Handbook of Geomorphology*. SAGE Publications Ltd, London. pp260-278.
- Lukas, S. and Bradwell, T. 2010. Reconstruction of a Lateglacial (Younger Dryas) mountain ice field in Sutherland, northwestern Scotland, and its palaeoclimatic implications. *Journal of Quaternary Science*. **25**: 567-580.
- Luoto, T.P. 2010. Spatial uniformity in depth optima of midges: evidence from sedimentary archives of shallow Alpine and boreal lakes. *Journal of Limnology*. **71**: 228-232.
- Luoto, T.P. 2012. Intra-lake patterns of aquatic insect and mite remains. *Journal of Paleolimnology*. **47**: 141-157.
- Luoto, T.P., Kaukolehto, M., Weckstrom, J., Korhola, A. and Väiliranta, M. 2014. New evidence of warm early-Holocene summers in subarctic Finland based on an enhanced regional chironomid-based temperature calibration model. *Quaternary Research*. **81**: 50-62.
- Mackereth, F.J.K. 1958. A portable core sampler for lake deposits. *Limnology and Oceanography*. **3**: 181-191.
- Marziali, L. and Rossaro, B. 2013. Response of chironomid species (Diptera, Chironomidae) to water temperature: effects on species distribution in specific habitats. *Journal of Entomological and Acarological Research*. **45**: 73-89.
- Massaferro, J. and Larocque-Tobler, I. 2013. Using a newly developed chironomid transfer function for reconstructing mean annual air temperature at Lake Potrok Aike, Patagonia, Argentina. *Ecological Indicators*. **24**: 201-210.
- Millet, L., Arnaud, F., Heiri, O., Magny, M., Verneaux, V., Desmet, M., 2009. Late-Holocene summer temperature reconstruction from chironomid assemblages of Lake Anterne, northern French Alps. *The Holocene*. **19**: 317-328.
- Nazarova, L., de Hoog, V., Hoff, U., Dirksen, O. and Diekmann, B. 2013. Late Holocene climate and environmental changes in Kamchatka inferred from the subfossil chironomid record. *Quaternary International*. **67**: 81-92.
- Nesje, A. 1992. A piston corer for lacustrine and marine sediments. *Arctic and Alpine Research*. **24**: 257-259.
- Oliver, D.R. 1971. Life history of the Chironomidae. *Annual Review of Entomology*. **16**: 211-230.
- Oliver, D.R. and Roussel, M.E. 1983. *The insects and arachnids of Canada, Part 11: The genera of larval midges of Canada*. Agriculture Canada Publication, Ottawa. pp263.
- Pinder, L.C.V. 1983. The larvae of Chironomidae (Diptera) of the Holarctic region – Introduction. In: Wiederholm, T. (Ed.) 1983. Chironomidae of the Holarctic region. Keys and diagnoses. Part I. Larvae. *Entomological Scandinavica Supplement*. **19**: 7-10.
- Pinder, L.C.V. 1995. The habitats of chironomid larvae. In: Armitage, P.D., Cranston, P.S. and Pinder, L.C.V. (Eds.) *The Chironomidae: Biology and ecology of non-biting midges*. Chapman and Hall, London. pp107-135.
- Quinlan, R. and Smol, J.P. 2001. Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. *Journal of Paleolimnology*. **26**: 327-342.
- Rae, J.G. 2013. Abiotic factors affect microhabitat selection and community dynamics in a sandy-bottom lotic chironomid midge assemblage. *Hydrobiologia*. **700**: 121-130.
- Reasoner, M.A. 1986. An inexpensive, lightweight percussion core sampling system. *Géographie Physique et Quaternaire*. **40**: 217-219.
- Rees, A.B.H., and Cwynar, L.C. 2010. A test of Tyler's Line - response of chironomids to a pH gradient in Tasmania and their potential

- as a proxy to infer past changes in pH. *Freshwater Biology*. **55**: 2521-2540.
- Renberg, I. 1991. The HON-Kajak sediment corer. *Journal of Paleolimnology*. **6**: 167-170.
- Renberg, I. and Hansson, H. 2010. Freeze corer No. 3 for lake sediments. *Journal of Paleolimnology*, **44**: 731-736.
- Rieradevall, M. and Brooks, S.J. 2001. An identification guide to subfossil Tanypodinae larvae (Insecta: Diptera: Chironomidae) based on cephalic setation. *Journal of Paleolimnology*. **25**: 81-99.
- Rosenberg, S. M., Walker, I.R. and Macpherson, J.B. 2005. Environmental changes at Port au Choix as reconstructed from fossil midges. *Newfoundland and Labrador Studies* **20**: 57-73.
- Satake, K. 1988. A handy impact corer for sampling lake surface sediment. *Hydrobiologia*. **169**: 259-264.
- Schmid, P.E. 1993. A key to the larval Chironomidae and their instars from Austrian Danube region streams and rivers. Part 1: Diamesinae, Prodiamesinae and Orthoclaadiinae. *Wasser und Abwasser Supplementband*. **3/93**: 1-513.
- Solem, J.O., Solem, T., Aagaard, K., and Hanssen, O. 1997. Colonization and evolution of lakes on the central Norwegian coast following deglaciation and land uplift 9500 to 7800 years B.P. *Journal of Paleolimnology*. **18**: 269-281.
- Summers, J.C., Ruhland, K.M., Kurek, J., Quinlan, R., Paterson, A.M. and Smol, P. 2012. Multiple stressor effects on water quality in Poplar Bay, Lake of the Woods, Canada: a midge-based assessment of hypolimnetic oxygen conditions over the last two centuries. *Journal of Limnology*. **71**: 34-44.
- van Hardenbroek, M., Heiri, O., Wilhelm, M.F. and Lotter, A.F. 2011. How representative are subfossil assemblages of Chironomidae and common benthic invertebrates for the living fauna of Lake De Waay, the Netherlands? *Aquatic Sciences*. **73**: 247-259.
- Vandergoes, M.J., Dieffenbacher-Krall, A.C., Newnham, R.M., Denton, G.H. and Blaauw, M. 2008. Cooling and changing seasonality in the Southern Alps, New Zealand during the Antarctic cold reversal. *Quaternary Science Reviews*. **27**: 589-601.
- Velle, G., Brodersen, K.P., Birks, H.J.B. and Willassen, E. 2012. Inconsistent results should not be overlooked. *Holocene*. **22**: 1501-1508.
- Walker, I.R. 2001. Midges: Chironomidae and related dipteral. In: Smol, J.P., Birks, H.J.B. and Last, W. (eds.) *Tracking Environmental Change Using Lake Sediments. Volume 4: Zoological Indicators*. Kluwer, Dordrecht. pp43-66.
- Wiederholm, T. (Ed.) 1983. Chironomidae of the Holarctic region. Keys and diagnoses. Part I. Larvae. *Entomological Scandinavica Supplement*. **19**: 1-457.
- Woodward, C.A. and Shulmeister, J. 2007. Chironomid-based reconstructions of summer air temperature from lake deposits in Lyndon Stream, New Zealand spanning the MIS 3/2 transition. *Quaternary Science Reviews*. **26**: 142-154.
- Zhang, E., Cao, Y., Langdon, P., Wang, Q., Shen, J. and Yang, X. 2013. Within-lake variability of subfossil chironomid assemblage in a large, deep subtropical lake (Lugu lake, Southwest China). *Journal of Limnology*. **72**: 117-126.

Pollen Analysis: Not Just a Qualitative Tool

Claire L. Twiddle¹

¹Geography and Environment, University of Aberdeen, Elphinstone Road, Aberdeen, AB24 3UF
(email: C.L.Twiddle@abdn.ac.uk)



ABSTRACT: Pollen analysis is seen as a largely qualitative method, but recent developments in analytical techniques, technologically and theoretically, have directed a move towards a greater use of quantitative interpretation. Use of these new methods allows more detailed information to be obtained from pollen data and in turn widens the scope for use of this technique. It is often assumed that pollen analysis only informs on the terrestrial changes in vegetation, but these analytical developments mean pollen data can be used to infer changes in climate, human activity and societal development, all of which are central to obtaining a holistic view of landscape change. This paper outlines the range of contributions pollen data can make to studies along with the considerations required to ensure reliable and informative data is obtained.

KEYWORDS: Fossil; Pollen; quantitative analysis; site selection

Introduction

Palaeoecology is the study of long-term vegetation patterns and encompasses a number of palaeoenvironmental techniques used to investigate past changes in these terrestrial environments, and pollen analysis is one of the most established methods. Being preserved in anoxic environments, pollen grains provide a valuable window into the past as identification to family, genus, and even species, level can be made from distinct morphological characteristics (Figure 1). First championed by von Post (1916) the theoretical and methodological developments that have occurred since this time have pushed the capabilities of the technique, and the current research agendas highlight further the potential that pollen analysis has to inform on key debates.

Sitting in a laboratory, at a microscope and counting pollen grains may not be as appealing as scaling mountains or wading in rivers, but the data obtained can be key to an understanding of human, climatic and wider environmental patterns and impacts in the past. The flora is intimately related to earth system processes and pollen analysis (palynology) can unlock the role of vegetation

in such systems over time. Information about the type, density and spatial patterns of past vegetation has implications for many areas of geographical science and assists in fine-tuning climate models. Importantly, global models are used to predict future changes, and a necessary stage is validation often using palaeoenvironmental, e.g. pollen, data (Anderson et al., 2006, Gaillard et al., 2010).



Figure 1: An image showing modern Picea pollen grains.

Pollen analysis is frequently used to investigate the environmental context of archaeological artefacts to enhance their interpretation, particularly with respect to

settlement, adoption of agriculture and woodland removal (Edwards and MacDonald, 1991; Edwards, 2004).

A further dimension comes from palynological contributions to landscape management via habitat conservation and the mitigation of predicted changes in climate (Birks, 1996, Willis et al., 2007, Willis and Birks, 2006). Current vegetation patterns are the result of complex interactions over many years and pollen analysis supplies longer-term views of change which can aid formulation of suitable action plans for the landscape.

Basics of pollen analysis

In order that pollen data can be used in the ways mentioned above, suitable samples have to be collected. This involves considerations of data collection, taphonomy / preservation and analysis.

Data collection and site selection

The main focus here will be upon lakes and peat bogs, as these often provide the most favourable conditions for preservation which are anoxic conditions preventing desiccation of the pollen grains. Such environments are used in the majority of pollen studies, but research have shown that palynomorphs (pollen and related microfossils) can be obtained from e.g. soils (Dimpleby, 1985),

speleothems (Caseldine et al., 2008), fluvial material (Brown, 1997; Brown, 1999) and marine cores (van Campo et al, 1982; Wilmshurst et al, 1999; McGlone, 2001). The preservation and differential source area of pollen from such contexts can hamper interpretation (Birks and Birks, 1980). In light of this, a number of issues have to be considered to ensure that the information obtained from palynology will be suitable to support any hypotheses being proposed.

A pollen sample combines locally and regionally derived pollen grains; the former is from vegetation immediately surrounding the depositional basin, whereas the latter is a homogenous representation of regional vegetation that is collected, carried and deposited by airflows. Also, in basins with inflows or steep catchments, a fluvial input can introduce a secondary long-distance component. These differing origins influence the recruitment of pollen between environments (Jacobson and Bradshaw, 1981; Sugita, 1994).

For example, an entire lake surface is the catchment for pollen transported by airflows, which is subsequently mixed through the water column to be deposited on the sediment surface at the bottom of the lake. In contrast, once pollen has been deposited on the surface of a peat bog there is little movement and behaves as a static point

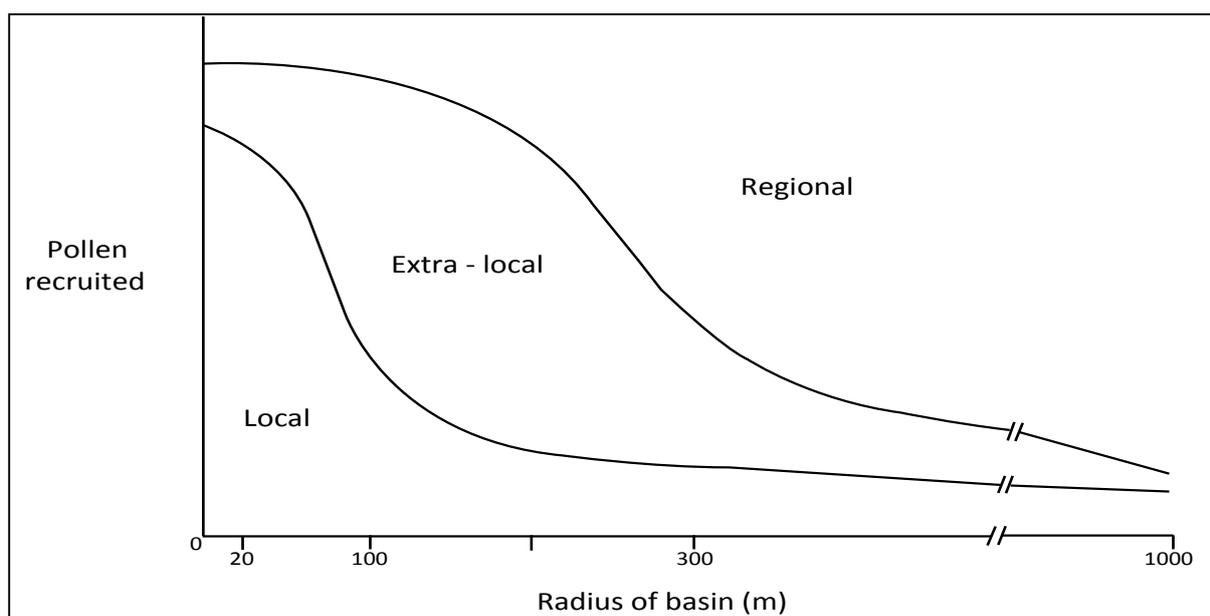


Figure 2: The recruitment of pollen to deposition basins (after Jacobson and Bradshaw, 1981).

source. Due to these differences, a pollen sample contained within a lake represents a much larger area of vegetation than that of a peat bog. Unsurprisingly, peat bogs often have a strong local representation from the mire vegetation (e.g. Cyperaceae [sedges]) as a pollen sample is a point source. In lakes, as the entire water surface collects pollen and is mixed before the pollen is deposited, the size of the lake also acts as a variable. For instance, the larger lake surface area the greater the amount of pollen contributed from the atmosphere, which in turn increases the area of the landscape that is reflected (cf. Figure 2) (Jacobson and Bradshaw, 1981; Sugita, 1994).

Simulations by Sugita (1994) show the effect that different lake radii have upon the signature for vegetation in the pollen assemblages. Figure 3 shows that the representation of the landscape from lakes

with a radius $>750\text{m}$ display more homogenous patterns than smaller lakes which show greater heterogeneity. Taxon-specific factors also have a role in determining the occurrence and quantity of a taxon in a pollen sample. Firstly, plants that are insect-pollinated (entomophilous) tend to be under-represented in the pollen record compared to those that are wind-pollinated (anemophilous) as the former produce less pollen which are not as well equipped for wind dispersal. Secondly, the amount of pollen produced differs between plants and this can influence their representation in the fossil record. For example *Pinus* (Pine) is a high pollen producer, whereas *Juniperus* (Juniper) is comparably lower - therefore the presence of low quantities of *Juniperus* in the pollen record may be highly significant. Some environmental factors influencing pollen productivity include: climate, specifically July temperature of previous

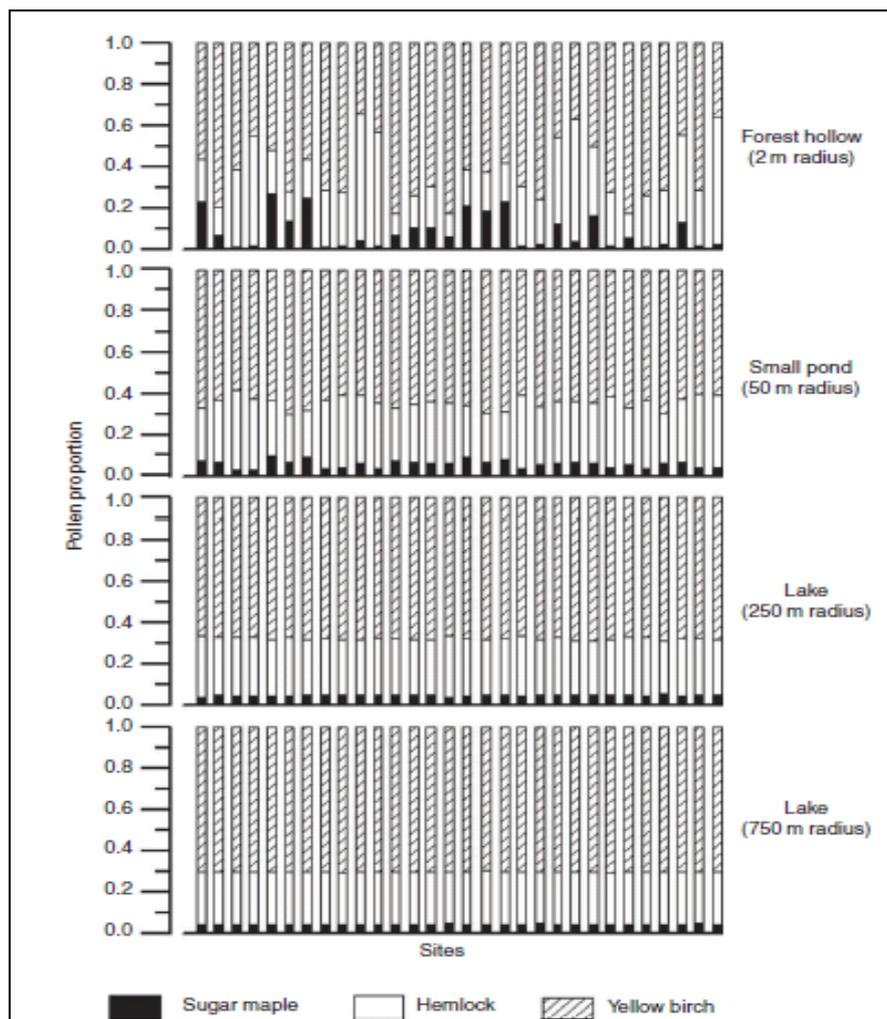


Figure 3: Simulated pollen proportions from different sized lakes within the same region (from Sugita, 1994).

years, which influences the flowering time (Barnekow *et al.*, 2007; Huusko and Hicks, 2009); vegetation patterning and structure which controls light availability; and location in relation to the taxon habitat range, as stressed individuals will apportion less energy to pollen production. Finally, as pollen grains are morphologically distinct, the size, shape and features can aid or hamper dispersal. Thus, *Pinus* pollen has two air bladders increasing its suitability and efficiency for wind dispersal, further influencing its over-representation in pollen samples. Research is underway to understand more fully the range of influences upon pollen productivity and the biases that are present in pollen assemblages (Broström *et al.*, 2008).

Taphonomy and preservation

Complications in deriving and interpreting pollen data arise from the depositional environment and specifically the preservation status of polleniferous material. In some situations, there can be mechanical, chemical and biological agents that breakdown the exine of the pollen grain. The severity of this damage can be recorded and used to further understand environmental conditions (Tweddle and Edwards, 2010), but in some cases the severity is such that it can induce bias into the pollen assemblage. In this case, it is very difficult to quantify the abundance of grains and species that have been affected although Bunting and Tipping (2000) devised a method to assess the degree of post-depositional influences and as such the amount of bias within the pollen assemblage. Studies have investigated the type and degree of degradation that can affect samples, and results suggest that the sporopollenin content of the grain's exine, which is species-specific, is a major control (Cushing, 1967; Havinga, 1984; Twiddle and Bunting, 2010). However, the processes responsible for the damage and the morphology of the grain also have a significant role and, as a result, details of this are uncertain.

A related issue is taphonomy, not only in terms of damage to grains but also their redistribution, and although this is primarily confined to lakes, peat bogs are not excluded (Clymo and Mackay, 1987). Processes such

as sediment focusing, scavenging and within-lake processes (e.g. thermal over-turning of the water body) redistribute grains and have been noted as influencing the relative proportions and concentration of grains (Davis, 1968; 1973; Davis and Brubaker, 1973; Davis *et al.*, 1984). Moreover, this can be species-specific, as for instance in the correlation between pollen release and the period of overturning in a lake (Davis, 1968). The occurrence and degree of influence these processes have is also exacerbated by lake basin morphology. Lehman (1975) showed that this characteristic gives rise to the uneven accumulation of sediment, and in turn the distribution of pollen over the sediment surface influencing the ability to replicate pollen records (Edwards, 1983). Also, within lakes fluvial input can distort the representation of the landscape vegetation, not only from inlets but also overland flow, as effectively shown in studies by Peck (1973), Bonny (1978) and more recently Wilmshurst and McGlone (2005). For this reason, closed basins, those with no inlet or outlet, are preferentially sought in order to minimise the risk of distortion from these sources. In many situations the issues described are difficult to detect and counteract. Sedimentological analyses and dating methods are often required, in addition to assessments of phenomena such as pollen condition, in order to assess adequately the likelihood and extent of such influences.

Data presentation

Qualitative presentation of data is normally made using percentages, but taxa are interrelated. Therefore, as taxon 1 varies taxon 2 mirrors the change regardless of whether these are occurring within the vegetation itself (Birks and Birks, 1980). Concentrations, the total grains deposited per unit volume of sediment, are often used to support interpretations from the percentage profile and to detect changes in the number of grains reaching the basin either as a result of floristic variations or changes in sedimentation (Figure 4). These are the most frequently used methods of data presentation. When a robust chronology is available, Pollen Deposition Rates (PDRs) or Pollen Accumulation Rates (PARs) – often called 'influx' – can be calculated. Defined by

Davis (1969) as 'the net number of grains accumulated per unit area of sediment surface per unit time', presented as $\text{cm}^{-2} \text{yr}^{-1}$; PDRs values give a better approximation of changes in the vegetation coverage or changes in the sedimentation rate and studies have demonstrated that variations within percentage data can be highlighted as artefacts of the method when PDRs are used as a comparison (Davis, 1967; Davis and Deevey, 1964; Hicks, 2001; Hyvärinen, 1975; Seppa and Hicks, 2006). Despite their benefits, limitations have to be considered and addressed on a site-specific basis (Bennett, 1994; Giesecke and Fontana, 2008; Pennington, 1996).

Analytical methods

The analytical methods used in pollen analysis have developed to provide semi-quantitative and quantitative data, a move from the more descriptive stance of some earlier studies. Use of these can require a greater investment, both in time and financially, during primary data collection and processing, but the output presents an alternative view of the raw data and can target specific hypotheses being addressed. Methods are available to directly obtain estimates of vegetation coverage and

landscape patterning or indirectly to derive climate information relationship.

Climate reconstruction

Pollen has been shown to respond to fluctuations in past temperatures and precipitation patterns because of its representation of vegetation patterns. This relationship, and abundant terrestrial sources, has encouraged the development and use of analytical techniques to derive past climatic information from pollen data.

One of the earliest analytical techniques proposed was the indicator species approach (Iversen, 1944) which uses overlapping climate ranges of specific taxa to derive estimates of past temperature regimes. From this the Modern Analogue Technique (Overpeck et al., 1985), the biomes approach (Prentice et al., 1996) and more recently the transfer function have all developed and have been applied in multiple studies to test the underlying methods and assumptions. Table 1 provides an introduction to the aims of the methods, key references associated with them and some of the limitations.

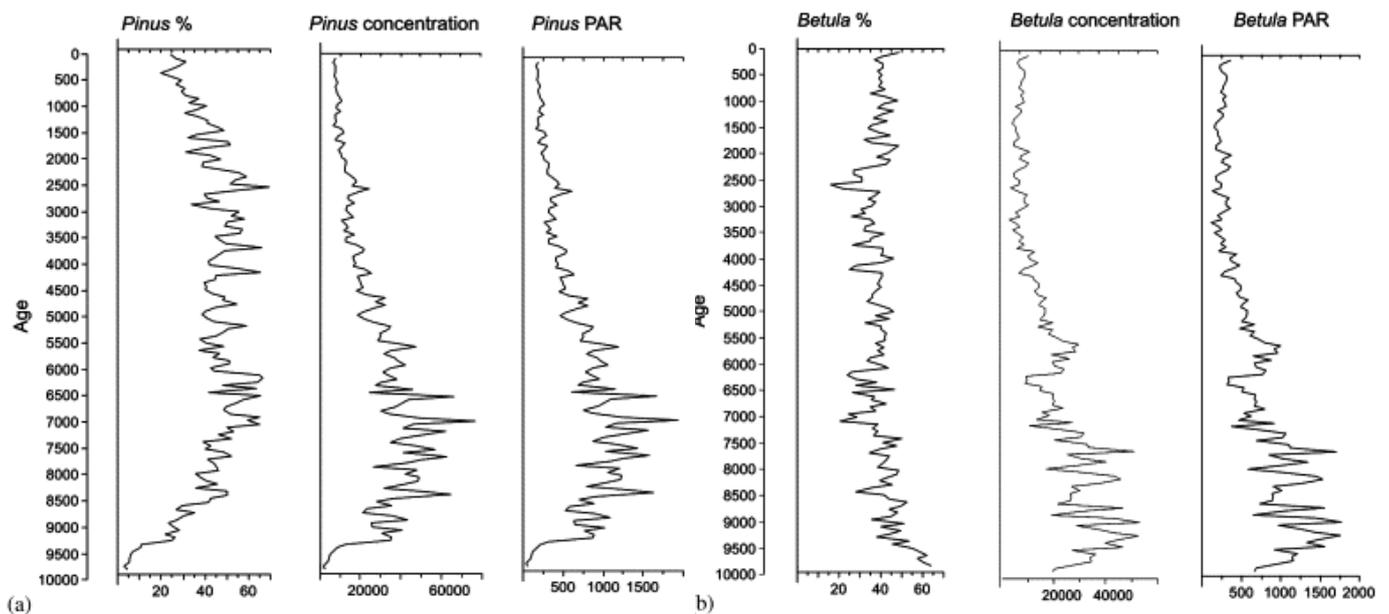


Figure 4: Differences between pollen data is presented using different methods including percentage, concentration and PAR format, (a) *Pinus* and (b) *Betula*. The data were collected from sites in northern Fennoscandia (Seppa and Hicks, 2006).

Method	Details	Limitations	Selected references
Indicator species	Based upon the presence/absence of individual taxa. A climatic range can be defined using the overlap of taxa within a single assemblage.	Pollen morphology cannot always distinguish to species level.	Iversen, 1944
Biomisation	Plant Functional Types (PFTs) are constructed using fossil pollen assemblages based upon the habitat preferences and biological function. These attributes along with present distribution of these types are used to infer past climatic variability by combining them to create biomes. Comparison with modern data can be used to infer changes in climate. More recently, process-based vegetation models such as BIOME3 and BIOME4 have been developed.	Defining the affinity cores to identify the PFT for each sample can be problematic. Also, a single sample can be applied to multiple PFTs, especially where the habitat tolerance of a species is wide or at the margins between two biomes.	Prentice et al, 1992; 1996; Edwards, 2007; Kaplan et al., 2003.
Modern Analogue Technique (MAT), Response surfaces	To identify modern environments analogous to those of the past, dissimilarity matrices are used to compare fossil pollen samples and those taken from a range of contemporary environments. The data used for the comparisons are pollen proportions from the samples which can be either the full range of taxa in each sample or a selected group of taxa e.g. representative proportion or main taxa. For those modern samples selected as suitable analogues, supplemental environmental data, e.g. climate, soils, vegetation structure, can be extrapolated to aid interpretation of the relevant fossil horizons.	Non-analogue situations can be encountered and limit the information that can be gained. These are frequently seen in studies that address Late-Glacial environments and those that have been subjected to intensive human interference. Error can be induced into the results if the size of the sampling locations for the fossil and modern sites are not harmonised. The climatic regions from which the modern samples are collected can be too large, such that some variation identified may be statistical artefacts.	Overpeck et al, 1985; Gavin, 2003; Wahl, 2004; Jackson and Williams, 2004; Bartlein et al, 1986; Cheddadi et al, 1998
Transfer function	A modern pollen dataset is used to derive a correlation, e.g. habitat range, of individual taxa for specific climatic variables. Using these known relationships, multivariate analysis is used to extrapolate this information to pollen data in fossil assemblages, e.g. WA, WA-PLS to reconstruct temperature and precipitation regimes.	The analysis is frequently based upon linear correlation between the pollen type and a climatic variable. Mostly, this is not the case and responses are unimodal so the taxa included and suitability of analysis applied can influence the interpretation of the results. Also, removing noise from the overall climate signal can be complex, but by transformation of the pollen data prior to the analysis the influence of this can be minimised. An edge effect can affect the reconstructions; this occurs when a taxon is present at the limit of its range or representation in the modern data.	Seppa and Birks, 2001; Seppa et al 2004; Seppa and Bennett, 2003; Birks and Seppa, 2004

Table 1: A summary of quantitative analytical methods that can be applied to pollen data

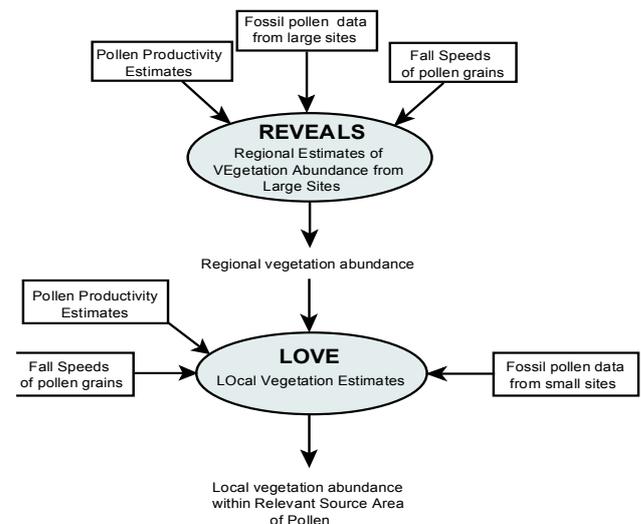
General limitations have been raised by numerous authors (MacDonald and Edwards, 1991; Seppa and Bennett, 2003). Firstly, the relationship between pollen and climate is not straightforward and rarely in equilibrium as a linear correlation between pollen and a single climatic factor, while biological influences are not accounted for. Although a number of studies have undertaken climate reconstructions based on pollen (Cheddadi et al., 1998; Peyron et al., 1998), critical evaluation of the reconstructed parameter is required to establish its value. Secondly, pollen morphology and most importantly species distinction, has models assume. Many of the reconstruction methods assume a linear or unimodal response to the climatic factor, being calculated - e.g. precipitation or summer temperature, but it can be difficult to ascertain a role in the performance of all the approaches. Most pollen types can be identified to genus level, but frequently species can occupy different habitats and climatic zones (e.g. species of *Juniperus* and *Betula*). When such taxa are present, the predictive capacity of the analytical method is compromised, especially when dealing with Late-Glacial sites because the climatic envelope of some taxa may not be strictly confined to low temperature regimes e.g. *Betula*. Finally, assessing the validity of the model output is critical to evaluating the reliability of the results obtained. Uncertainty estimators and validation tools can be used to assess the model performance, but assessment of the latter is complex and limits the value that can be placed upon the results obtained (Birks and Seppa, 2004). Investment in Bayesian methods to improve error estimations, spatial and multiproxy representation of climate from pollen data could provide a solution and enhance interpretations (Ohlwein and Wahl, 2012).

Vegetation data

One of the recent developments in pollen analysis is the availability of quantitative modelling software. Arising from the understanding that pollen production and dispersal is not even between species, mechanistic models provide a correction for this bias allowing a more faithful account of changes within the vegetation patterns. Currently, two models are available which are

constructed using very similar algorithms, but different approaches.

The Landscape Reconstruction Algorithm (LRA: Sugita, 2007a, b) uses pollen counts from both large and small lakes to calculate vegetation coverage within defined spatial extent (Figure 5). As a pollen sample from a lake is composed of pollen recruited from a regional (~100km radius) and locally derived areas (~10km) (Figure 2), the LRA uses sites from large lakes to estimate regional vegetation coverage using REVEALS, the first stage of the process. By defining the influence of regional vegetation, it is possible to obtain a purely local signal from smaller sites used in the second stage of the model, LOVE. Restrictions on the spatial extent represented are controlled by the size of the sites used within the model and this provides the freedom to generate data specific to the hypothesis being addressed. By isolating these spatial scales, changes become more visible leading to more informed interpretations. Validation studies have been conducted using this software, showing the potential for improving understanding of past landscape changes (Soepboer et al., 2010; Sugita et al., 2010) along with the necessary cautions and considerations that must be made when using such models (Hellman et al., 2007).



Sugita, 2007a, b

Figure 5: Components of the Landscape Reconstruction Model (Sugita, 2007 a, b).

The Multiple Scenario Approach (Bunting and Middleton, 2009) creates images of past landscapes and is a highly informative tool, especially for archaeological applications (Bunting et al., 2007). Incorporating elements of MAT, numerous grids (hypothetical landscapes) are created from which pollen assemblages at target site(s) can be generated. Using dissimilarity coefficients, the most probable landscapes can be selected when compared to actual pollen data from the target site, and used as scenarios to address the original hypotheses. When using multiple target sites, the model can be used to construct potential spatial patterns of past landscapes (Bunting et al., 2007; Pratt, 2007).

Although the output generated differs between the models, there are *a priori* specifications that have to be considered for both approaches. Firstly, the models require parameter data in the form of fall speeds of pollen grains, atmospheric properties (standard data are often used) and pollen productivity estimates. Studies have shown that changes in these components can be significant to model output, especially with productivity estimates (Bunting et al., 2004; Neilsen and Sugita, 2005). Secondly, a robust chronology is essential to ensure that when multiple sites are used, the equivalent horizons are selected.

Conclusion

Pollen is often considered a qualitative tool that has limits in its interpretative capacity, but this is far from the truth. This paper has tried to show that consideration of specific project aims in light of the strengths and limitations of pollen analysis can yield highly informative quantitative data, far from the descriptive results often sought from pollen analysis. Development of analytical tools and theoretical understanding is encouraging pollen data to be reconsidered as a quantitative technique which can be effectively applied to address key debates. Where this has been done, the results show that pollen analysis is far from a descriptive tool, and further research will help pollen archives reach their full potential.

References

- Anderson, N.J., Bugmann, H., Dearing, J.A. and Gaillard, M.J., 2006. Linking palaeoenvironmental data and models to understand the past and predict the future. *TRENDS in Ecology and Evolution*, **21**: 696 - 704.
- Bennett, K.D., 1994. Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences. *The Holocene*, **4**: 337 - 348.
- Birks, H.J.B., 1996. Contributions of Quaternary Palaeoecology to Nature Conservation. *Journal of Vegetation Science*, **7**: 89-95.
- Birks, H.J.B. and Birks, H.H., 1980. Quaternary Palaeoecology. Edward Arnold, London, 289 pp.
- Birks, H.J.B. and Seppa, H., 2004. Pollen-based reconstructions of late-Quaternary climate in Europe - progress, problems, and pitfalls. *Acta Palaeobotanica*, **44**: 317 - 334.
- Bonny, A.P., 1978. The effect of pollen recruitment processes on pollen distribution over the sediment surface of a small lake in Cumbria. *Journal of Ecology*, **66**: 385 - 416.
- Broström, A. et al., 2008. Pollen productivity estimates of key European plant taxa for quantitative reconstruction of past vegetation: a review. *Vegetation History and Archaeobotany*, **17**: 461.
- Brown, A., 1997. Alluvial geoarchaeology: floodplain archaeology and environmental change. Cambridge University Press, Cambridge.
- Brown, A.G., 1999. Biodiversity and pollen analysis: modern pollen studies and the recent history of a floodplain woodland in S-W-Ireland. *Journal of Biogeography*, **26**: 19-32.
- Bunting, J., Middleton, R. and Twiddle, C.L., 2007. Wetland records of a coastal cultural landscape in north-west Scotland: an application of the Multiple Scenario Approach to landscape reconstruction. In: B. Barber et al. (Editors), *Archaeology from the wetlands: recent perspectives*. WARP Occasional Paper 18, Society of Antiquaries of Scotland, pp. 109 - 117.

- Bunting, M.J., Gaillard, M.J., Sugita, S., Middleton, R. and Broström, A., 2004. Vegetation structure and pollen source area. *The Holocene*, **14**: 651-660.
- Bunting, M.J. and Middleton, R., 2009. Equifinality and uncertainty in the interpretation of pollen data: the Multiple Scenario Approach to reconstruction of past vegetation mosaics. *The Holocene*, **19**: 799.
- Caseldine, C.J., McGarry, S.F., Baker, A., Hawkesworth, C. and Smart, P.L., 2008. Late Quaternary speleothem pollen in the British Isles. *Journal of Quaternary Science*, **23**: 193 - 200.
- Cheddadi, R., Lamb, H.F., Guiot, J. and van der Kaars, S., 1998. Holocene climatic change in Morocco: a quantitative reconstruction from pollen data. *Climate Dynamics*, **14**: 883 - 890.
- Clymo, R.S. and Mackay, D., 1987. Upwash and downwash of pollen and spores in the unsaturated surface layer of *Sphagnum*-dominated peat. *New Phytologist*, **105**: 175 - 183.
- Cushing, E.J., 1967. Evidence for differential pollen preservation in late Quaternary sediments in Minnesota. *Review of Palaeobotany & Palynology*, **4**: 87 - 101.
- Davis, M.B., 1967. Pollen accumulation rates at Rogers Lake, Connecticut, during Late- and Postglacial time. *Review of Palaeobotany & Palynology*, **2**: 219 - 230.
- Davis, M.B., 1968. Pollen Grains in Lake Sediments: Redeposition Caused by Seasonal Water Circulation. *Science*, **162**: 796.
- Davis, M.B., 1969. Climatic changes in Southern Connecticut recorded by pollen deposition at Rogers Lake. *Ecology*, **50**: 409 - 422.
- Davis, M.B., 1973. Redeposition of Pollen Grains in Lake Sediment. *Limnology and Oceanography*, **18**: 44.
- Davis, M.B. and Brubaker, L.B., 1973. Differential Sedimentation of Pollen Grains in Lakes. *Limnology and Oceanography*, **18**: 635.
- Davis, M.B. and Deevey, E.S., 1964. Pollen Accumulation Rates: Estimates from Late-Glacial sediment of Rogers Lake. *Science*, **145**: 1293 - 1295.
- Davis, M.B., Moeller, R.E. and Ford, J., 1984. Sediment focusing and pollen influx. In: E.Y. Haworth and J.W.G. Lund (Editors), *Lake sediments and environmental history*. Leicester University Press, Leicester, pp. 261 - 293.
- Dimbleby, G.W., 1985. The palynology of archaeological sites. Academic Press, 176 pp.
- Edwards, K.J., 1983. Quaternary palynology: multiple profile studies and pollen variability. *Progress In Physical Geography*, **7**: 587 - 609.
- Edwards, K.J., 2004. People, environmental impacts and the changing landscapes of Neolithic and early Bronze Age times. In: I.A. Shepherd and G.J. Barclay (Editors), *Scotland in ancient Europe*. Society of Antiquities of Scotland, Edinburgh, pp. 55 - 71.
- Edwards, K.J. and MacDonald, G.M., 1991. Holocene palynology: II human influence and vegetation change. *Progress in Physical Geography*, **15**: 364 - 391.
- Edwards, M.E., 2007. BIOME model of vegetation reconstruction. In: S. Elias (Editor), *Encyclopaedia of Quaternary Science*. Elsevier, pp. 2551 - 2561.
- Gaillard, M.J. et al., 2010. Holocene land-cover reconstructions for studies on land cover-climate feedbacks. *Climate of the Past*, **6**: 483 - 499.
- Giesecke, T. and Fontana, S.L., 2008. Revisiting pollen accumulation rates from Swedish lake sediments. *The Holocene*, **18**: 293 - 305.
- Havinga, A.J., 1984. A 20 year experimental investigation into the differential corrosion susceptibility of pollen and spores in various soil types. *Pollen et Spores*, **26**: 541 - 558.
- Hellman, S., Gaillard, M.-J., Broström, A. and Sugita, S., 2007. The REVEALS model, a new tool to estimate past regional plant abundance from pollen data in large lakes: validation in southern Sweden. *Journal of Quaternary Science*, **23**: 21 - 41.

- Hicks, S., 2001. The use of annual arboreal pollen deposition values for delimiting tree-lines in the landscape and exploring models of pollen dispersal. *Review of Palaeobotany and Palynology*, **117**: 1.
- Hyvärinen, H., 1975. Absolute and relative pollen diagrams from northernmost Fennoscandia. *Fennia*, **142**: 1 - 23.
- Iversen, J., 1944. *Viscum, Hedera and Ilex* as climatic indicators. *Geologiska Föreningens Förhandlingar*, **66**: 463 - 483.
- Jacobson, G.L. and Bradshaw, R., 1981. The selection of sites for palaeoenvironmental studies. *Quaternary Research*, **16**: 80 - 96.
- Kaplan, J.O. et al., 2003. Climate change and arctic ecosystems II: modelling palaeodata-model comparisons and future projections. *Journal of Geophysical Research*, **108**: 1 - 17.
- Lehman, J.T., 1975. Reconstructing the rate of accumulation of lake sediments: The effects of sediment focusing. *Quaternary Research*, **5**: 541 - 550.
- MacDonald, G.M. and Edwards, K.J., 1991. Holocene palynology 1: principles, population and community ecology, palaeoclimatology. *Progress in Physical Geography*, **15**: 261 - 289.
- McGlone, M. A., 2001. A late Quaternary pollen record from marine core P69, southeastern North Island, New Zealand, *New Zealand Journal of Geology and Geophysics*, **44**, 69 - 77
- Nielsen, A.B. and Sugita, S., 2005. Estimating relevant source area of pollen for small Danish lakes around AD 1800. *The Holocene*, **15**: 1006-1020.
- Ohlwein, C. and Wahl, E.R., 2012. Review of probabilistic pollen-climate transfer methods. *Quaternary Science Reviews*, **31**: 17 - 29.
- Overpeck, J.T., Webb III, J.A. and Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: Dissimilarity coefficients and the method of modern analogues. *Quaternary Research*, **23**: 87 - 109.
- Peck, R.M., 1973. Pollen budget studies in a small Yorkshire catchment. In: H.J.B. Birks and R.G. West (Editors), *Quaternary Plant Ecology*. Blackwell, pp. 43 - 60.
- Pennington, W. and Tutin, T.G., 1996. Limnic sediments and the taphonomy of Lateglacial pollen assemblages. *Quaternary Science Reviews*, **15**: 501.
- Peyron, O. et al., 1998. Climatic reconstruction in Europe for 18,000 YR B.P. from pollen data. *Quaternary Research*, **49**: 183 - 196.
- Pratt, S., 2007. Long term landscape dynamics in Caledonian pine forest. University of Edinburgh, Unpublished PhD thesis.
- Prentice, I.C., Guiot, J., Huntley, B., Jolly, D. and Cheddadi, R., 1996. Reconstructing biomes from palaeoecological data: A general method and its application to European pollen data at 0 and 6 ka. *Climate Dynamics*, **12**: 185.
- Seppa, H. and Bennett, K.D., 2003. Quaternary pollen analysis: recent progress in palaeoecology and palaeoclimatology. *Progress in Physical Geography*, **27**: 548 - 580.
- Seppa, H. and Hicks, S., 2006. Integration of modern and past pollen accumulation rate (PAR) records across the arctic tree-line: a method for more precise vegetation reconstructions. *Quaternary Science Reviews*, **25**: 1501.
- Soepboer, W., Sugita, S. and Lotter, A.F., 2010. Regional vegetation-cover changes on the Swiss Plateau during the past two millennia: A pollen-based reconstruction using the REVEALS model. *Quaternary Science Reviews*, **29**: 472-483.
- Sugita, S., 1994. Pollen Representation of Vegetation In Quaternary Sediments - Theory And Method In Patchy Vegetation. *Journal of Ecology*, **82**: 881-897.
- Sugita, S., 2007a. Theory of quantitative reconstruction of vegetation I. Pollen from large lakes reveals regional vegetation. *The Holocene*, **17**: 229 - 241.
- Sugita, S., 2007b. Theory of quantitative reconstruction of vegetation II. All you need is LOVE. *The Holocene*, **17**: 243 - 257.
- Sugita, S., Parshall, T., Calcote, R. and Walker, K., 2010. Testing the Landscape Reconstruction Algorithm for spatially explicit reconstruction of vegetation in northern

Michigan and Wisconsin. *Quaternary Research*, **74**: 289 - 300.

Twiddle, J.C. and Edwards, K.J., 2010. Pollen preservation zones as an interpretative tool in Holocene palynology. *Review of Palaeobotany and Palynology*, **161**: 59 - 76.

Twiddle, C.L. and Bunting, J., 2010. Experimental investigations into the preservation of pollen grains. *Review of Palaeobotany and Palynology*, **162**: 621 - 630.

von Post, L., 1916. Skosträdpollen i Sydsvenska Tormosselagerföljder. *Forhandlingar Vid*, **16**: 434 - 465.

Willis, K.J. et al., 2007. How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **362**: 175-186.

Willis, K.J. and Birks, H.J.B., 2006. What is natural? The need for a long-term perspective in biodiversity conservation. *Science*, **314**: 1261-1265.

Wilmshurst, J.M and McGlone, M.S., 2005. Corroded Pollen and Spores as Indicators of Changing Lake Sediment Sources and Catchment Disturbance [*Journal of Paleolimnology*](#), **34**, 503-517.

[Wilmshurst](#), J. M, [Eden](#), D.N, and Froggit, P.C., 1999, Late Holocene forest disturbance in Gisborne, New Zealand: A comparison of terrestrial and marine pollen records New Zealand, *Journal of Botany*, **37**, 523 - 540

Van Campo, E., Duplessy, J. C, and Rossignol-Strick, M., 1982. Climatic conditions deduced from a 150-kyr oxygen isotope-pollen record from the Arabian Sea, *Nature*, **296**, 56-59

Luminescence Dating

Claire L. Mellett¹

¹ School of Environmental Science, University of Liverpool (mellettc@liverpool.ac.uk)



ABSTRACT: Luminescence dating is a unique chronometric tool as it dates sediments and landforms directly by establishing the time elapsed since grains were last exposed to daylight. Sediments from various terrestrial and shallow marine environments, ranging from a few years to over one hundred thousand years in age can be dated using luminescence techniques. Over the last decade, methodological and technological developments have improved the reliability and precision of the technique, thus encouraging widespread application in a variety of environments and helping resolve questions regarding the timing and rate of geomorphic processes. The field of luminescence is continually advancing and ongoing research focuses on extending the age range and exploring new applications. This paper outlines the principles behind luminescence dating and introduces the current and most widely applied methodological approach to using luminescence dating in geomorphology.

KEYWORDS: luminescence, OSL, chronology, Quaternary, environmental change, sediments

Introduction

Constraining the age of sediments and landforms and determining the sequence of events driving landscape development is fundamental in the study of geomorphology and allows geomorphologists to test theoretical models and quantify rates of landscape change. A number of dating tools targeting various materials and spanning different age ranges are available to support geomorphological research (refer to other chapters in this edition). Of these, radiocarbon dating (^{14}C) is historically the most commonly used technique. However, ^{14}C dating is limited to environments where organic matter is preserved in the sediment and as many geomorphic systems are predominantly siliciclastic, its use is often restricted. As an alternative, Luminescence dating uses minerals that are common in most environments making the technique widely applicable in variety of landscapes. Further, it is particularly important in geomorphology as it dates sediments and landforms directly and gives numerical ages without the need for independent age control (see Lian and Roberts, 2006).

This paper provides a brief introduction into the basic principles that allow age determination using luminescence, focusing on common laboratory practices and data analysis techniques in Optically Stimulated Luminescence (OSL) dating. It also acts a guide to assist geomorphologists in selecting the most suitable depositional settings, site locations and sampling methodologies. Overall this technical guide is intended to provide readers with an informed platform of knowledge from which they are able to acquire more information from the published literature or approach individual laboratories for further advice.

Basic Principles

Luminescence dating determines the age since burial of sediment by measuring the total amount of stored signal resulting from exposure of the sediment to a known annual dose of background radiation. The concept of luminescence dating relies on defects in the crystal lattice of dosimeter minerals, most commonly quartz and feldspar, to trap energy produced during the interaction between

environment. The lower age limit of luminescence dating is in the range of years (see Madsen and Murray, 2009 for a review), e.g. 10 ± 3 a for modern dunes in Denmark (Madsen et al., 2007). The upper age limit depends on the dose required to saturate the luminescence signal which is a function of the minerals capacity to store charge and the dose-rate received. Routine OSL dating of quartz can give ages up to 150 ka (Stokes, 1999). Where the dose-rate is low and/or the saturation dose is high, there is the potential to extend this upper age limit towards hundreds of thousands of years (e.g. 107.8 ± 5.2 ka: Mellett et al., 2012, and 507 ± 41 ka: Watanuki et al., 2005). Quartz typically saturates at a much lower dose than feldspar. However, the dose-rate of potassium-rich feldspars can be high due to internal dose contribution from the decay of radioactive isotopes within the crystal lattice (Adamic and Aitken 1998). Given the variability outlined above, an age range for luminescence dating is difficult to define and often only reveals itself when samples are being analysed in the laboratory.

Sample selection

Choosing the right sediments

The most suitable sediments for luminescence dating are those that have been exposed to sufficient daylight to enable bleaching and resetting of the luminescence signal. This can be achieved in a variety of terrestrial and shallow water environments making luminescence dating one of the most widely applicable chronometric tools used in geomorphology. The first principles of OSL dating were developed using aeolian sediments from low latitude desert environments where insolation levels are high and there is increased likelihood that sediments will undergo repeated erosion-deposition cycles (see Wintle, 1993 for a review). However, improved understanding of mineral characteristics and bleaching regimes in a variety of sedimentary environments (see Rhodes, 2011 pp 468-470) means it is now possible to produce reliable luminescence ages from sediments deposited in fluvial, coastal and shallow marine, glaciofluvial, periglacial and colluvial environments (see Wallinga, 2002; Jacobs, 2008; Thrasher et al., 2009a; Bateman, 2008; Fuchs and Lang, 2009 for reviews

respectively). When making decisions about the most suitable sediments for luminescence dating it is vital to have an independent understanding of depositional history obtained through lithological, stratigraphic and morphological analyses.

Sand or silt sized sediments are required for luminescence dating. Grain size fractions 4-11 μm (silt) and ~63-250 μm (sand) should be targeted to account for differences in the laboratory dose-rate received by a sample according to grain size (Aitken, 1985). Further, it is advisable to avoid sediments that are rich in organic matter or carbonates as they may have received heterogeneous dose-rates.

Quartz vs feldspar?

When quartz or feldspar minerals are exposed to light the luminescence signal is released over a short period of time (typically 10s of seconds). The amount of time required to remove this signal and the style of signal decay is mineral dependent (Fig. 2). Quartz bleaches more easily than feldspar when exposed to natural light (Godfrey-Smith et al., 1988). Therefore, quartz is the preferred mineral when dating environments where insufficient bleaching, and hence age overestimation, may present a problem (e.g. Wallinga et al., 2001; Jain et al., 2004). In environments where quartz has been eroded only recently, low sensitivity can hinder reliable age estimates (Sawakuchi et al., 2011) and feldspar may be better suited. However, feldspars are vulnerable to the effect of anomalous fading (Wintle, 1973) and can underestimate ages as a result. Methodological advancements such as the development of single-grain OSL dating (Duller, 2008a) have improved the reliability of ages obtained from environments where incomplete bleaching is a problem. However, it is pertinent to consider depositional history and local mineralogy when selecting the most suitable mineral for luminescence dating.

Sampling methodology

The number and location of samples for luminescence dating is governed by the geomorphic problem that is being addressed and is therefore site specific. Despite this, there are a number of criteria that must be

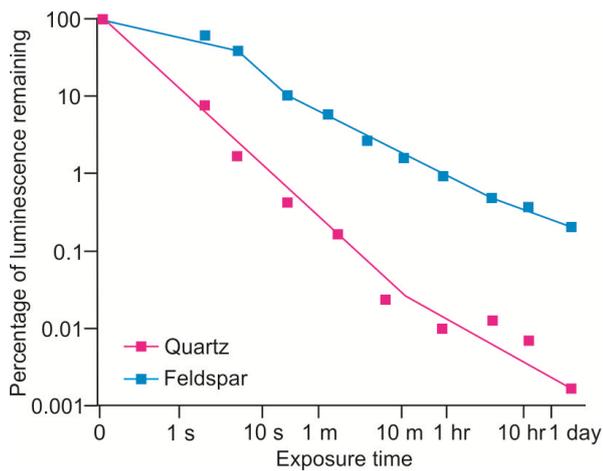


Figure 2: Reduction of Quartz and Feldspar luminescence signals during exposure to light (redrawn from Duller et al., 2008b). Note the luminescence is released much faster in quartz when compared to feldspar.

met to ensure luminescence ages from sediments are reliable for dating geomorphic events or processes. Firstly, it is important to take a sample from the centre of a lithofacies or stratigraphic unit (e.g. A and B Fig. 3b.) to avoid heterogeneity in the dose rate (e.g. C Fig. 3b). Disturbed ground that shows evidence of bioturbation or water leaching (e.g. D Fig. 3b) should be avoided to ensure no post-depositional mixing or significant changes in water content as these can influence the accuracy of the age produced. Further, an understanding of the sediment composition surrounding the sample is vital for determining dose rate. Therefore, sampling at the base of sections (or cores) where the underlying geology is unknown should be avoided (e.g. E Fig. 3b). To test the reliability and chronostratigraphic significance of ages where no independent chronological tool is available (e.g. radiocarbon dating), multiple samples within a lithofacies or stratigraphic unit (e.g. B Fig. 3b) may prove useful.

When sampling from cores the above criteria also apply. However, further care must be employed to avoid sampling sediment in close proximity (approximately 1 cm) to the core barrel or casing as this may have been disturbed during the coring process (Fig. 4a). It is advisable to avoid areas in the core where there is visible evidence of disturbance.

After choosing sampling locations, the most important methodological constraint is to recover the sediment and transport it to the laboratory without exposure to daylight. If sampling using coring apparatus, this can be achieved using opaque core liners and subsequent sub-sampling under safe light conditions in the laboratory whilst ensuring exposed sediment from the top and bottom of the core is discarded. When sampling exposed sediment sections (e.g. Fig 3a), an opaque PVC or metal cylinder is hammered into the face of the exposed section until it is completely filled with sediment (N.B. the diameter and length of cylinder depends on the lithofacies being sampled) (Fig. 4b). The cylinder is then excavated and sealed with black opaque plastic liners to ensure minimum exposure to daylight and transported to the laboratory. In some cases, to avoid unnecessary exposure to daylight an opaque fabric or liner can be draped over the sampling location whilst the sediment is being extracted (Fig. 4c).

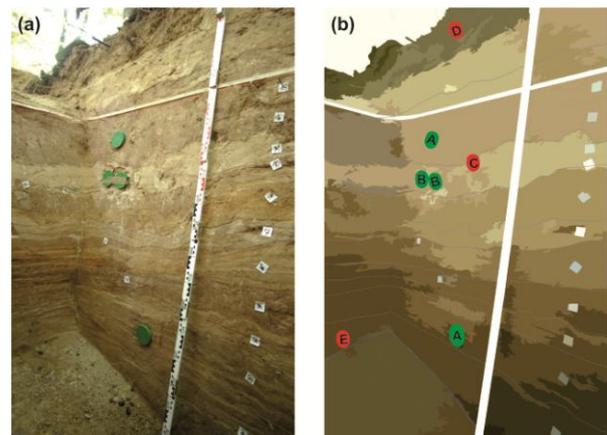


Figure 3: (a) Luminescence sampling of an excavated stratigraphic section. Photograph courtesy of Andreas Lang. (b) schematic stratigraphic section showing ideal sample locations (green) and unsuitable sample locations (red) as discussed in the main text.

The above section outlines procedures for sampling material for D_e determination (one half of the age equation). An additional bulk sample is required to determine dose-rate if in situ gamma spectrometry is not being undertaken (see the next section). This sample can be excavated whilst exposed to daylight (either from the exposed section or a core) and should be taken from the area directly surrounding the original sample to ensure an accurate representation of

environmental dose. Further, a sub-sample of sediment should be extracted to determine water content (an important component of the dose-rate calculation). These samples must be stored in airtight bags and are best kept refrigerated to ensure limited evaporation.

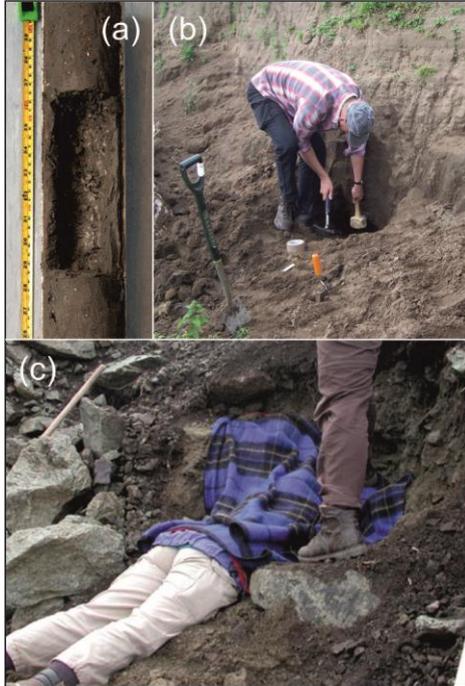


Figure 4: (a) An example of a luminescence sample taken from a core. (b) Sampling an exposed section by hammering a cylinder into the face. Photograph courtesy of Richard Chiverrell. (c) An opaque fabric or liner can be used to minimise exposure to daylight. Photograph courtesy of Barbara Mauz.

Dose-rate determination

To obtain an age using luminescence dating, the age equation (Eq. 1.) requires a measure of the radiation received by the sample per year i.e. the dose rate. There are two major sources of ionising radiation in the environment that need to be determined; (i) the radioactive decay of radionuclides U, Th and K and their daughter nuclides, and; (ii) cosmic rays.

Annual dose received from radionuclides

This can be achieved by either; (i) measuring the concentration of radioactive elements in the sediment then calculating the dose-rate using the conversion factors outlined in Adamiec and Aitken (1998), or; (ii) measuring the radioactivity directly by counting the

emissions of alpha, beta and gamma using field or laboratory spectrometers. Both methodologies assume the dose-rate has not changed significantly since burial. See Duller (2008b) for a more thorough review of these techniques.

Measuring dose-rate in the laboratory as outlined above does not account for the presence of water within sediment pore spaces which absorbs radiation. To correct for this effect, water content needs to be determined and incorporated into the dose-rate calculation (Aitken, 1998). To account for fluctuations in water content that may have occurred during burial, an understanding of geological history (e.g. changes in climate or water table) is required to constrain the variability (e.g. Mellett et al., 2012). Finally, it is also important to assess if disequilibrium in the U decay series is present (see Olley et al., 1996) to ensure high degrees of accuracy are met when determining dose rate.

Annual dose received from cosmic rays

The dose received from cosmic rays is relatively low when compared to that received from radionuclides. However, in low dose environments and at high altitudes and latitudes, the cosmic dose contribution can be significant. Cosmic dose varies with longitude, latitude, altitude and burial depth and can be calculated using the equations outlined in Prescott and Hutton (1994).

Equivalent dose (D_e) determination

Determining equivalent dose (D_e) is arguably the most time consuming stage in luminescence dating. At this point in the dating process geomorphologists should be working in close collaboration with a luminescence laboratory that will provide advice and guidance about independent laboratory practises and the most suitable techniques according to individual samples. This section gives a brief introduction into conventional procedures used to determine D_e .

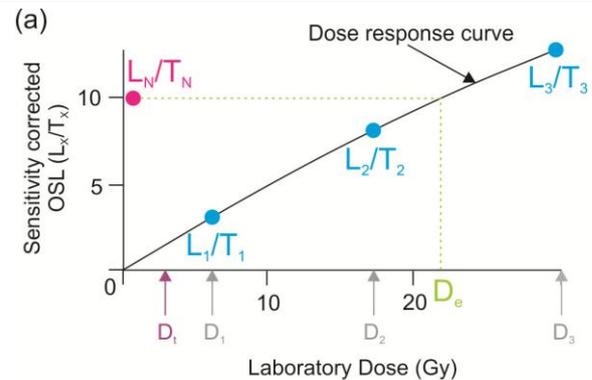
Extracting quartz or feldspar

The procedures for extracting quartz or feldspar from a bulk sample will depend on the individual laboratory. Typically, samples

are separated to the desired grain size (sand or silt) then organic matter and calcium carbonate are removed using H_2O_2 and HCl . Quartz, feldspar and heavy minerals are separated according to density using heavy liquid or magnetic procedures. If using quartz for OSL dating, the resulting fraction is etched using HF (to remove the outer part of the grain affected by alpha radiation). The above are carried out under subdued red-light conditions.

The single-aliquot-regenerative-dose (SAR) protocol

D_e represents the amount of dose accumulated in the environment since the last exposure of the mineral to daylight. The SAR protocol (Murray and Roberts, 1998; Murray and Wintle, 2000) is a procedure to measure D_e whilst accounting for changes in the mineral (sensitivity change) that can occur due to irradiation, heating or light stimulation in the laboratory. The SAR protocol is widely used to determine D_e in quartz (Wintle and Murray 2006). The protocol measures the natural luminescence (L_N) produced by the mineral when stimulated by light (i.e. it completely bleaches the signal) and then effectively tries to recreate its dose history by giving incremental doses of radiation (Dose 1, Dose 2, Dose 3 Fig. 5) and regenerating the luminescence accumulated. After each given dose, the luminescence is measured (L_1 L_2 L_3 Fig. 5). With the exception of the first cycle (i.e. L_N), each time luminescence is measured during the SAR protocol, it is corrected for sensitivity changes by taking the ratio of the luminescence signal (L_x) to the response to a fixed test dose (T_x). A dose response curve (Fig. 5) is constructed by plotting the sensitivity corrected luminescence (e.g. L_x/T_x). D_e is determined at the point on the X axis where the natural luminescence (L_N) intercepts the dose response curve (Fig. 5). Determining ages using the SAR protocol is considered both accurate and precise when compared to independent age controls (e.g. Rittenour, 2008; Rhodes, 2011). However, it is important to ensure the assumptions inherent to the protocol (Murray and Wintle 2000; Wintle and Murray 2006) are met which can be achieved by carrying out a number of tests in the laboratory (see Duller et al., 2008b for a review).



(b)

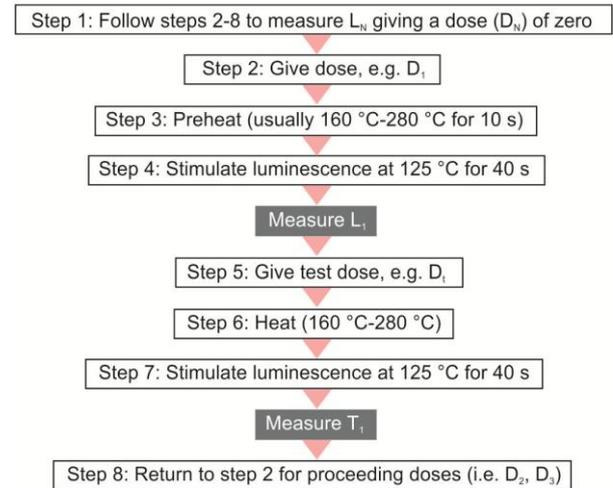


Figure 5: (a) Determination of D_e using the SAR protocol where the natural dose (L_N) and regenerative dose (L_x) are corrected for sensitivity changes using a small test dose (T_x). (b) Generalized quartz SAR protocol (after Murray and Wintle, 2000).

Single aliquot or single grain?

An aliquot is a subsample of grains taken from the bulk sample. In luminescence dating aliquots are mounted on 1 cm diameter discs for measurement. The number of grains present on these discs depends on the area of the aliquot covered (e.g. 1 mm to 10 mm) and the grain size of the subsample (Duller et al., 2008a). For example, a 1 mm aliquot of grain size 100 μm will comprise on average 10 individual grains whilst a 1 mm aliquot of grain size 300 μm will hold only 1 or 2 grains (Duller, 2008a). When using single aliquots, D_e is the sum of the luminescence emitted by all grains present on the disc (Fig. 6a). However, not all grains within the aliquot produce the same, if any, luminescence in response to the same dose (e.g. Duller et al., 2000), thus reducing the precision of D_e values. Therefore, understanding the number of grains present within an aliquot and how

many of those grains produce luminescence is important.

Measurement of multiple single aliquots can result in a large spread (overdispersion) of D_e values (Fig. 6b) which can be a result of (i) variations in the degree of bleaching, or; (ii) mixing of grains during or post-deposition. Typically, using the SAR protocol large numbers of aliquots are measured and the D_e is determined using a variety of statistical analyses (see the section below). However, in some cases the spread in D_e values is too large due to poor bleaching and single-grain techniques are better suited (e.g. Olley et al., 2004; Duller, 2006).

Advancements in instrumentation have made it feasible to measure the OSL signal from single grains thus demonstrating the variability in dose distribution between grains from the same sample. This technique is most suitable for samples where exposure of grains to daylight is variable (e.g. fluvial and glaciogenic). However, single grain measurements are time intensive and typically <5% of grains dominate the luminescence signal giving results that are similar to those obtained through single aliquots (Duller, 2008b). The decision to use small aliquots or single grains is sample specific and should be undertaken in collaboration with a laboratory.

Statistical analyses

There is typically a degree of variability in the D_e values from different aliquots or different grains within the same sample. A number of statistical methods exist to identify a dominant component related to the last bleaching event within complex distributions (e.g. Fig. 6c) and hence improve the reliability of age estimates. These include the Central Age Model (CAM) and The Minimum Age Model (MAM) (Galbraith et al., 1999), and the Finite Mixture Model (Galbraith and Green, 1990). The most appropriate age model is sample dependent and can be assessed according to a number of statistical criteria (see Bailey and Arnold, 2006; Rodnight et al., 2006; Rowan et al., 2012; Thrasher et al., 2009b for examples). However, these criteria are not inclusive and statistical treatment can be governed by the geological context of individual samples (Galbraith and Roberts, 2012). After statistical assessment of D_e

distributions, the most appropriate D_e has been established and is available for input into the age equation (Eq. 1).

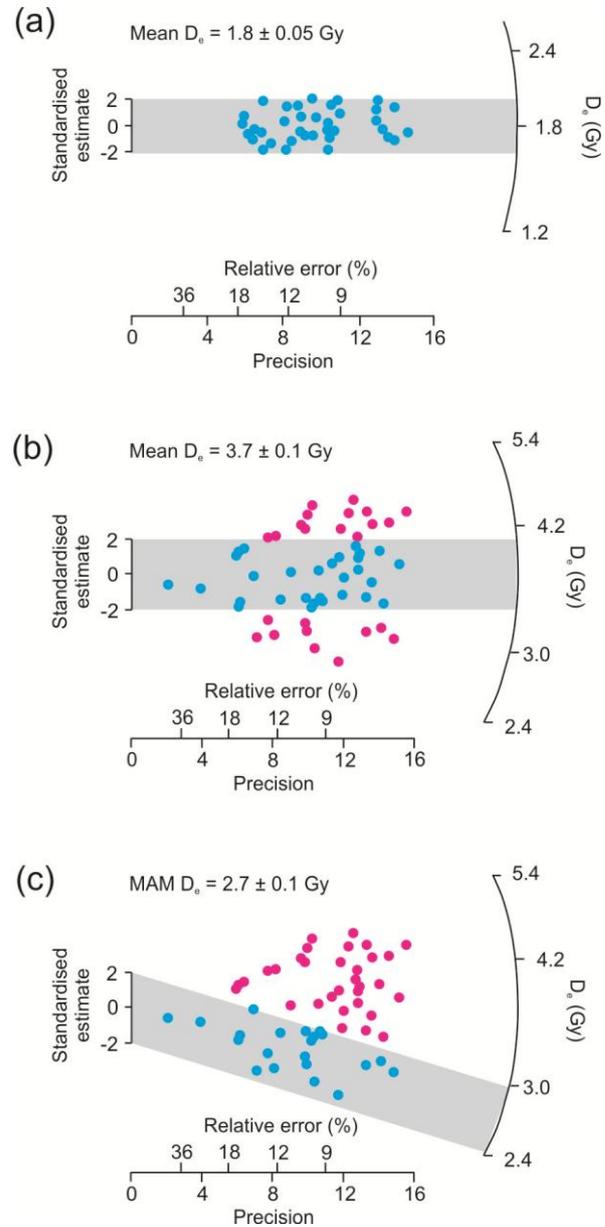


Figure 6: Radial plots of D_e values determined using SAR of quartz (modified from Mauz et al., 2010). Refer to Galbraith and Roberts (2012) for an explanation on how to interpret radial plots. (a) D_e is determined by taking the arithmetic mean of values shown in blue, those in pink are rejected. (b) For this sample, due to a broad distribution of values the arithmetic mean is not an appropriate representation of D_e (c) Statistical analysis of the distribution shown in (b) led to the application of the Minimum Age Model (MAM) to determine D_e for the age calculation. Note the difference in D_e determined from the same sample using arithmetic mean (b) and MAM (c).

Recent advances and applications

Age range extension

Age range extension, particularly the upper limit has been a priority amongst researchers over recent years. The development of experimental procedures that remove the effect of anomalous fading in feldspar minerals (post IR-IRSL) (Thomsen et al., 2008; Buylaert et al., 2009) means it may be possible to target the luminescence signal corresponding to the dose received during burial without incorporating less stable components of the signal. As the saturation level of feldspar is much higher than quartz, this methodological development is pushing the upper limit with the potential to extend the age range by a factor of 4-5 (Buylaert et al., 2012).

Dating sediments from challenging environments

The range of environments that can be dated using luminescence dating and examples of the successful application of the technique are extensive (see Rhodes, 2011 for a summary). Aeolian and fluvial sediments are commonly dated as bleaching regimes in these environments are relatively well understood (e.g. Fujioka et al., 2009; Wallinga et al., 2010). However, in many marine, coastal, colluvial and glacial environments complete zeroing of the luminescence signal can be problematic. Despite this, ongoing developments in luminescence dating techniques are allowing the production of reliable ages from these more challenging environments (e.g. Murari et al., 2007; Mauz et al., 2010; Thrasher et al., 2009b).

Dating gravel rich sediment is problematic as it is difficult to confidently estimate dose-rate in heterogeneous sediments. However, it has been demonstrated that when gravels are dominated by inert lithologies (such as chalk), sand within the gravel matrix can provide reliable age estimates (e.g. Mellett et al., 2012). A further advancement has been made in dating gravel clasts directly (Simms et al., 2012). Whilst this application is in its pilot stage, if successful it has the potential to extend the range of landforms that can be dated using OSL.

Thermochronology

A limitation of luminescence dating is that as it dates sediment directly, it provides an age of depositional processes only. In geomorphology, erosional regimes are equally important in shaping the landscape and there is a need to date bedrock surfaces. The development of OSL-thermochronology, a technique that measures the exhumation rate of bedrock according to its cooling history, has been used to measure denudation rates in New Zealand (Herman et al., 2010). However, the use of OSL dating in determining the timing of event scale erosion processes directly i.e. fluvial incision of bedrock, has yet to be developed.

Rapid age determination

Luminescence dating is a time and labour intensive technique. The time taken between sampling in the field and obtaining an age can be up to 6 months depending on the individual laboratory and sample characteristics. In geomorphology, this can be problematic when undertaking restricted field campaigns. With this in mind, development of the 'range finder' protocol has enabled rapid approximations of OSL ages which can be used to guide sampling strategies and establish the resolution of dating campaigns (Roberts et al., 2009; Durcan et al., 2010).

Finally, the use of portable OSL readers, whilst they do not give an absolute age, can improve the speed of luminescence dating by establishing luminescence characteristics in the field (e.g. Munyikwa et al., 2012). This procedure identifies what sediments are dateable in the field helping laboratories avoid intensive analyses on samples that ultimately do not produce any data.

References

- Adamic, G., Aitken, M.J., 1998. Dose rate conversion factors: update. *Ancient TL*, **16**: 37-50.
- Aitken, M.J., 1985. *Thermoluminescence Dating*. Academic Press, London.
- Aitken, M.J., 1998. *An introduction to optical dating: the dating of Quaternary sediments by the use of photon-stimulated luminescence*. Oxford University Press, 267 p.

- Bailey, R.M., Arnold, L.J., 2006. Statistical modelling of single grain quartz D_e distributions and an assessment of procedures for estimating burial dose. *Quaternary Science Reviews*, **25**: 2475-2502.
- Bateman, M.D., 2008. Luminescence dating of periglacial sediments and structures. *Boreas*, **37**: 574-588.
- Buylaert, J.P., Murray, A.S., Thomsen, K.J., Jain, M., 2009. Testing the potential of an elevated temperature IRSL signal from K-feldspar. *Radiation Measurements*, **44**: 560-565.
- Buylaert, J.P., Jain, M., Murray, A.S., Thomsen, K.J., Thiel, C., Sohbaty, R., 2012. A robust feldspar luminescence dating method for Middle and Late Pleistocene sediments. *Boreas*, **41**: 435-451.
- Duller, G.A.T., Bøtter-Jensen, L., Murray, A.S., 2000. Optical dating of single sand-sized grains of quartz: Sources of variability. *Radiation Measurements*, **31**: 453-457.
- Duller, G.A.T., 2006. Single grain optical dating of glacial deposits. *Quaternary Geochronology*, **1**: 296-304.
- Duller, G.A.T., 2008a. Single-grain optical dating of Quaternary sediments: why aliquot size matters in luminescence dating. *Boreas*, **37**: 589-612.
- Duller, G.A.T., 2008b. Luminescence dating: guidelines on using luminescence dating in archaeology. English Heritage, 43 p.
- Durcan, J.A., Roberts, H.M., Duller, G.A.T., Alizai, A.H., 2010. Testing the use of range-finder OSL dating to inform field sampling and laboratory processing strategies. *Quaternary Geochronology*, **5**: 86-90.
- Fuchs, M., Lang, A., 2009. Luminescence dating of hillslope deposits--A review. *Geomorphology*, **109**: 17-26.
- Fujioka, T., Chappell, J., Fifield, L.K., Rhodes, E.J. 2009. Australian desert dune fields initiated with Pliocene-Pleistocene global climatic shift. *Geology*, **37**:51-54.
- Galbraith, R.F., Green, P.F., 1990. Estimating the component ages in a finite mixture. *International Journal of Radiation Applications and Instrumentation. Nuclear Tracks and Radiation Measurements*, **17**: 197-206.
- Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., Olley, J.M., 1999. Optical Dating of Single and Multiple Grains of Quartz from Jinmium Rock Shelter, Northern Australia: Part I, Experimental Design and Statistical Models. *Archaeometry*, **41**: 339-364.
- Galbraith, R.F., Roberts, R.G., 2012. Statistical aspects of equivalent dose and error calculation and display in OSL dating: An overview and some recommendations. *Quaternary Geochronology*, **11**: 1-27.
- Godfrey-Smith, D.I., Huntley, D.J., Chen, W.H., 1988. Optical dating studies of quartz and feldspar sediment extracts. *Quaternary Science Reviews*, **7**: 373-380.
- Herman, F., Rhodes, E.J., Braun, J., Heiniger, L., 2010. Uniform erosion rates and relief amplitude during glacial cycles in the Southern Alps of New Zealand, as revealed from OSL-thermochronology. *Earth and Planetary Science Letters*, **297**: 183-189.
- Jacobs, Z., 2008. Luminescence chronologies for coastal and marine sediments. *Boreas*, **37**: 508-535.
- Jain, M., Murray, A.S., Botter-Jensen, L., 2004. Optically stimulated luminescence dating: how significant is incomplete light exposure in fluvial environments? *Quaternaire*, **15**: 143-157.
- Lian, O.B., Roberts, R.G., 2006. Dating the Quaternary: progress in luminescence dating of sediments. *Quaternary Science Reviews*, **25**: 2449-2468.
- Madsen, A.T., Murray, A.S., Andersen, T.J., 2007. Optical dating of dune ridges on Rømø, a barrier island in the Wadden Sea, Denmark. *Journal of Coastal Research* **23**:1259-1269
- Madsen, A.T., Murray, A.S., 2009. Optically stimulated luminescence dating of young sediments: A review. *Geomorphology*, **109**: 3-16.
- Mauz, B., Baeteman, C., Bungenstock, F., Plater, A.J., 2010. Optical dating of tidal sediments: Potentials and limits inferred from the North Sea coast. *Quaternary Geochronology*, **5**: 667-678.
- Mellett, C.L., Mauz, B., Plater, A.J., Hodgson, D.M., Lang, A., 2012. Optical dating of drowned landscapes: A case study from the English Channel. *Quaternary Geochronology*, **10**: 201-208.

- Munyakwa, K., Brown, S., Kitabwalla, Z., 2012. Delineating stratigraphic breaks at the bases of postglacial eolian dunes in central Alberta, Canada using a portable OSL reader. *Earth Surface Processes and Landforms*, **37**: 1603-1614.
- Murari, M.K., Achyuthan, H., Singhvi, A.K., 2007. Luminescence studies on the sediments laid down by the December 2004 tsunami event: prospects for the dating of palaeo tsunamis and for the estimation of sediment fluxes. *Current Science*, **92**:367–71.
- Murray, A.S., Roberts, R.G., 1998. Measurement of the equivalent dose in quartz using a regenerative-dose single-aliquot protocol. *Radiation Measurements*, **29**: 503-515.
- Murray, A.S., Wintle, A.G., 2000. Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements*, **32**: 57-73.
- Olley, J.M., Murray, A., Roberts, R.G., 1996. The effects of disequilibria in the uranium and thorium decay chains on burial dose rates in fluvial sediments. *Quaternary Science Reviews*, **15**: 751-760.
- Olley, J.M., Pietsch, T., Roberts, R.G., 2004. Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz. *Geomorphology*, **60**: 337-358.
- Prescott, J.R., Hutton, J.T., 1994. Cosmic ray contributions to dose rates for luminescence and ESR dating: Large depths and long-term time variations. *Radiation Measurements*, **23**: 497-500.
- Roberts, H.M., Durcan, J.A., Duller, G.A.T. 2009. Exploring procedures for the rapid assessment of optically stimulated luminescence range-finder ages. *Radiation Measurements*, **44**: 582-587.
- Rhodes, E.J., 2011. Optically Stimulated Luminescence Dating of Sediments over the Past 200,000 Years. *Annual Review of Earth and Planetary Sciences*, **39**: 461-488.
- Rittenour, T.M., 2008. Luminescence dating of fluvial deposits: Applications to geomorphic, palaeoseismic and archaeological research. *Boreas*, **37**: 613-635.
- Rodnight, H., Duller, G.A.T., Wintle, A.G., Tooth, S., 2006. Assessing the reproducibility and accuracy of optical dating of fluvial deposits. *Quaternary Geochronology*, **1**: 109-120.
- Rowan, A.V., Roberts, H.M., Jones, M.A., Duller, G.A.T., Covey-Crump, S.J., Brocklehurst, S.H., 2012. Optically stimulated luminescence dating of glaciofluvial sediments on the Canterbury Plains, South island, New Zealand. *Quaternary Geochronology*, **8**: 10-22.
- Sawakuchi, A.O., Blair, M.W., DeWitt, R., Faleiros, F.M., Hyppolito, T., Guedes, C.C.F., 2011. Thermal history versus sedimentary history: OSL sensitivity of quartz grains extracted from rocks and sediments. *Quaternary Geochronology*, **6**: 261-272.
- Simms, A.R., Ivins, E.R., DeWitt, R., Kouremenos, P., Simkins, L.M., 2012. Timing of the most recent Neoglacial advance and retreat in the South Shetland islands, Antarctic Peninsula: Insights from raised beaches and Holocene uplift rates. *Quaternary Science Reviews*, **47**: 41-55.
- Stokes, S., 1999. Luminescence dating applications in geomorphological research. *Geomorphology*, **29**: 153-171.
- Thomsen, K.J., Murray, A.S., Jain, M., Bøtter-Jensen, L., 2008. Laboratory fading rates of various luminescence signals from feldspar-rich sediment extracts. *Radiation Measurements*, **43**: 1474-1486.
- Thrasher, I.M., Mauz, B., Chiverrell, R.C., Lang, A., 2009a. Luminescence dating of glaciofluvial deposits: A review. *Earth-Science Reviews*, **97**: 133-146.
- Thrasher, I.M., Mauz, B., Chiverrell, R.C., Lang, A., Thomas, G.S.P., 2009b. Testing an approach to OSL dating of Late Devensian glaciofluvial sediments of the British Isles. *Journal of Quaternary Science*, **24**: 785-801.
- Wallinga, J., Murray, A.S., Duller, G.A.T., Törnqvist, T.E., 2001. Testing optically stimulated luminescence dating of sand-sized quartz and feldspar from fluvial deposits. *Earth and Planetary Science Letters*, **193**: 617-630.
- Wallinga, J., 2002. Optically stimulated luminescence dating of fluvial deposits: a review. *Boreas*, **31**: 303-322.

Wallinga, J., Hobo, N., Cunningham, A.C., Versendaal, A.J., Makaske, B., Middelkoop, H., 2010. Sedimentation rates on embanked floodplains determined through quartz optical dating. *Quaternary Geochronology*, **5**:170–75.

Watanuki, T., Murray, A.S., Tsukamoto, S., 2005. Quartz and polymineral luminescence dating of Japanese loess over the last 0.6 Ma: Comparison with an independent chronology. *Earth and Planetary Science Letters*, **240**: 774-789.

Wintle, A.G., 1973. Anomalous Fading of Thermo-luminescence in Mineral Samples. *Nature*, **245**: 143-144.

Wintle, A.G., 1993. Luminescence dating of aeolian sands: an overview. In: Pye, K. (Editor), *The dynamics and environmental context of aeolian sedimentary systems*. Geological Society Special Publication London, pp. 49-58.

Wintle, A.G., Murray, A.S., 2006. A review of quartz optically stimulated luminescence characteristics and their relevance in single-aliquot regeneration dating protocols. *Radiation Measurements*, **41**: 369-391.

Lichenometry

Hazel E. Trenbirth¹

¹Geography Department, Swansea University (447801@swansea.ac.uk)



ABSTRACT: Lichenometry has been widely used to date rock surfaces since it was developed by Roland Beschel in the 1950s. Two methods have been developed: first, the indirect method, which requires the availability of substrates of known ages and the measurement of lichens growing on them, from which a correlation is established between the size of the lichen and surface age; the second, is direct lichenometry, which requires the measurement of the growth rate of individual lichens in real time and construction of a growth curve. Both methods are reviewed here with reference to species, sampling and measurement techniques and dating-curve construction. Several different variants of the indirect approach are available. While the direct approach has always been regarded as based on sounder biological principles, there are greater practical limitations to surmount.

KEYWORDS: Lichens, chronology, environmental reconstruction, rock surfaces, Holocene.

Introduction

Lichenometry is one of a number of chronological tools used to estimate late-Holocene timescales accurately. It is particularly useful in arctic-alpine environments above the tree line where some crustose lichens grow very slowly, have great longevity (Armstrong, 2004), and where the lack of suitable organic material makes other dating techniques (e.g. dendrochronology and radiocarbon dating) less suitable or not possible.

Roland Beschel pioneered the use of lichens for dating surfaces. As part of his doctoral research completed in 1958 at the University of Innsbruck, he studied the ecology of lichens and measured diameters of lichens on dated tombstones in Austrian cemeteries, determining the growth rates of a number of fast-growing species. Beschel published a series of articles on the subject in the 1950s (Benedict, 2009).

Lichenometry has been used worldwide to date rock and boulder strewn surfaces. Bradwell and Armstrong (2007) identified a

range of landforms that can be dated by lichenometry including river channels, flood deposits, lake shorelines, raised beaches, rock falls, debris flows and moraines. Other reviews include those of Locke *et al.* (1979), Innes (1985), Matthews (1994), Noller and Locke (2000), Benedict (2009) and Armstrong and Bradwell (2010).

The basic principle involved in using lichens for dating is that if the relationship between the size and age of lichens is known, then the age of a surface can be inferred from the size of the lichens present (Innes, 1985). However, a distinction is usually made between the differing methods of estimating the rate of lichen growth; the direct approach involves monitoring measurement of lichen growth rates, and the more commonly used approach of indirect lichenometry involves inferred growth rates from the sizes of lichens on surfaces of known age.

Species selection

The choice of lichen species used in dating is influenced by a sufficient abundance of

individual specimens at the site (Mottershead, 1980). Armstrong and Bradwell (2010) produced a list of selected publications using 27 different species of crustose lichens in direct lichenometry. The most commonly used taxon is the slow growing, long lived, and widely distributed saxicolous (grows on rocks) crustose lichen *Rhizocarpon geographicum* (Loso & Doak, 2006). However, many published accounts that claim to use *Rhizocarpon geographicum* are actually using the larger grouping, *Rhizocarpon* subgenus *Rhizocarpon* (Matthews, 2005).

Measurement techniques

Measuring the growth of a lichen thallus needs to be quick, inexpensive and accurate in order for a large number of thalli to be measured (Armstrong, 1976). The two parameters commonly measured for crustose lichens are the area and the diameter, with the diameter of nearly circular thalli being the most common. In practice, the diameter of non-circular thalli may be defined as the long axis, the short axis or an average of the two. For example, Trenbith and Matthews (2010) measured the long axis using Miltutoyo dial callipers (instrumental precision ± 0.05 mm) (Figure 1). Lichen diameters have also been measured by a ruler (Phillips, 1963; Bradwell & Armstrong, 2007), digital callipers (Lowell *et al.*, 2005; McCarthy, 2003), tracing the thalli outline (Miller & Andrews, 1972; Haworth *et al.*, 1986) and photography (Hooker & Brown, 1977; Proctor, 1983; Rogerson *et al.*, 1986; Bradwell, 2010). Indirect lichenometry usually measures to the nearest mm (with a ruler or steel tape) and does not require such high precision measurements as the direct method.

A photographic technique for monitoring lichen growth is described by Benedict (2008). Photographs are taken over a period of time. Before digital photography the negatives were enlarged in the chemical darkroom and scanned at 600 ppi for use in Adobe Photoshop 6.0. The images were adjusted so that the millimetre scale included in each photograph was reproduced at the actual size. Original and repeat photographs were viewed side by side to ensure

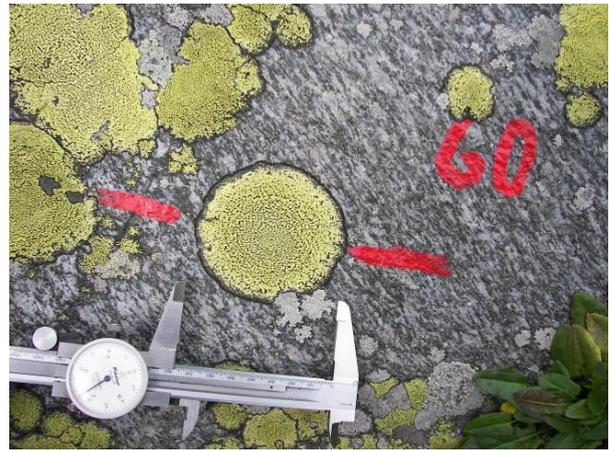


Figure 1. *Rhizocarpon* subgenus lichen with red paint lines marking the measured longest axis (Trenbith & Matthews, 2010).

measurements were made in the same locations. Distances were measured in pixels. Bradwell (2010) adapted this method for a digital camera. The images were enlarged in an image analysis application (Adobe Photoshop 8.0) and accurately overlaid. Scale bars (Figure 2) were used to determine precise size. On-screen measurements were made to an accuracy of 0.01 mm.

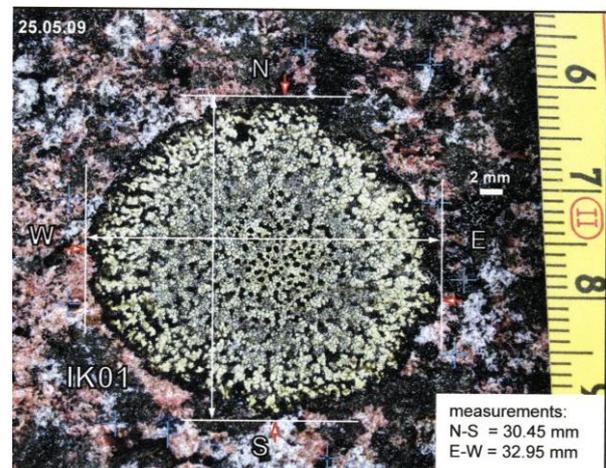


Figure 2. Measurement parameters used by Bradwell (2010), measured on screen in Adobe Photoshop.

The area of lichens can also be used to calculate lichen growth rates. Prior to the use of computer software for calculating area Rydzak (1961) traced the outlines of thalli on plastic sheets and then retraced these at a later date (Figure 3). The surface area of each thallus traced was measured using a planimeter; the procedure is repeated in the

next period of measurement and the increment was calculated in mm². Miller (1973) found measuring the thallus area of *Alectoria minuscula* to have an accuracy of ± 0.5 mm using the tracing method compared to an accuracy of ± 0.01 mm using a photometric method. The latter can be inaccurate over short periods (Armstrong, 1976).

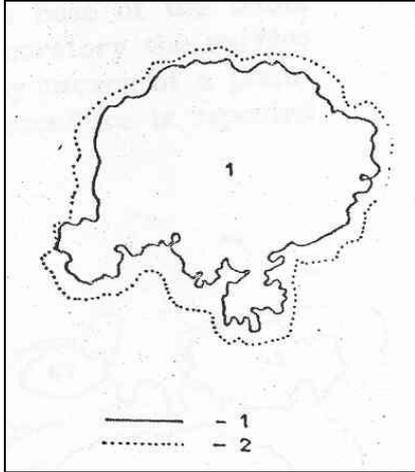


Figure 3. Change in the outline of a thallus of *Lecanora saxicola* Ach between 1957 (1) and 1960 (2) (Rydzak, 1961).

Sampling techniques

The majority of workers attempting to date a substrate use the largest thalli, which has resulted in many debates about the sampling strategy. There is a need to recognise abnormal thalli sizes as there are a number of factors affecting the relationship between the largest lichen and those that are slightly smaller. These include: age of substrate; chemical composition of the substrate; lichen population dynamics; microenvironment; lichen species; presence of anomalously old thalli; and area of the substrate searched (Innes, 1983b). Matthews (1994) pointed out that the use of a large surface area is likely to produce more realistic results as lichen sizes vary over a moraine surface of a single age. To avoid the problem of abnormally large thalli, Calkin and Ellis (1980) discounted any thalli 20% larger than the next largest. Maizels and Petch (1985) proposed the use of the mean of the 100 largest lichen thalli from 100 m² quadrats. Although this is more time consuming, it has the advantage that it is likely to cancel out any extreme values representing localised conditions promoting

excessive growth. Locke *et al.* (1979) recommended using a search area of 400 m², measurement of the ten largest thalli, and the mean of the five largest thalli (including the largest). This method provides a compromise approach in which the effect of extreme measurements may be reduced.

Indirect lichenometry

This requires the recognition of substrates of different known ages and the measurement of lichens growing on them (Mottershead, 1980). A correlation is established between the size of the lichen and the surface age, based on lichen measurements from surfaces of known age (Matthews, 1994). Substrates of known age (control points) that have been used in indirect lichenometry include: anthropogenic surfaces, such as gravestones (Innes, 1983a); stone walls (Benedict, 1967); mine spoil heaps (Karlén, 1973); abandoned farmsteads (Caseldine, 1983); a whaling monument (Werner, 1990); and natural surfaces dated by tree rings and varves (Noller & Locke, 2000). Historical documents can also be used to date surfaces (Anderson & Solid, 1971). A growth curve is then constructed from the lichen sizes from surfaces of known age. A selection of growth rates from different regions is shown in Figure 4.

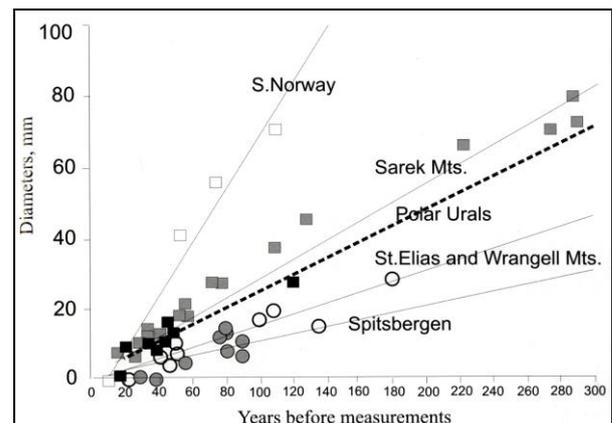


Figure 4. *Rhizocarpon* subgenus *Rhizocarpon* growth rates in several subpolar regions (Solomina *et al.*, 2010): Spitsbergen (grey circles) (Werner, 1990); St. Elias and Wrangell Mts, southern Alaska (open circles) (Denton & Karlén, 1973); Sarek Mountains (grey squares) (Karlén & Denton, 1975); Southern Norway (open squares) (Bickerton & Matthews, 1992); Polar Urals (black squares) (Solomina *et al.*, 2010).

It is essential that the sampling design used to derive the control points for the growth curve is the same as that used on the surfaces of unknown age (Innes, 1985).

There have been numerous studies using the indirect approach to lichenometry, involving numerous different methods adapted to local conditions. This makes it impossible to recommend a standard technique.

One method frequently used is to take the largest lichen or the mean measurements of lichen diameters or the radius. Hughes (2007) measured *Aspicilia calcarea* agg. lichens in Montenegro to date recent behaviour of the Debeli Namet glacier. The mean of the five largest lichens from a total sample size of 30 lichens measured at random from 50 gravestones and monuments was used to construct the lichen age/size relationship (Figure 5). The mean diameter of the five largest lichens from a random sample of 30 lichens was then calculated for each sediment ridge to compare with the size/age relationship established from the gravestones and monuments to determine the approximate age of the landform.

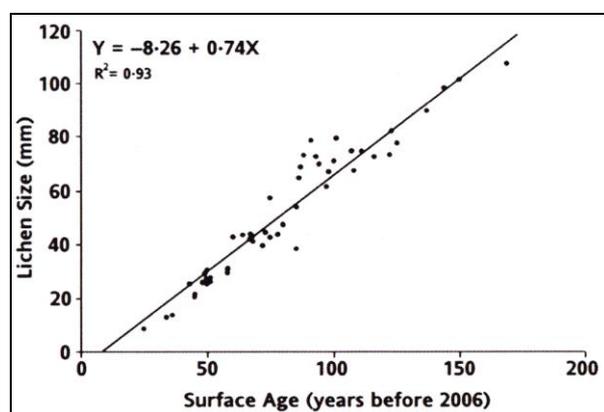


Figure 5. Regression of surface age against the mean size of the five largest lichen thalli (50 samples) obtained by measuring lichens on monuments and gravestones of known age (Hughes, 2007).

Cook-Talbot (1991) calculated the mean diameters of the 5, 10 and 50 largest lichens on clasts in stone circle borders and centres in Norway as part of an investigation into palaeoenvironmental reconstruction in an alpine periglacial environment.

The largest lichen method was used by O'Neal and Schoenenberger (2003) to produce a *Rhizocarpon geographicum* growth curve for the Cascade Range of Washington and northern Oregon, USA. The largest lichen was identified on each man-made structure and each moraine for 22 sites. The maximum diameters of individual lichens with circular or nearly circular thalli were measured on boulders with surface area > 0.3 m² or on rock walls. For the man-made structures the largest lichens on individual blocks were measured. These structures were selected through the altitudinal range of the glacier moraines to ensure no bias between the setting and the altitude.

The accuracy of such methods is difficult to estimate. Matthews (2005) compared the lichenometric dates based on indirect lichenometry (using the mean of the five largest lichens) to the control points of known age. The surfaces, which were about 50 years old, had an accuracy of ±6 years whereas surfaces about 230 years old had an accuracy of ±35 years.

One indirect technique uses size-frequency distributions. Benedict (1967) was the first to use this approach when dating a native Indian wall in the Colorado Front Range. This method has been adopted by several studies including: Andersen and Sollid (1971), Innes (1983b), Cook-Talbot (1991), Bradwell (2004), McKinzey *et al.* (2004), Bradwell *et al.* (2006), Golledge *et al.* (2010) and Roberts *et al.* (2010). In South Georgia, Roberts *et al.* (2010) measured the diameters of 872 lichens on boulders on five moraines. These lichens were analysed using the size-frequency method. The lichens were measured in pre-defined areas of approximately equal dimensions. Size-frequency analysis of the data was then used to determine whether using the largest five or the largest ten lichens could be used for dating. The lichen population size-frequency data were plotted as histograms, Q-Q plots (which are quantile probability plots showing all data and illustrating where each data point deviates from the theoretical normal distribution) and also plotted using a class size of 3 mm expressed as log₁₀ percentage frequency against diameter size (Figure 6).

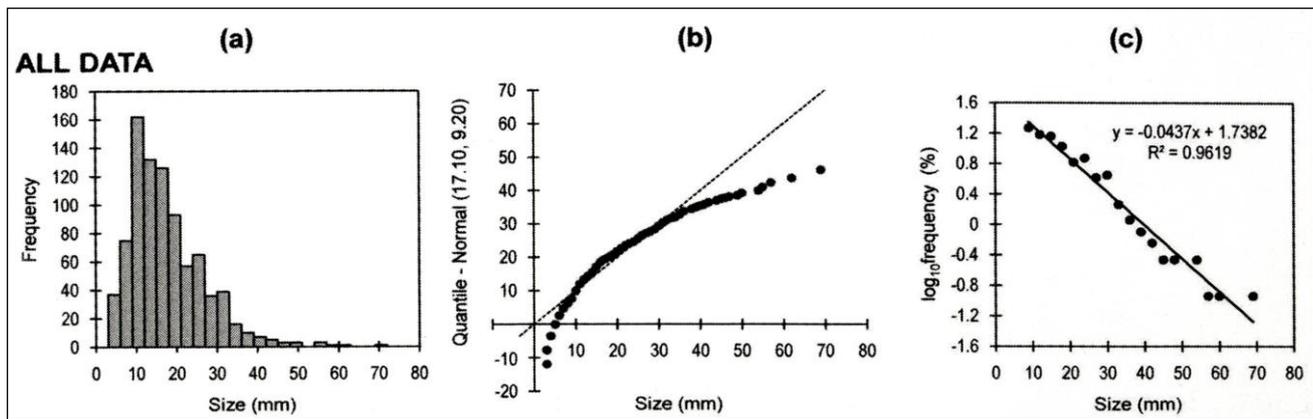


Figure 6. Lichen size data for all data shown as: (a) histograms plotted with a class size of 3mm; (b) Q-Q plots; (c) lichen population size-frequency data plotted using a class size of 3 mm expressed as \log_{10} percentage frequency against diameter size (Roberts *et al.*, 2010).

The results showed a notable increase in lichen size with increased distance from the cirque headwall. In the absence of an age-size curve for South Georgia the largest lichen, size-frequency method and long-term growth rates established on the nearest Antarctic localities were used to establish lichen growth-rate ranges, and likely age ranges for the moraines.

Percentage cover measurements of lichens is another indirect lichenometry technique. The premise is that the total surface area of a rock covered by one or more species of lichen increases through time. The deposit with a greater lichen cover than another has been taken as being older. Innes (1986) concluded from studying two glacier forelands in Norway that this technique is subjective as the results obtained by different observers are not reproducible. Substantial variations in cover were found on the moraines and this method appears to be more sensitive to environmental variations. The technique appears to be unreliable but has considerable potential for the dating of smaller boulders or surfaces that cannot be dated by conventional lichenometric techniques. Several studies have used this technique including Carroll (1974) who used the percentage of lichen cover as one relative dating technique to date Quaternary deposits in Arikaree Cirque, Colorado.

McCarroll (1993) proposed a new technique for using lichenometry on surfaces comprising material deposited at different times, as traditional techniques are used for

surfaces deposited in a single event. This approach uses simulation-modelling in an attempt to translate the lichen-size frequency distributions obtained from diachronous surfaces into the age-frequency of surface boulders. A record of snow avalanche activity was obtained from using this approach in western Norway (McCarroll, 1993, 1994, 1995; McCarroll *et al.*, 1995) and rock fall activity (McCarroll, 1994; McCarroll *et al.*, 1998).

A new statistical model for lichenometry has been developed using the generalized extreme value (GEV) distribution, building a Bayesian hierarchical value model (Cooley *et al.*, 2006). It uses the largest lichen measurements and applies statistical theory of extreme values. It offers several advantages, including: accounting for the uncertainty associated with the estimates of moraine ages; accommodating any growth curve function; and allowing for spatial variation in lichen growth. Further studies on this method include Jomelli *et al.* (2007, 2008), Naveau *et al.* (2007), Rabatel *et al.* (2008) and Chenet *et al.* (2010). Although the GEV is computationally intensive (Bradwell, 2009), it appears to enable a better quantification of uncertainty (Jomelli *et al.*, 2010).

A selection of recent applications, locations and different indirect lichenometric techniques is shown in Table 1. Different methods have been used depending on the availability of lichens and the terrain on which they are growing.

Table 1. Aspects of selected indirect lichenometric dating studies: location, area surveyed, area searched, method used, lichen size range and the authors.

Location	Area surveyed	Search area	Method	Lichen size range	Author
Bolivian Andes	15 glacier forelands	Blocks, 2678 lichens	Largest lichen	9-41 mm	Rabatel <i>et al.</i> (2008)
Patagonian Andes	6 glacier forelands	Entire surface each moraine	Largest lichen Size-frequency	0.6-13.5 cm (largest lichen)	Garibotti & Villalba (2009)
Montenegro	5 cirques	Sites 25 m long, 10 m wide	Mean 5 largest	67.5-142.3 mm	Hughes (2010)
Vancouver Island, Canada	2 glacier forelands		Largest lichen Mean 5 largest	33.2-97.4 mm (largest lichen)	Lewis & Smith (2004)
Iceland	Glacier forelands	30 m ² / site	Largest lichen Size-frequency	19-95 mm (largest lichen)	Bradwell (2004)
Norway	16 glacier forelands	25x8 m/site	Largest lichen Mean 5 largest	14-158 mm (largest lichen)	Matthews (2005)
Cascade Range, USA	Glacier forelands		Largest lichen	11.8-50 mm	O'Neal & Schoenenberger (2003)
New Zealand	Glacier forelands	Sample areas 10 to 100 m ²	Largest lichen Mean 5 largest Mean 10 largest 98% quantile		Lowell <i>et al.</i> (2005)
Iceland	Rock Glacier	Entire surface	Size-frequency	<5 to 80 mm (size class)	Hamilton & Whalley (1995)
Austria	Talus slope	300 boulders	Mean 5 largest % cover	5.1-27.9 mm (mean 5 largest)	Sass (2010)
Norway	Talus slope	100 boulders, 25 sites	Largest lichen Size-frequency	16.6-37.2 mm	McCarroll <i>et al.</i> (2001)
Antarctic	Former snow-patch areas	13 sites	Size-frequency		Golledge <i>et al.</i> (2010)
Poland	Debris flow		Largest lichen		Jonasson <i>et al.</i> (1991)
Iceland	Proglacial river terrace	Entire surface of each terrace	Mean 5 largest	14.3-67.4 mm	Thompson & Jones (1986)
Sweden	Raised beaches	All rocks and surfaces	Largest lichen	19-358 mm	Broadbent & Bergqvist (1986)
Norway	Lake shoreline	25 m sections, 6 sites	Mean 5 largest	59-310 m	Matthews <i>et al.</i> (1986)
Lake District, England	Flood event	All deposit measured	Mean 5 largest		Johnson & Warburton (2002)
France	Flood event, 4 rivers	Blocks in riverbeds	Mean 5 largest Size-frequency	0.1-81 mm	Gob <i>et al.</i> (2010)
Cumbria, England	Valley floor development	7 zones	Largest lichen	4->95 mm	Harvey <i>et al.</i> (1984)
Australia	Archaeological structures		Largest lichen	5.7-74.4 mm	Müller (2005)
Central Asia, USSR	Seismic dislocations		Statistical model		Smirnova & Nikonov (1990)
New Zealand	Earthquake blocks, rockfall event		Largest lichen	0.2-1m	Bull & Brandon (1998)

Direct lichenometry

Observations of individual lichens at repeated intervals over time are required for the direct approach. Several years of measurement are needed to assess the growth rate. A growth curve is constructed from the growth measurements of lichens of varying size.

Relatively few studies have adopted the direct approach, largely because of practical difficulties associated with the slow-growing crustose lichens that are most commonly used for dating purposes. Bradwell and Armstrong (2007) reviewed 13 studies (including their own) that used the direct approach for the *Rhizocarpon* subgenus (Hausmann 1948; Ten Brink 1973; Hooker 1980; Armstrong 1983, 2005; Proctor 1983; Rogerson *et al.* 1986; Haworth *et al.* 1986; Matthews 1994; Winchester and Chaujar 2002; McCarthy 2003; Sancho and Pintado 2004). These studies showed considerable variability in the estimated mean annual diameter growth rates ranging from 0.08 to 1.47 mm/yr. The number of years over which the measurements were made ranged from 1.5 to 12 years, with only two records > 6 years. Furthermore, the majority of sample sizes have been small at single sites. This has resulted in little information on temporal variability and the effects of habitat variation.

A recent exception is the study of Trenbirth and Matthews (2010) who monitored diameter measurements over 25 years for 2,795 individuals of the *Rhizocarpon* subgenus at 47 sites on 18 glacier forelands in southern Norway. Individual lichens were selected with well-defined orbicular thalli, free of competition from other lichens and located on stable boulders or, in a few cases, on bedrock. Red marker lines were painted on the adjacent rock (Figure 1), to indicate the position of the measured long axis of each thallus, and an identification number was painted close to the lichen. The long axis was then measured using Mitutoyo dial callipers (instrumental precision ± 0.05 mm). The mean annual growth rate for the 47 sites ranged from 0.43 mm/yr to 0.87 mm/yr. The main between-site pattern in these data related to surface age, with the growth rate declining with increasing surface age.

Lichen growth curves (age-calibration curves), which relate lichen size to lichen age, were constructed for three sites using the method of Armstrong (1976) with initial 5.0-mm lichen-size classes (Figure 7). Wider class intervals were used for lichen sizes above about 100 mm and also for lichens <10 mm diameter at one of the sites (Styggedalsbreen) because of the small number of lichens. Growth rate at the

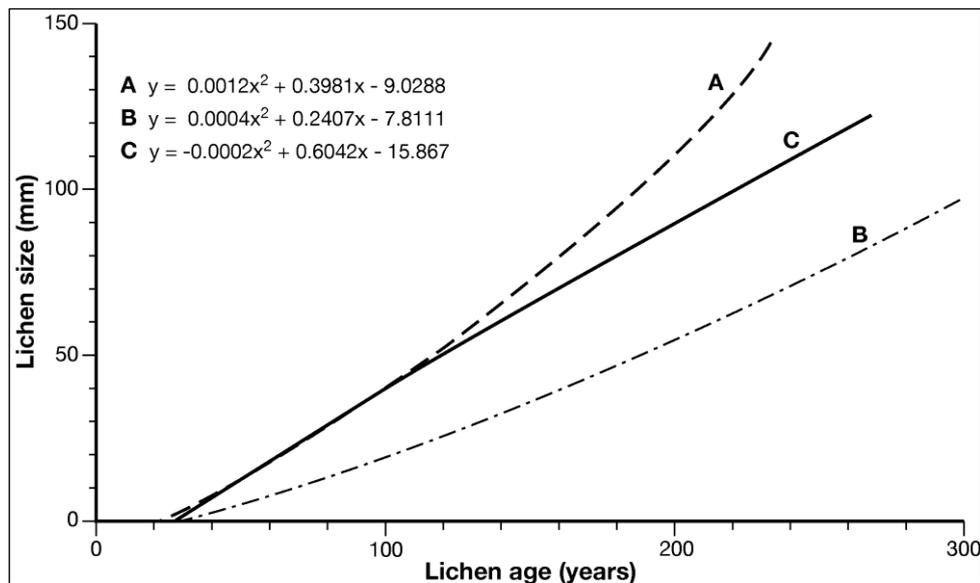


Figure 7. Growth curves constructed for direct measurements: (A) Storbreen AD1750 high altitude site; (B) Styggedalsbreen AD1750 site; and (C) Nigardsbreen AD1750 site (Trenbirth & Matthews, 2010).

midpoint of each size class was calculated from the data. The time taken for each 5.0-mm growth increment, estimated from the annual growth rate within each size class, was plotted against the size-class mid-point. Second-order polynomial curves were fitted to these mid-point values to represent the generalised growth curve. A lag-time of 15 years was assumed on the basis of previous indirect lichenometric dating in southern Norway. The growth curves (Figure 7) reflect the widely differing growth rates measured at the three sites.

Advantages and limitations

Indirect lichenometry requires surfaces of known age. Therefore indirect lichenometry cannot be employed in areas where there are no surfaces of known age (Smirnova & Nikonov, 1990). Jochimsen (1973) criticised lichenometry as numerous potential sources of error arise from ecological factors. Noller and Locke (2000) drew attention to one of the unresolved dilemmas of lichenometry, namely lichen thallus sensitivity to climate. Growth curves that have been developed for one region will not necessarily be applicable elsewhere as environmental factors make them area-specific (Walker, 2005).

Benedict (1990) suggested that growth curves derived from direct monitoring of annual growth of individual lichens were likely to be unrepresentative due to current climate being atypically warm with no indication of imminent cooling. This should affect dating of Little Ice Age moraines that have experienced mean temperatures substantially cooler than present. Matthews (1994) and Trenbirth and Matthews (2010) concluded that direct measurements produce extremely variable growth rates for the slow-growing species of lichens used for dating purposes. This point combined with the small number of thalli measured over relatively few years leads to an inability to determine accurately the form of the lichen growth curve. Some direct studies have therefore produced results that conflict with indirect studies at the same location.

Conclusion

Lichenometry is an inexpensive and widely applicable tool for estimating surface ages. Its application is straightforward and does not require the user to undergo intensive training or use sophisticated instruments. It has the ability to date surfaces during the last 500 years a period when radiocarbon dating is least efficient (Armstrong, 2004). A major advantage is the ubiquitous presence of lichens on many rocky substrates (Noller & Locke, 2000). Lichenometry is a useful proxy dating method and has been successfully used on its own or in combination with other dating methods (e.g. Solomina and Calkin, 2003).

References

- Andersen JL, Sollid JL. 1971. Glacial chronology and glacial geomorphology in the marginal zones of the glaciers, Midtdalsbreen and Nigardsbreen, south Norway. *Norsk Geografisk Tidsskrift* **25**: 1-38.
- Armstrong RA. 1976. Studies of the growth rates of lichens. In *Lichenology: Progress and Problems*, Brown DH, Hawksworth DL, Bailey RH (eds). Academic Press: London; 309-322.
- Armstrong RA. 1983. Growth curve of the lichen *Rhizocarpon geographicum*. *New Phytologist* **73**: 913-918.
- Armstrong RA. 2004. Lichens, lichenometry and global warming. *The Microbiologist* **5**: 32-35.
- Armstrong RA. 2005. Growth rate measurements of *Rhizocarpon geographicum* in the Snoqualimie Pass, Washington, USA. *Arctic, Antarctic and Alpine Research* **37**: 411-415.
- Armstrong RA, Bradwell, T. 2010. Growth of crustose lichens: a review. *Geografiska Annaler* **92** A(1): 3-17.
- Benedict JB. 1967. Recent glacial history of an alpine area in the Colorado Front Range, U.S.A. I. Establishing a lichen-growth curve. *Journal of Glaciology* **6**(48): 817-832.
- Benedict JB. 1990. Experiments on lichen growth. I. Seasonal patterns and environmental controls. *Arctic and Alpine Research* **22**(3): 244-254.
- Benedict JB. 2008. Experiments on Lichen Growth, III. The shape of the Age-Size Curve.

- Arctic, Antarctic, and Alpine Research* **40**(1): 15-26.
- Benedict JB. 2009. A review of lichenometric dating and its applications to archaeology. *American Antiquity* **74**(1): 143-172.
- Bickerton RW, Matthews JA. 1992. On the accuracy of lichenometric dates: an assessment based on the 'Little Ice Age' moraine sequence of Nigardsbreen, southern Norway. *The Holocene* **2**: 227-237.
- Bradwell T. 2004. Lichenometric dating in southeast Iceland: the size-frequency approach. *Geografiska Annaler* **86** A(1): 31-41.
- Bradwell T. 2009. Lichenometric dating: a commentary, in the light of some recent statistical studies. *Geografiska Annaler* **91** A(2): 61-69.
- Bradwell T. 2010. Studies on the growth of *Rhizocarpon geographicum* in NW Scotland, and some implications for lichenometry. *Geografiska Annaler* **92** A(1): 41-52.
- Bradwell T, Armstrong RA. 2007. Growth rates of *Rhizocarpon geographicum* lichens: a review with new data from Iceland. *Journal of Quaternary Science* **22**(4): 311-320.
- Bradwell T, Dugmore AJ, Sugden DE. 2006. The Little Ice Age glacier maximum in Iceland and the North Atlantic Oscillation: evidence from Lambatungnajökull, southeast Iceland. *Boreas* **35**: 61-80.
- Broadbent ND, Bergqvist KI. Lichenometric chronology and archaeological features on raised beaches: Preliminary results from the Swedish North Bothnian coastal region. *Arctic and Alpine Research* **18**: 297-306.
- Bull WB, Brandon MT. 1998. Lichen dating of earthquake-generated regional rock-fall events, Southern Alps, New Zealand. *Geological Society of America Bulletin* **110**(1): 60-84.
- Calkin PE, Ellis JM. 1980. A lichenometric dating curve and its application to Holocene glacier studies in the Central Brooks Range, Alaska. *Arctic and Alpine Research* **12**(3): 245-264.
- Carroll T. 1974. Relative Age Dating Techniques and a Late Quaternary Chronology, Arikaree Cirque, Colorado. *Geology* **2**: 321-325.
- Caseldine CJ. 1983. Resurvey of the margins of Gljúfurajökull and the chronology of recent deglaciation. *Jökull* **33**: 111-118.
- Chenet M, Roussel E, Jomelli V, Grancher D. 2010. Asynchronous Little Ice Age glacial maximum extent in southeast Iceland. *Geomorphology* **114**: 253-260.
- Cook-Talbot JD. 1991. Sorted circles, relative-age dating and palaeoenvironmental reconstruction in an alpine periglacial environment, eastern Jotunheimen, Norway: lichenometric and weathering-based approaches. *The Holocene* **1**(2): 128-141.
- Cooley D, Naveau P, Jomelli V, Rabatel A, Grancher D. 2006. A Bayesian hierarchical extreme value model for lichenometry. *Environmetrics* **17**: 555-574.
- Denton GH, Karlén W. 1973. Lichenometry: its application to Holocene moraine studies in Southern Alaska and Swedish Lapland. *Arctic and Alpine Research* **5**: 347-372.
- Garibotti IA, Villalba R. 2009. Lichenometric dating using *Rhizocarpon* subgenus *Rhizocarpon* in the Patagonian Andes, Argentina. *Quaternary Research* **71**: 271-283.
- Gob F, Bravard JP, Petit F. 2010. The influence of sediment size, relative grain size and channel slope on initiation of sediment motion in boulder bed rivers. A lichenometric study. *Earth Surface Processes and Landforms* **35**: 1535-1547.
- Golledge NR, Everest JD, Bradwell T, Johnson JS. 2010. Lichenometry on Adelaide Island, Antarctic Peninsula: size-frequency studies, growth rates and snowpatches. *Geografiska Annaler* **92** A(1): 111-124.
- Hamilton SJ, Whalley WB. 1995. Preliminary results from the lichenometric study of the Nautardalur rock glacier, Tröllaskagi, northern Iceland. *Geomorphology* **12**: 123-132.
- Harvey AM, Alexander RW, James PA. 1984. Lichens, soil development and the age of Holocene valley floor landforms: Howgill Fells, Cumbria. *Geografiska Annaler* **66A** (4): 353-366.
- Hausmann EH. 1948. Measurements of the annual growth rate of two species of rock lichens. *Bulletin of the Torrey Botanical Club* **75**: 116-117.

- Haworth LA, Calkin PE, Ellis JM. 1986. Direct measurement of lichen growth in the Central Brooks Range, Alaska, U.S.A., and its application to lichenometric dating. *Arctic and Alpine Research* **18**: 289-296.
- Hooker TN. 1980. Factors affecting the growth of Antarctic crustose lichens. *British Antarctic Survey Bulletin* **50**: 1-19.
- Hooker TN, Brown DH. 1977. A photographic method for accurately measuring the growth of crustose and foliose saxicolous lichens. *Lichenologist* **9**: 65-75.
- Hughes PD. 2007. Recent behaviour of the Debeli Namet glacier, Durmitor, Montenegro. *Earth Surface Processes and Landforms* **32**(10): 259-267.
- Hughes PD. 2010. Little Ice Age glaciers in the Balkans: low altitude glaciation enabled by cooler temperatures and local topoclimatic controls. *Earth Surface Processes and Landforms* **35**: 229-241.
- Innes JL. 1983a. Use of an aggregated *Rhizocarpon* 'species' in lichenometry: an evaluation. *Boreas* **12**: 183-190.
- Innes JL. 1983b. Size frequency distribution as a lichenometric technique: an assessment. *Arctic and Alpine Research* **15**(3): 285-294.
- Innes JL. 1985. Lichenometry. *Progress in Physical Geography* **9**(2): 187-254.
- Innes JL. 1986. The use of percentage cover measurements in lichenometric dating. *Arctic and Alpine Research* **18**(2): 209-216.
- Jochimsen M. 1973. Does the size of lichen thalli really constitute a valid measure for dating glacial deposits? *Arctic and Alpine Research* **5**(4): 417-424.
- Johnson RM, Warburton J. 2002. Flooding and geomorphic impacts in a mountain torrent: Raise Beck, central Lake District, England. *Earth Surface Processes and Landforms* **27**: 945-969.
- Jomelli V, Grancher D, Naveau P, Cooley D, Brunstein D. 2007. Assessment study of lichenometric methods for dating surfaces. *Geomorphology* **86**: 131-143.
- Jomelli V, Grancher D, Brunstein D, Solomina O. 2008. Recalibration of the yellow *Rhizocarpon* growth curve in the Cordillera Blanca (Peru) and implications for LIA chronology. *Geomorphology* **93**: 201-212.
- Jomelli V, Naveau P, Cooley D, Grancher D, Brunstein D, Rabatel A. 2010. Letter To The Editor: A Response To Bradwell's Commentary ON Recent Statistical Studies In LICHENOMETRY. *Geografiska Annaler* **92** A: 485-487.
- Jonasson C, Kot M, Kotarba A. 1991. Lichenometrical studies and dating of debris flow deposits in the High Tara Mountains, Poland. *Geografiska Annaler* **73** A(3-4): 141-146.
- Karlén W. 1973. Holocene glacier and climatic variations, Kebnekaise Mountains, Swedish Lapland. *Geografiska Annaler* **55** A: 29-63.
- Karlén W, Denton G. 1975. Holocene glacial variations in Sarek National Park, Northern Sweden. *Boreas* **5**: 25-56.
- Lewis DH, Smith DJ. 2004. Little Ice Age glacial activity in Strathcona Provincial Park, Vancouver Island, British Columbia, Canada. *Canadian Journal of Earth Sciences* **41**: 285-297.
- Locke WW, Andrews JT, Webber PJ. 1979. *A manual for lichenometry*. British Geomorphological Research Group: London; 47pp.
- Loso MG, Doak DF. 2006. The biology behind lichenometric dating curves. *Oecologia* **147**: 223-229.
- Lowell TV, Schoenenberger K, Deddens JA, Denton GH, Smith C, Black J, Hendy CH. 2005. *Rhizocarpon* calibration curve for the Aoraki/Mount Cook area of New Zealand. *Journal of Quaternary Science* **20**(4): 313-325.
- Maizels JK, Petch JR. 1985. Age determination of intermoraine areas, Austerdalen, southern Norway. *Boreas* **14**: 51-65.
- Matthews JA. 1994. Lichenometric dating: a review with particular reference to 'Little Ice Age' moraines in southern Norway. In *Dating in exposed and surface contexts*, Beck C. (ed). University of New Mexico Press: Albuquerque; 185-212.
- Matthews JA. 2005. 'Little Ice Age' glacier variations in Jotunheimen, southern Norway: a study in regionally controlled lichenometric dating of recessional moraines with implications for climate and lichen growth rates. *The Holocene* **15**(1): 1-19.

- Matthews JA, Dawson AG, Shakesby RA. 1986. Lake shoreline development, frost weathering and rock platform erosion in an alpine periglacial environment, Jotunheimen, southern Norway. *Boreas* **15**: 33-50.
- McCarroll D. 1993. Modelling late-Holocene snow-avalanche activity: incorporating a new approach to lichenometry. *Earth Surface Processes and Landforms* **18**: 527-539.
- McCarroll D. 1994. A new approach to lichenometry: dating single age and diachronous surfaces. *The Holocene* **4**: 383-396.
- McCarroll D. 1995. Lichens: lichenometric dating of diachronous surfaces. *Earth Surface Processes and Landforms* **20**: 829-831.
- McCarroll D, Matthews JA, Shakesby RA. 1995. Late Holocene snow-avalanche activity in southern Norway: interpreting lichen size-frequency distributions using an alternative to simulation modelling. *Earth Surface Processes and Landforms* **20**: 465-471.
- McCarroll D, Shakesby RA, Matthews JA. 1998. Spatial and Temporal Patterns of Late Holocene Rockfall Activity on a Norwegian Talus Slope: a Lichenometric and Simulation-Modelling Approach. *Arctic and Alpine Research* **30**(1): 51-60.
- McCarroll D, Shakesby RA, Matthews JA. 2001. Enhanced Rockfall Activity during the Little Ice Age: Further Lichenometric Evidence from a Norwegian Talus. *Permafrost and Periglacial Processes* **12**: 157-164.
- McCarthy DP. 2003. Estimating lichenometric ages by direct and indirect measurement of radial growth: a case study of *Rhizocarpon* agg. At the Illecillewaet Glacier, British Columbia. *Arctic and Alpine Research* **35**: 203-213.
- McKinzey KM, Orwin JF, Bradwell T. 2004. Re-dating the moraines at Skálafellsjökull and the Heinabergsjökull using different lichenometric methods: implications for the timing of the Icelandic Little Ice Age maximum. *Geografiska Annaler* **86** A(4): 319-335.
- Miller GH, Andrews JT. 1972. Quaternary history of northern Cumberland Peninsula, East Baffin Island, N.W.T., Canada Part VI: Preliminary lichen growth curve for *Rhizocarpon geographicum*. *Geological Society of America Bulletin* **83**: 1133-1138.
- Miller GH. 1973. Variations in lichen growth from direct measurements: Preliminary curves for *Alectoria minuscula* from eastern Baffin Island, N.W.T., Canada. *Arctic and Alpine Research* **5**(4): 333-339.
- Mottershead DN. 1980. Lichenometry – some recent applications. In *Timescales in Geomorphology*, Davidson RA, Lewin J (eds). Wiley & Sons Ltd.: Chichester; 95-108.
- Müller G. 2005. Growth of the Lichen *Buellia albula* and its Potential for Lichenometric Dating in South-eastern Australia. *Geographical Research* **43**(3): 267-273.
- Naveau P, Jomelli V, Cooley D, Delphine G, Rabatel A. 2007. Modeling Uncertainties in Lichenometry Studies. *Arctic, Antarctic and Alpine Research* **39**(2): 277-285.
- Noller JS, Locke WW. 2000. Lichenometry. In *Quaternary Geochronology: Methods and Applications*, Noller JS, Sowers JM, Lettis WR (eds). American Geophysical Union: Washington; 261-272.
- O'Neal MA, Schoenenberger KR. 2003. A *Rhizocarpon geographicum* growth curve for the Cascade Range of Washington and northern Oregon, USA. *Quaternary Research* **60**: 233-241.
- Phillips HC. 1963. Growth rate of *Parmelia isidiosa* (Mull. Arg.) Hale. *Journal of the Tennessee Academy of Science* **38**(3): 95-96.
- Proctor MCF. 1983. Sizes and growth-rates of thalli of the lichen *Rhizocarpon geographicum* on the moraines of the Glacier De Valsorey, Valais, Switzerland. *Lichenologist* **15**(3): 249-261.
- Rabatel A, Francou B, Jomelli V, Naveau P, Grancher D. 2008. A chronology of the Little Ice Age in the tropical Andes of Bolivia (16°S) and its implications for climate reconstruction. *Quaternary Research* **70**: 198-212.
- Roberts SJ, Hodgson DA, Shelley S, Royles J, Griffiths HJ, Deen TJ, Thorne AS. 2010. Establishing lichenometric ages for nineteenth- and twentieth-century glacier fluctuations on South Georgia (South Atlantic). *Geografiska Annaler* **92** A(1): 125-139.
- Rogerson RJ, Evans DJA, McCoy WD. 1986. Five-year growth of rock lichens in a low-

arctic mountain environment, northern Labrador. *Géographie physique et Quaternaire* **XL**: 85-91.

Rydzak J. 1961. Investigations on the Growth Rate of Lichens. *Annales Universitatis Mariae Curie-Sklodowska Lubin-Polonia* **16**: 1-15.

Sancho LG, Pintado A. 2004. Evidence of high annual growth rate for lichens in the maritime Antarctic. *Polar Biology* **27**: 312-319.

Sass O. 2010. Spatial and temporal patterns of talus activity – a lichenometric approach in the Stubai Alps, Austria. *Geografiska Annaler* **92** A(3): 375-391.

Smirnova TY, Nikonov AA. 1990. A revised lichenometric method and its application dating great past earthquakes. *Arctic and Alpine Research* **22**(4): 375-388.

Solomina O, Calkin PE. 2003. Lichenometry as Applied to Moraines in Alaska, U.S.A., and Kamchatka, Russia. *Arctic, Antarctic, and Alpine Research* **35**(2): 129-143.

Solomina O, Ivanov M, Bradwell T. 2010. Lichenometric studies on moraines in the Polar Urals. *Geografiska Annaler* **92** A(1): 81-99.

Ten Brink NW. 1977. Colonisation, growth, succession and competition. In *Lichen Ecology*, Seward MRD (ed). Academic Press: London; 31-68.

Thompson A, Jones A. 1986. Rates and causes of Proglacial river terrace formation in southeast Iceland: an application of lichenometric dating techniques. *Boreas* **15**: 231-246.

Trenbith HE, Matthews JA. 2010. Lichen growth rates on glacier forelands in southern Norway: Preliminary results from a 25-year monitoring programme. *Geografiska Annaler* **92** A(1): 19-39.

Walker M. 2005. *Quaternary Dating Methods*. Antony Rowe Ltd.: Chippenham.

Werner A. 1990. Lichen growth rates for the northwest coast of Spitsbergen, Svalbard. *Arctic and Alpine Research* **22**(2): 129-140.

Winchester V, Chaujar RK. 2002. Lichenometric dating of slope movements, Nant Ffrancon, North Wales. *Geomorphology* **47**: 61-74.

4.2.8. Dendrochronology

Max C.A. Torbenson¹

¹ Department of Geosciences, University of Arkansas (mtorbens@uark.edu)



ABSTRACT: Dendrochronology is a dating technique that utilises the common growth signal in trees of the same species growing in the same area under similar conditions. Cross-dating is achieved by matching ring-width patterns between individual trees. Exact calendric dates can be produced for subfossil material if a temporal overlap exists with modern samples. The death and establishment of trees, as well as ring-widths and anatomical features in the wood, have been used to study past geomorphic episodes including mass movement activity, earthquakes, volcanic eruptions, fire, glacier movement, and flooding. In addition to providing high-precision dates for known disturbances, evidence from tree rings can help identify previously unknown events and be used to study the synchronicity of abrupt environmental changes across space. This paper summarizes the theory and practice of tree-ring dating and reviews its application in dendro-geomorphological research.

KEYWORDS: tree rings, chronology, cross-dating, Common Era, Holocene

Introduction

The growth of trees has attracted scientific interest for thousands of years. Theophrastus, a student of Aristotle, noted that stones forced into a tree became “hidden” with time (Studhalter, 1956), indicating that new wood is formed close to the exterior of the tree. Leonardo da Vinci is said to have discovered the annual nature of tree rings, as well as their connection to climate (Stallings *et al.*, 1937). It was not until the early 20th century, however, that dendrochronology and tree rings as a dating method was fully developed into a science, largely through the works of A.E. Douglass (Douglass, 1919; Baillie 1982).

Dendrochronology has been used extensively in a range of scientific fields over the past 100 years, including archaeology (Baillie, 1982), ecology (Schweingruber, 1996), and palaeoclimatology (Fritts 1976). Tree rings provided calibration during the development of the radiocarbon dating because of their ability to record atmospheric carbon fluctuations with high dating resolution (Stuiver and Becker, 1993; Friedrich *et al.*, 2004). Prehistoric artefacts have been dated, with annual and sometimes intra-annual precision, using dendrochronology (e.g. Visser, 2015). The

relationship between climate variability and tree growth has facilitated continental to hemispheric reconstructions of drought and temperature (Cook *et al.*, 1999; Esper *et al.*, 2002), and tree rings have provided regional and local perspectives of past ecological and climatic shifts across the world (Swetnam and Betancourt, 1998; Salzer *et al.*, 2009; Buckley *et al.*, 2010). The high precision dating offered by dendrochronology can also be useful when measuring rates and timing of geomorphic processes (Strunk, 1997; Stoffel and Bollschweiler, 2008; Koprowski *et al.*, 2010). Trees can be killed by geomorphic processes and/or become established on recently modified surfaces, and the time of death/germination will therefore serve as a temporal indicator of abrupt landscape change. Surviving trees can also record evidence of these events in the wood they form, and by establishing an absolute chronology for each tree ring it is possible to assign calendric dates to such markers.

Principles of dendrochronology

Trees cover approximately 30% of Earth’s land surface and are found on all continents except for Antarctica (Pan *et al.*, 2013). The distribution of any given species is governed by the

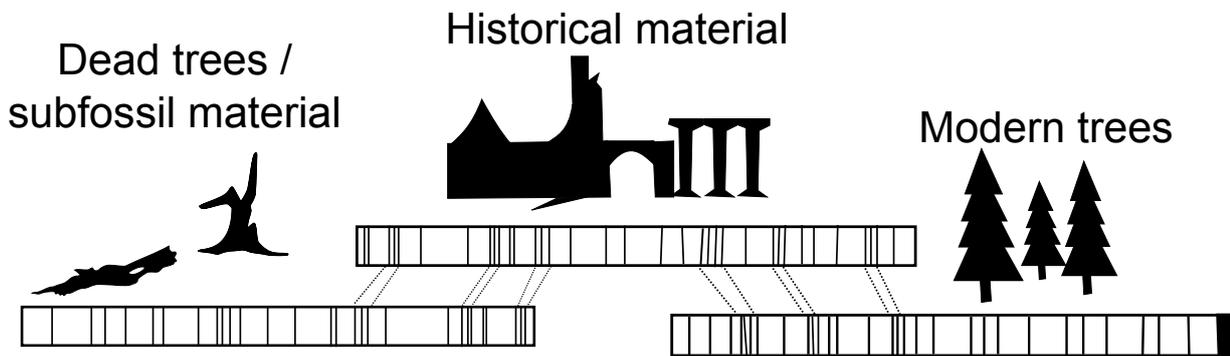


Figure 1: Schematic of matching ring-width patterns in overlapping series from different contexts.

conditions a tree can grow and reproduce under, known as the ecological amplitude of the species (Fritts, 1976). Trees can be divided into two main subgroups: angiosperms (flowering trees that include broad-leaves such as oaks and ashes), and gymnosperms (seed-producing trees that include all conifer species). Angiosperms generally require higher temperatures and precipitation totals than gymnosperms but there is considerable geographical overlap between the two (Woodward *et al.*, 2004).

In regions where distinct seasonal climate forces plants into periodic dormancy, trees tend to form annual rings (Fritts, 1976). Radial growth is influenced by a range of factors including climate, soil, topography, and competition for resources with nearby trees (Cook, 1987). Individuals growing at the same site under similar conditions should therefore display similar year-to-year variability in growth. By comparing and matching ring-width patterns between trees it is possible to cross-date each ring in a given set of samples (Stokes and Smiley, 1968). Trees growing at the edge of the species' ecological amplitude show stronger coherence in inter-annual growth and thus have higher potential for dendrochronological dating.

The outermost ring in a living tree represents the most recent growing season (if sampling a Northern Hemisphere tree in January 2015, the outermost ring will correspond to the growth of 2014) and becomes an "anchor" in time for preceding patterns of growth. Cross-dating thus allows for the assignment of absolute calendric dates for every ring in samples recovered from dead trees and timbers (Figure 1), provided that there is a temporal overlap in ring patterns and that the growth forcing common to all sampled is strong enough. In addition to sequences of narrow and wide rings, samples can also be

matched by comparing distinctive anatomical features in the wood such as frost damage and light-coloured latewood bands (Filion *et al.*, 1986; Schweingruber *et al.*, 1990).

Sampling and processing

Field methods

Material collected for dendrochronology can be divided into two main categories: increment cores and cross-sections (or parts of cross-sections) (Stokes and Smiley, 1968). The least destructive method to obtain tree-ring samples is coring the tree with an increment borer, a tool designed to extract a small core (Figure 2) without inflicting significant mechanical injury to the tree (van Mantgem and Stephenson, 2004).



Figure 2. An increment core is extracted from an *Abies magnifica* tree at Lassen Volcanic National Park, California.

In general, coring should be done at a right angle on a section of wood that does not contain distorted growth (Stokes and Smiley, 1968). The borer should be aimed at what is believed to be the pith of the tree in order to maximize the number of rings sampled and to

ease the subsequent analysis (Figure 3). More than one core is usually collected because of locally absent rings (LARs, or “missing” rings; St. George *et al.*, 2013) and extra “false” rings (Schweingruber, 1996). Depending on the objective of the study, coring close to visible injuries may be appropriate (Stoffel and Bollschweiler, 2008).

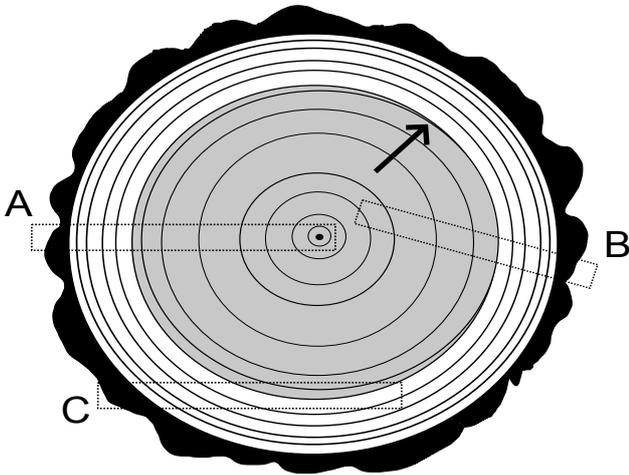


Figure 3: Placement of: (A) an increment core that includes all elements of radial growth (pith, heartwood (grey), sapwood (white), and bark); (B) a core with a locally absent ring (arrow), as well as missing pith; and (C) an off-centred core of little dendrochronological value.

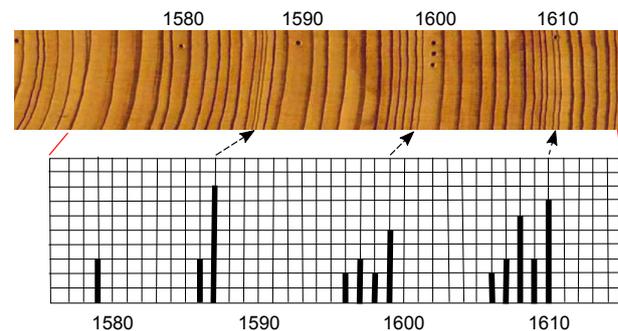
Cross-sections are taken perpendicular to the stem with a manual saw or chainsaw, they provide tree-ring information across the whole circumference of growth. Unless the tree is rotten at the centre, cross-sections also guarantee that the full number of rings (from pith to bark) is present in the sample. However, collecting cross-sections from living trees means killing the tree, and the practice should be reserved for fallen/dead trees. Furthermore, the volume and weight of cross-sections can become an issue when sampling remote locations. Taking a ‘wedge’ or ‘plunge’ cut (Arno and Sneek, 1977) is a compromise and frequently done when studying fire history through tree rings (Swetnam, 1996; Kipfmüller and Baker, 2000).

Sample preparation and cross-dating

After collection, samples are air-dried before mounted on wooden platforms for stability. The surface of cores is prepared using a scalpel or sandpaper of progressively finer grade until a fine polish is achieved. The use of a planer and belt-sander can ease preparation of large

cross-sections. For subfossil materials, rubbing powdered chalk into prepared surfaces can improve the clarity of ring boundaries (Pilcher *et al.*, 1995), and the use of rubbing alcohol can similarly help define rings in burnt samples.

A simple ring count can produce an inaccurate age for individual rings because of the possible presence of LARs and/or “false” rings in a core or cross-section, and checking samples against each other is therefore required to establish true dates. There are several approaches to cross-dating, including skeleton plotting (Stokes and Smiley, 1968), the list method (Yamaguchi, 1991), and the memorization method (Speer, 2010). Common for these techniques is the use of “event” and “pointer” years (or marker rings/years). Event years are those years for which a tree displays abnormal or conspicuous growth in terms of width or structure (Figure 4; Schweingruber *et al.* 1990). A year is labelled a pointer year when a group of trees display event years in the same year, and the pointer years aid in the assignment of dates to samples of unknown age.



*Figure 4: An example of event years (below) in a *Taxodium distichum* sample (above) from Blackwater River, Virginia. Note the narrow ring in 1587, and the sequence of five consecutive narrow rings between 1606 and 1610.*

When absolute dates have been assigned to each ring in a sample, ring-widths are measured. A sliding stage paired with a microscope is commonly used, but high-resolution scans of the core or cross-section surface can be used together with computer software such as WinDendro (Regent Instruments Canada Inc., 2009). However, a microscope is recommended when studying species with low growth rates as some finer features of the samples can be missed due to limited scanning resolution (Maxwell *et al.*, 2011).

Statistical testing of dates produced by cross-dating can be performed by analysing measurements in COFECHA (Holmes, 1983). The averaging of measurements and construction of a 'master' chronology can be done in ARSTAN (Cook 1985). There are several other computer packages for tree-ring analysis, including *dpLR* for the R language environment (Bunn, 2008) and David Meko's '[Tree-Ring MATLAB Toolbox](#)'. A master chronology can be used to date samples from dead trees (Figure 5). The frequency of LARs varies over space and species boundaries (St. George *et al.*, 2013), and for some study regions it is sufficient to count and measure rings and produce absolute dates based on statistical analysis (Baillie, 1982; Holmes, 1983). The minimum number of overlapping years required to provide an absolute date varies depending on species and strength of the common signal (Yamaguchi, 1986), where there are few overlapping rings short sequences of tree rings may cause spurious and incorrect matches (Hillam *et al.*, 1987).

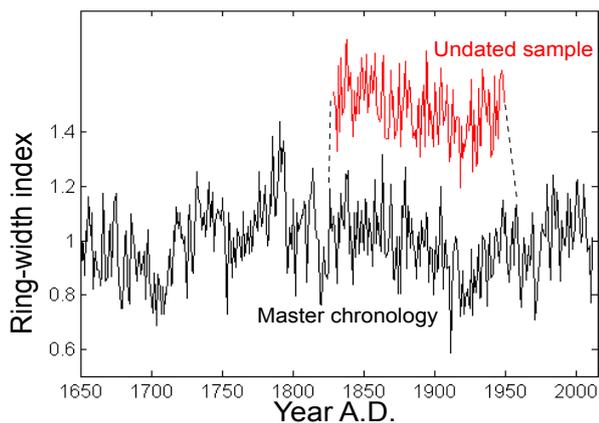


Figure 5: Alignment of a sample of unknown date (a floating chronology) with a local master chronology to produce calendar dates.

The inner-ring date of a sample only provides the age for above where the core was taken, and to identify the total tree age the number of years' growth below the core must be added (Schweingruber, 1996). Hence, coring close to the root crown is recommended in order to minimize the uncertainty of establishment dates (McCarthy *et al.*, 1991).

Applications in geomorphology

Landslide and rockfall events

The notion that trees can record evidence of landslides and debris flows dates back to the

late 19th century (McGee, 1893). Such events can cause scarring and death, but also create new surfaces for tree establishment (Shroder, 1979; Stoffel *et al.*, 2006). Furthermore, mass movements will have direct or indirect effects on the growth rate of nearby trees (Schweingruber, 1996; Stoffel *et al.*, 2006). Partial burial, loss of limbs, or decapitation will cause abrupt growth reductions, and the tilting of stems can cause eccentric growth (leading to the formation of reaction wood in the downslope side of conifers and tension wood in the upslope side of broad-leafed trees) on the tree's downslope/upslope side (Panshin and de Zeeuw, 1970; Stoffel, 2008). Trees left unharmed will benefit from the death of neighbouring trees through less competition for resources and their radial growth rate will increase (Schweingruber, 1996). Rockfall activity can have similar effects but unlike debris flows and landslides, during which large masses of material have a uniform impact on a given area, rockfalls will affect the landscape discretely (Stoffel and Perret, 2006). Some tree species respond to mechanical wounding by increased resin flow in the stem (Ruel *et al.*, 1998). Traumatic resin ducts (TRDs; Figure 6), often formed in the growth ring near an injury, can provide evidence of past debris flows and rockfall activity (Bollschweiler *et al.*, 2008), including events that are not recorded in other proxies (Stoffel, 2008). Because a range of other factors can cause TRDs (Schweingruber, 2007; Sheppard *et al.*, 2008), thresholds may need to be established in order to identify geomorphic events (Stoffel *et al.*, 2005).

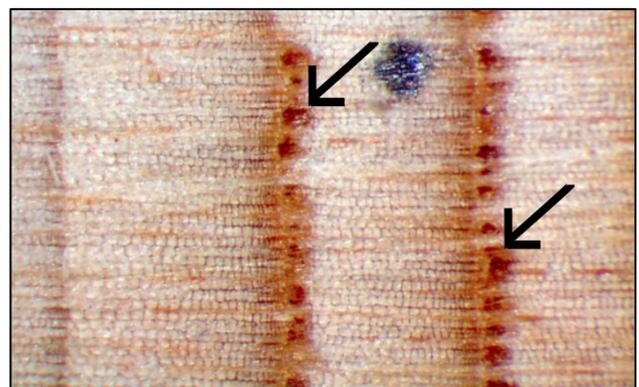


Figure 6: Traumatic resin ducts (arrows) in *Abies magnifica* from Crater Lake, Oregon. Growth left-to-right.

Tree selection and sampling strategies are particularly important for the study of landslides, debris flows, and rockfalls (Trappmann and Stoffel, 2013; Corona *et al.* 2014). Stoffel *et al.*

(2013) present an excellent guide on field sampling for studying mass movements through tree rings. Large sample sizes from a confined area will not only allow for the dating of individual events but can also provide spatial and temporal patterns of rockfall activity over time, which can in turn inform hazard management (Stoffel *et al.*, 2005; Corona *et al.*, 2014).

Glaciology

The establishment of trees can be used to date moraines and other glacial deposits (Luckman, 1988). Germinating trees from nearby areas will colonize a surface as it becomes ice-free and able to sustain vegetation (Sigafos and Hendricks, 1969). The pith of the oldest tree on a landform will therefore produce a minimum age for the most recent glacial retreat in that area (Lawrence, 1950). Sampling several locations on a glacial forefield can provide rates for ice front-recession over time (Smith *et al.*, 1995; Winchester *et al.*, 2014). However, the lag-time between glacial recession and successful tree establishment (or ecesis) must be considered (Lawrence, 1950; McCarthy and Luckman, 1993). If similar changes have occurred during periods for which satellite imagery or historical photographs exist, an ecesis interval can be estimated (McCarthy and Luckman 1993; Winchester *et al.*, 2014).

Dating glacial retreats through tree rings is often done in combination with lichenometry to minimize the uncertainties associated with either method (Karlén, 1984; Smith *et al.*, 1995; Winchester and Harrison, 2000; Wiles *et al.*, 2002; Reyes *et al.*, 2006; Trenbirth, 2010).

Together, it is postulated that the two dating methods can provide sufficiently high temporal resolution of glacial activity to allow comparison with paleoclimatic proxies (Luckman, 1993; Mood and Smith, 2015) and to aid reconstruction of past mass-balances (Wood *et al.*, 2011). Glacial advances have also been dated through dendrochronology (e.g. Nicolussi and Schlüchter, 2012).

Seismology

The impact of earthquakes on vegetation is multifaceted (Page, 1970). Trees may fall over and/or be buried by rapid sedimentation or landslides activated by the seismic activity (Jacoby, 1997). The apparently sudden death

of trees along the West Coast of United States has been linked to a large Cascadia earthquake at A.D. 1700 (Atwater and Yamaguchi, 1991; Jacoby *et al.*, 1995). In areas with extensive landslide activity, tree-ring dating can show synchronicity across space, and highlight events of great magnitude. The precision of dendroseismology allowed Fillion *et al.* (1991) to attribute two landslides in eastern Canada to a 1663 earthquake.

Abrupt changes in growth rates following an earthquake have been observed in several studies across different forest ecosystem (Jacoby *et al.*, 1988; Veblen *et al.*, 1992; Bekker, 2004). The ecology and topographic setting of the species studied is of importance because trees can have opposite excursions of growth in response to the same event (Stahle *et al.*, 1992).

Volcanic events

Tree rings have been used to date major and minor volcanic eruptions for periods when observations are sparse or non-existent. During large events, gases and dust particles from the volcano will reach the atmosphere and influence climate on a global scale by limiting incoming solar radiation (Rampino and Self, 1982). Subsequent cooling can affect trees either over a full growing season causing low growth over the year or through frost damage caused by a late frost at any point of the growing season (Figure 7).

LaMarche and Hirschboeck (1984) recorded frost rings in *Pinus longaeva* trees in the western United States during years of known eruptions. Using tree-ring density chronologies, Briffa *et al.* (1998) were able to show the impact of major volcanic events on periods of cooling in the Northern Hemisphere since the 1400's. The synchronous effect on high latitude or high altitude trees has since been demonstrated for the past 5000 years, with the tree-ring record displaying high correlation with volcanic events identified in the Greenland ice cores (Salzer and Hughes, 2007).

Eruptions with less pronounced impacts can also be dated if local samples exist (Biondi *et al.*, 2003). Subfossil materials from Kamchatka, Russia, have been used to date an eruption of the Shiveluch volcano to 1756-58 (Solomina *et al.*, 2008). Pyroclastic surges caused by the 1842-43 eruption of Mount St. Helens,

Washington, are thought to explain the anomalous series of narrow rings in nearby trees (Yamaguchi and Lawrence, 1993). In addition to ring-widths, variations in the chemistry of individual growth rings may also contain information on past volcanic activity (Pearson *et al.* 2005). Cinder cone eruptions have shown to increase Na, S, and P in rings of adjacent trees (Sheppard *et al.*, 2008; 2009). Thus, the chemical composition of wood can be used to strengthen the interpretation of ring-width anomalies and dating produced by cross-dating, or even to identify the source of eruption (Pearson *et al.*, 2009).

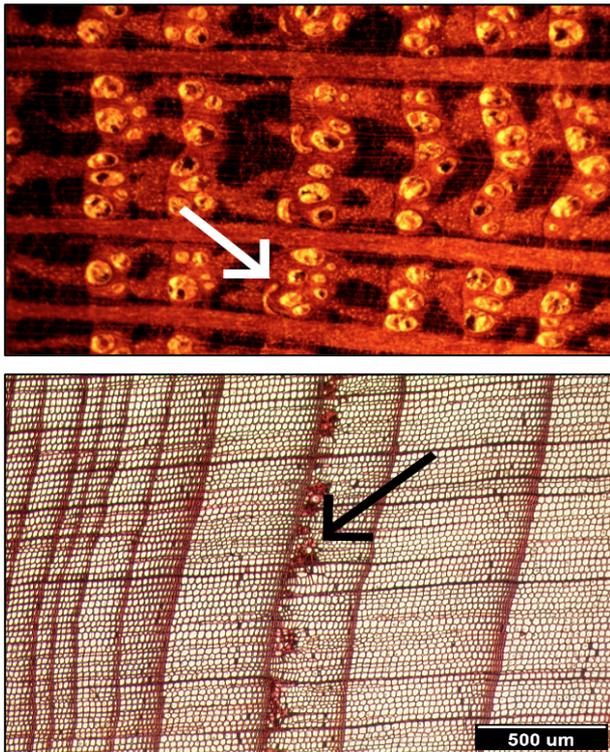


Figure 7: Frost damage (arrows) during the early growing season of year 1826 in *Quercus alba* from Arkansas (top), and *Juniperus virginiana* from Oklahoma (bottom). Growth left-to-right. Photos courtesy of the Tree-Ring Laboratory, University of Arkansas.

Flooding

Flood events can cause distinct anatomical features in the wood of rings formed during inundation (Schweingruber, 2007). Yanosky (1983; 1984) studied *Fraxinus americana* and *F. pennsylvanica* on the Potomac River, Washington D.C., and noted that there were considerably larger cells in the latewood of 1972 than in any other year. This year

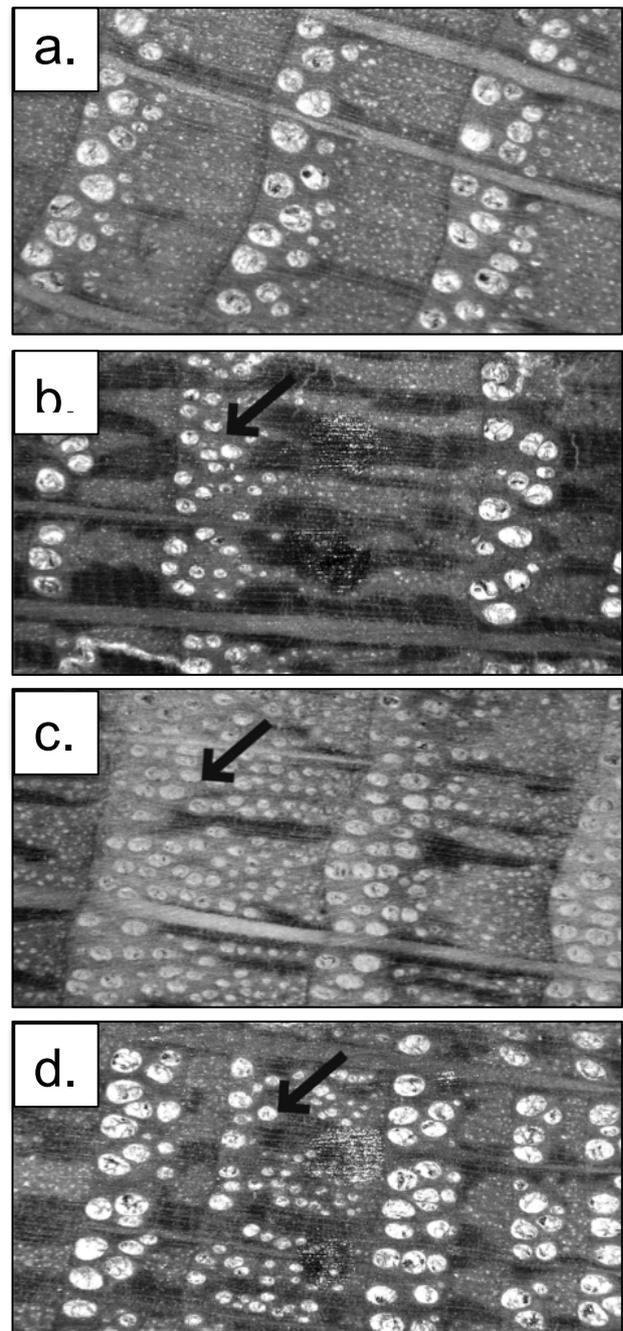


Figure 8. (a) Normal growth rings in *Quercus macrocarpa*, with large vessels in the earlywood, (b) shrunken earlywood vessels, (c) earlywood vessels extending into the latewood, and (d) both shrunken and extending vessels. Growth left-to-right. Photos courtesy of Dr. Scott St. George.

coincided with the third largest flood on record, and as Yanosky looked further back in time the growth features occurred in many other known flood years and they were therefore dubbed 'flood rings'. More recently, anatomical signatures of flood in *Quercus macrocarpa* have been identified, including shrunken earlywood vessels and earlywood vessels extending into the latewood (St. George *et al.*,

2002; Wertz *et al.*, 2013; Figure 8). The use of older wood material from historical buildings and subfossil deposits has allowed the reconstruction of a flood history for Red River, Manitoba, extending further back in time than observational data (St. George and Nielsen, 2000). Sampling trees from different locations can also produce a spatial reconstruction of flood extent. Although the most successful studies have used ring-porous broad-leaf species, conifer softwoods are also known to exhibit features attributable to floods (Yamamoto *et al.*, 1987; Ballesteros *et al.*, 2010).

Caveats and considerations

Although the vast majority of tree species outside the tropics form annual rings, not all show coherent inter-annual growth variability across space and these are therefore not suitable for cross-dating exercises. An extensive list of species and their usefulness in dendrochronology is provided by Grissino-Mayer (1993). The temporal range to which trees can be dated also differs depending on species and region. In areas where subfossil or archaeological materials do not exist, the germination date of the oldest living trees becomes the limit to which a sample can be dated. Although the trees of some species can live for several millennia (Brown, 1996), tree-ring chronologies that date back over 1000 years are rare outside the southwestern United States (St. George, 2014). Over 2500 tree-ring chronologies from across the world are publically available through the International Tree-Ring Data Bank (Grissino-Mayer and Fritts, 1997). However, tree rings can still provide valuable chronological information even when there are no possibilities of calendric dates. Cross-dating between samples can indicate synchronicity over space, and samples of annual tree-growth can facilitate the 'wiggles-matching' of radiocarbon dates (Bronk Ramsey, 2009).

Some samples may not contain every ring from pith to bark because of poor preservation or sampling difficulties. This can be detrimental when assigning death and/or germination dates of trees. Where a core does not reach the pith, the curvature of the innermost rings can be used to estimate pith position and the number of rings not present (Applequist, 1958; Wong and Lertzman, 2001). If bark is not present,

death dates can be estimated based on the number of sapwood rings still attached to the sample (Baillie, 1982; Hillam *et al.*, 1987). Sapwood is more susceptible to decay over time (Panshin and de Zeeuw, 1970), thus the outer rings of trees were sometimes removed from timber used in historic buildings (Baillie, 1982) and in cases where no sapwood is present it is only possible to produce a *terminus post quem*.

Conclusions

Shared year-to-year variability in ring-width allows for the assignment of calendric or cross-dates to sequences of tree rings. Dendrochronology does not only offer the potential of unrivalled dating precision but can also inform researchers about the spatial extent of past landscape change. The temporal range for which trees can be useful varies between study areas but trees grow to be over 200 years old in most regions. When subfossil or historical material is available it is possible to use tree rings for dating on considerably longer timescales. Because signs of geomorphic processes recorded in trees can be caused by several different factors, it is recommended to use multiple lines of evidence, including death and establishment, abrupt changes in growth rate, anatomical features, and in some cases chemical composition. Correctly used and in combination with other sources of information, dendrogeomorphology can be a powerful tool for understanding the processes that form the landscapes around us.

References

- Applequist MB. 1958. A simple pith locator for use with off-center increment cores. *Journal of Forestry* **56**: 141.
- Arno SF, Sneek KM. 1977. *A method for determining fire history in conifer forests in the mountain west*. US Department of Agriculture, Forest Service General Technical Report INT-42, Intermountain Forest and Range Experiment Station, Ogden, Utah.
- Atwater BF, Yamaguchi DK. 1991. Sudden, probably coseismic submergence of Holocene trees and grass in coastal Washington State. *Geology* **19**: 706-709.
- Ballesteros JA, Stoffel M, Bollschweiler M, Bodoque JM, Diez-Herrero A. 2010. Flash-flood impacts cause changes in wood anatomy of

- Alnus glutinosa, Fraxinus angustifolia and Quercus pyrenaica. *Tree Physiology* **30**: 773-781.
- Baillie MGL. 1982. *Tree-ring dating and archaeology*. Croom Helm, London, UK.
- Bekker MF. 2004. Spatial variation in the response of tree rings to normal faulting during the Hebgen Lake Earthquake, Southwestern Montana, USA. *Dendrochronologia* **22**: 53-59.
- Biondi F, Galindo Estrada I, Gavilanes Ruiz J C, Torres AE. 2003. Tree growth response to the 1913 eruption of Volcán de Fuego de Colima, Mexico. *Quaternary Research* **59**: 293-299.
- Bollschweiler M, Stoffel M, Schneuwly DM, Bourqui K. 2008. Traumatic resin ducts in Larix decidua stems impacted by debris flows. *Tree Physiology* **28**: 255-263.
- Briffa KR, Jones PD, Schweingruber FH, Osborn TJ. 1998. Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* **393**: 450-455.
- Bronk Ramsey C. 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**: 337-360.
- Brown PM. 1996. OLDLIST: A database of maximum tree ages. In: Dean JS, Meko DM & Swetnam TW (Eds.) *Tree Rings, Environment, and Humanity*. Radiocarbon, University of Arizona, Tucson, AZ.
- Buckley BM, Anchukaitis KJ, Penny D, Fletcher R, Cook ER, Sano M, Nam LC, Wichienkeo A, Minh TT, Hong TM. 2010. Climate as a contributing factor in the demise of Angkor, Cambodia. *Proceedings of the National Academy of Science* **107**: 6748-6752.
- Bunn AG. 2008. A dendrochronology program library in R (dplR). *Dendrochronologia* **26**: 115-124.
- Cook ER. 1987. The decomposition of tree-ring series for environmental studies. *Tree-Ring Bulletin* **47**: 37-59.
- Cook ER. 1985. A time series analysis approach to tree-ring standardization. Ph.D. dissertation. The University of Arizona, Tucson, AZ.
- Cook ER, Meko DM, Stahle DW, Cleaveland MK. 1999. Drought reconstructions of the continental United States. *Journal of Climate* **12**: 1145-1162.
- Corona C, Lopez Saez J, Stoffel M. 2014. Defining optimal sample size, sampling design and thresholds for dendrogeomorphic landslide reconstructions. *Quaternary Geochronology* **22**: 72-84.
- Douglass AE. 1919. *Climatic cycles and tree growth*. Vol. I. Carnegie Institute of Washington, Washington, D.C.
- Esper J, Cook ER, Schweingruber FH. 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* **295**: 2250-2253.
- Filion L, Quinty F, Bégin C. 1991. A chronology of landslide activity in the valley of Rivière du Gouffre, Charlevoix, Quebec. *Canadian Journal of Earth Sciences* **28**: 250-256.
- Filion L, Payette S, Gauthier L, Boutin Y. 1986. Light rings in subarctic conifers as a dendrochronological tool. *Quaternary Research* **26**: 272-279.
- Friedrich M, Remmele S, Kromer B, Hofmann J, Spurk M, Kaiser KF, Orcel C, Küppers M. 2004. The 12,460-year Hohenheim oak and pine tree-ring chronology from central Europe – a unique annual record for radiocarbon calibration and paleoenvironment reconstructions. *Radiocarbon* **46**: 1111-1122.
- Fritts HC. 1976. *Tree rings and climate*. Academic Press, London, UK.
- Grissino-Mayer HD. 1993. An updated list of species used in tree-ring research. *Tree-Ring Bulletin* **53**: 16-42.
- Grissino-Mayer HD, Fritts HC. 1997. The International Tree-Ring Data Bank: an enhanced global database serving the global scientific community. *The Holocene* **7**: 235-238.
- Hillam J, Morgan RA, Tyers I. 1987. Sapwood estimates and the dating of short ring sequences. *British Archaeological Reports* **17**: 165-185.
- Holmes RL. 1983. Computer-assisted quality control in tree-ring dating and measurement. *Tree-Ring Bulletin* **43**: 69-78.
- Jacoby GC. 1997. Application of tree ring analysis to paleoseismology. *Review of Geophysics* **35**: 109-124.
- Jacoby GC, Carver G, Wagner W. 1995. Trees and herbs killed by an earthquake ~300 yr ago at Humboldt Bay, California. *Geology* **23**: 77-80.

- Jacoby GC, Sheppard PR, Sieh KE. 1988. Irregular recurrence of large earthquakes along the San Andreas Fault: Evidence from trees. *Science* **241**: 196-199.
- Karlén W. 1984. Dendrochronology, mass balance and glacier front fluctuations in northern Sweden. In: Mörner N-A & Karlén W (Eds.) *Climatic changes on a yearly to millennial basis: Geological, historical and instrumental records*. Springer, New York, NY.
- Kipfmüller KF, Baker WL. 2000. A fire history of a subalpine forest in south-eastern Wyoming, USA. *Journal of Biogeography* **27**: 71-85.
- LaMarche Jr. VC, Hirschboeck KK. 1984. Frost rings in trees as records of major volcanic eruptions. *Nature* **307**: 121-126.
- Lawrence DB. 1950. Estimating dates of recent glacier advances and recession rates by studying tree growth layers. *Eos, Transactions, American Geophysical Union* **31**: 243-248.
- Luckman, BH. 1993. Glacier fluctuations and tree-ring records for the last millennium in the Canadian Rockies. *Quaternary Science Reviews* **12**: 441-450.
- Luckman BH. 1988. Dating the moraines and recession of Athabasca and Dome Glaciers, Alberta, Canada. *Arctic and Alpine Research* **20**: 40-54.
- Koprowski M, Winchester V, Zielski A. 2010. Tree reactions and dune movements: Slowinski National Park, Poland. *CATENA* **81**: 55-65.
- Maxwell RS, Wixom JA, Hessler AE. 2011. A comparison of two techniques for measuring and crossdating tree rings. *Dendrochronologia* **29**: 237-243.
- McCarthy DP, Luckman BH. 1993. Estimating ecesis for tree-ring dating of moraines: a comparative study from the Canadian Cordillera. *Arctic and Alpine Research* **25**: 63-68.
- McCarthy DP, Luckman BH, Kelley PE. 1991. Sampling height-age error correction for spruce seedlings in glacial forefields, Canadian Cordillera. *Arctic and Alpine Research* **23**: 451-455.
- McGee WJ. 1893. A fossil earthquake. *Geological Society of America Bulletin* **4**: 411-414.
- Mood BJ, Smith DJ. 2015. Latest Pleistocene and Holocene behaviour of Franklin Glacier, Mt. Waddington area, British Columbia Coast Mountains, Canada. *The Holocene* **25**: 784-794.
- Nicolussi K, Schlüchter C. 2012. The 8.2 ka event – Calendar-dated glacier response in the Alps. *Geology* **40**: 819-822.
- Page R. 1970. Dating episodes of faulting from tree rings: Effects of the 1958 rupture of the Fairweather Fault on tree growth. *Geological Society of America Bulletin* **81**: 3085-3094.
- Pan Y, Birdsey RA, Philips OL, Jackson RB. 2013. The structure, distribution, and biomass of the world's forests. *Annual Review of Ecology, Evolution, and Systematics* **44**: 593-622.
- Panshin AJ, de Zeeuw C. 1970. *Textbook of Wood Technology*. McGraw-Hill, New York, NY.
- Pearson CL, Dale DS, Brewer PW, Kuniholm PI, Lipton J, Manning SW. 2009. Dendrochemical analysis of a tree-ring growth anomaly associated with the Late Bronze Age eruption of Thera. *Journal of Archaeological Science* **36**: 1206-1214.
- Pearson C, Manning SW, Coleman M, Jarvis K. 2005. Can tree-ring chemistry reveal absolute dates for past volcanic eruptions? *Journal of Archaeological Science* **32**: 1265-1274.
- Pilcher JR, Baillie MGL, Brown DM, McCormac FG, MacSweeney PB, McLawrence AS. 1995. Dendrochronology of subfossil pine in the north of Ireland. *Journal of Ecology* **83**: 665-671.
- Rampino MR, Self S. 1982. Historic eruptions of Tambora (1815), Krakatau (1883), and Agung (1963), their stratospheric aerosols, and climatic impact. *Quaternary Research* **18**: 127-143.
- Regent Instruments Canada Inc. 2009. WINDENDRO for Tree-ring Analysis.
- Reyes AV, Wiles GC, Smith DJ, Barclay DJ, Allen S, Jackson S, Larocque S, Laxton S, Lewis D, Calkin PE, Clague JJ. 2006. Expansion of alpine glaciers in Pacific North America in the first millennium AD. *Geology* **34**: 57-60.
- Ruel JJ, Ayers MP, Lorio PL. 1998. Loblolly pine responds to mechanical wounding with increased resin flow. *Canadian Journal of Forest Research* **28**: 596-602.
- St. George S. 2014. An overview of tree-ring width records across the Northern Hemisphere. *Quaternary Science Reviews* **95**: 132-150.

- St. George S, Nielsen E. 2000. Signatures of high-magnitude 19th century floods in *Quercus macrocarpa* (Michx.) tree rings along the Red River, Manitoba, Canada. *Geology* **28**: 899-902.
- St. George S, Ault TR, Torbenson MCA. 2013. The rarity of absent growth rings in Northern Hemisphere forests outside the American Southwest. *Geophysical Research Letters* **40**: 3727-3731.
- St. George S, Nielsen E, Conciatori F, Tardif J. 2002. Trends in *Quercus macrocarpa* vessel areas and their implications for tree-ring paleoflood studies. *Tree-Ring Research* **58**: 3-10.
- Salzer MW, Hughes MK. 2007. Bristlecone pine tree rings and volcanic eruptions over the last 5000 yr. *Quaternary Research* **67**: 57-68.
- Salzer MW, Hughes MK, Bunn AG, Kipfmüller KF. 2009. Recent unprecedented tree-ring growth in bristlecone pine at the highest elevations and possible causes. *Proceedings of the National Academy of Science* **106**: 20348-20353.
- Schweingruber FH. 2007. *Wood structure and environment*. Springer-Verlag, Berlin.
- Schweingruber FH. 1996. *Tree rings and environment: dendroecology*. Paul Haupt, Bern, Switzerland.
- Schweingruber FH, Eckstein D, Serre-Bachet F, Bräker OU. 1990. Identification, presentation and interpretation of event years and pointer years in dendrochronology. *Dendrochronologia* **8**: 9-38.
- Sheppard PR, Ort MH, Anderson KC, Clynne MA, May EM. 2009. Multiple dendrochronological responses to the eruption of Cinder Cone, Lassen Volcanic National Park, California. *Dendrochronologia* **27**: 213-221.
- Sheppard PR, Ort MH, Anderson KC, Elson MD, Vázquez-Selem L, Clemens AW, Little NC, Speakman RJ. 2008. Multiple dendrochronological signals indicate the eruption of Parícutin volcano, Michoacán, Mexico. *Tree-Ring Research* **64**: 97-108.
- Shroder JF. 1979. Dendrogeomorphological analysis of mass movement on Table Cliffs Plateau, Utah. *Quaternary Research* **9**: 168-185.
- Sigafoos RS, Hendricks EL. 1969. The time interval between stabilization of alpine glacial deposits and establishment of tree seedlings. *United States Geological Survey Professional Paper* **650B**, 89-93.
- Smith DJ, McCarthy DP, Colenutt ME. 1995. Little Ice Age glacial activity in Peter Lougheed and Elk Lakes provincial parks, Canadian Rocky Mountains. *Canadian Journal of Earth Sciences* **32**: 579-589.
- Solomina O, Pavlova I, Curtis A, Jacoby GC, Ponomareva V, Pevzner M. 2008. Constraining recent Shiveluch volcano eruptions (Kamchatka, Russia) by means of dendrochronology. *Natural Hazards and Earth System Sciences* **8**: 1083-1097.
- Speer JH. 2010. *Fundamentals of Tree-ring Research*. The University of Arizona Press, Tucson, AZ.
- Stahle DW, van Arsdale RB, Cleaveland MK. 1992. Tectonic signal in baldcypress trees at Reelfoot Lake, Tennessee. *Seismological Research Letters* **63**: 439-447.
- Stallings Jr. WS, Schulman E, Douglass AE. 1937. Some early papers on tree-rings: I. J. Keuchler. II. J.C. Kapteyn. *Tree-Ring Bulletin* **3**: 27-29.
- Stoffel M. 2008. Dating past geomorphic processes with tangential rows of traumatic resin ducts. *Dendrochronologia* **26**: 53-60.
- Stoffel M, Bollschweiler M. 2008. Tree-ring analysis in natural hazards research – an overview. *Natural Hazards Earth System Sciences* **8**: 187-202.
- Stoffel M, Perret S. 2006. Reconstructing past rockfall activity with tree rings: Some methodological considerations. *Dendrochronologia* **24**: 1-15.
- Stoffel M, Butler DR, Corona C. 2013. Mass movements and tree rings: a guide to dendrogeomorphic field sampling and dating. *Geomorphology* **200**: 106-120.
- Stoffel M, Bollschweiler M, Hassler G-R. 2006. Differentiating past events on a cone influenced by debris-flow and snow avalanche activity – a dendrogeomorphological approach. *Earth Surface Processes and Landforms* **31**: 1424-1437.
- Stoffel M, Schneuwly D, Bollschweiler M, Lièvre I, Delaloye R, Myint M, Monbaron M. 2005. Analyzing rockfall activity (1600-2002) in a protection forest – a case study using dendrogeomorphology. *Geomorphology* **68**: 224-241.

- Stokes MA, Smiley TL. 1968. *An introduction to tree-ring dating*. University of Chicago Press, Chicago, IL.
- Strunk H. 1997. Dating geomorphological processes using dendrogeomorphological methods. *Catena* **31**: 137-151.
- Studhalter RA. 1956. Early history of crossdating. *Tree-Ring Bulletin* **21**: 31-35.
- Stuiver M, Becker B. 1993. High-precision decadal calibration of the radiocarbon time scale, AD 1950-6000 BC. *Radiocarbon* **35**: 35-65.
- Swetnam TW. 1996. Fire and climate history in the Central Yenisey region, Siberia. In: Goldammer JG & Furyaev VV (Eds.) *Fire in ecosystems of Boreal Eurasia*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Swetnam TW, Betancourt JL. 1998. Mesoscale disturbance and ecological response to decadal climate variability in the American Southwest. *Journal of Climate* **11**: 3128-3148.
- Trappmann D, Stoffel M. 2013. Counting scars on tree stems to assess rockfall hazards: a low effort approach, but how reliable? *Geomorphology* **180-181**: 180-186.
- Trenbirth HE. 2010. Section 4.2.7: Lichenometry. In: Clarke LE & Nield JM (Eds.) *Geomorphological Techniques* (Online Edition). British Society for Geomorphology, London, UK.
- van Mantgem PJ, Stephenson NL. 2004. Does coring contribute to tree mortality? *Canadian Journal of Forest Research* **34**: 2394-2398.
- Veblen TT, Kitzberger T, Lara A. 1992. Disturbance and forest dynamics along a transect from Andean rain forest to Patagonian shrubland. *Journal of Vegetation Science* **3**: 507-520.
- Visser RM. 2015. Imperial timber? Dendrochronological evidence for large-scale road building along the Roman limes in the Netherlands. *Journal of Archaeological Science* **53**: 243-254.
- Wertz EL, St. George S, Zeleznik JD. 2013. Vessel anomalies in *Quercus macrocarpa* tree rings associated with recent floods along the Red River of the North, United States. *Water Resources Research* **49**: 630-634.
- Wiles GC, Jacoby GC, Davi NK, McAllister RP. 2002. Holocene glacier fluctuations in the Wrangell Mountains, Alaska. *Geological Society of America Bulletin* **114**: 896-908.
- Winchester V, Harrison S. 2000. Dendrochronology and lichenometry: Colonization, growth rates and dating of geomorphological events on the east side of the North Patagonian Icefield, Chile. *Geomorphology* **34**: 181-194.
- Winchester V, Sessions M, Valdivia Cerda J, Wünderlich O, Clemmens S, Glasser NF, Nash M. 2014. Post-1850 changes in Glacier Benito, North Patagonian Icefield, Chile. *Geografiska Annaler* **96**: 43-59.
- Wong CM, Lertzman KP. 2001. Errors in estimating tree age: implications for studies of stand dynamics. *Canadian Journal of Forest Research* **31**: 1262-1271.
- Wood LJ, Smith DJ, Demuth MN. 2011. Extending the Place Glacier mass-balance record to AD 1585, using tree rings and wood density. *Quaternary Research* **76**: 305-313.
- Woodward FI, Lomas MR, Kelly CK. 2004. Global climate and the distribution of plant biomes. *Philosophical Transactions of the Royal Society B* **359**: 1465-1476.
- Yamaguchi DK. 1991. A simple method for cross-dating increment cores from living trees. *Canadian Journal of Forest Research* **21**: 414-416.
- Yamaguchi DK. 1986. Interpretation of cross correlation between tree-ring series. *Tree-Ring Bulletin* **46**: 47-54.
- Yamaguchi DK, Lawrence DB. 1993. Tree-ring evidence for 1842-1843 eruptive activity at the Goat Rocks dome, Mount St. Helens, Washington. *Bulletin of Volcanology* **55**: 264-272.
- Yamamoto F, Kozłowski TT, Wolter KE. 1987. Effects of flooding on growth, stem anatomy, and ethylene production of *Pinus haepensis* seedlings. *Canadian Journal of Forest Research* **17**: 69-79.
- Yanosky TM. 1984. Documentation of high summer flows on the Potomac River from the wood anatomy of ash trees. *Water Resources Bulletin* **20**, 241-250.
- Yanosky TM. 1983. *Evidence of floods on the Potomac River from anatomical abnormalities in the wood of flood-plain trees*. US Geological Survey Professional Paper v. 1296.

Cosmogenic nuclide analysis

Christopher M. Darvill¹

¹ Department of Geography, Durham University, South Road, Durham, DH1 3LE, UK
(christopher.darvill@durham.ac.uk)



ABSTRACT: Cosmogenic nuclides can be used to directly determine the timing of events and rates of change in the Earth's surface by measuring their production due to cosmic ray-induced reactions in rocks and sediment. The technique has been widely adopted by the geomorphological community because it can be used on a wide range of landforms across an age range spanning hundreds to millions of years. Consequently, it has been used to successfully analyse exposure and burial events; rates of erosion, denudation and uplift; soil dynamics; and palaeo-altimetric change. This paper offers a brief outline of the theory and application of the technique and necessary considerations when using it.

KEYWORDS: Cosmogenic nuclide, dating, chronology, landscape change, Quaternary

Introduction

Geochronology allows the quantification of rates of landscape change and timing of geomorphic events, bridging the gap between geomorphological evidence and environmental or climatic variability over time. Cosmogenic nuclide analysis involves the measurement of rare isotopes which build-up in rock minerals predictably over time due to bombardment of the upper few metres of the Earth's surface by cosmic rays. In this way, the exposure, burial and altimetric change of surficial rocks and sediments can be assessed. The six most commonly used cosmogenic isotopes, ¹⁰Be, ²⁶Al, ³⁶Cl, ¹⁴C, ³He and ²¹Ne, have allowed dating on the scale of hundreds to millions of years, and they can be used to address a wide range of geomorphological problems due to their production in commonly occurring minerals (Dunai, 2010). Unlike some other techniques, cosmogenic nuclide analysis can be applied directly to the rock or sediment in question rather than providing indirect bracketing information, and can be applied to numerous environmental situations. Consequently, the technique has been enthusiastically adopted

by the geomorphological community in addressing issues such as the timing of glacial advances, fault-slip rates, bedrock/basin erosion and sediment burial (Cockburn and Summerfield, 2004). However, there are a number of practical and theoretical concerns that need to be considered when applying the technique.

This chapter will briefly explain how cosmogenic nuclides are produced, describe the range of geomorphological applications and highlight some of the practical and theoretical concerns that should be considered. The aim is to help the reader understand if the technique is applicable to their own study; it is a short introduction to a complex and continually-developing subject. For more detail, there exist several extensive reviews of the technique and literature, notably: Gosse and Phillips (2001); Cockburn and Summerfield (2004); Ivy-Ochs and Kober (2008); Dunai (2010); Granger *et al.* (2013). For glacial studies see Balco (2011); burial dating see Granger (2006) and Dehnert and Schlüchter (2008); and landscape denudation studies see Granger (2007).

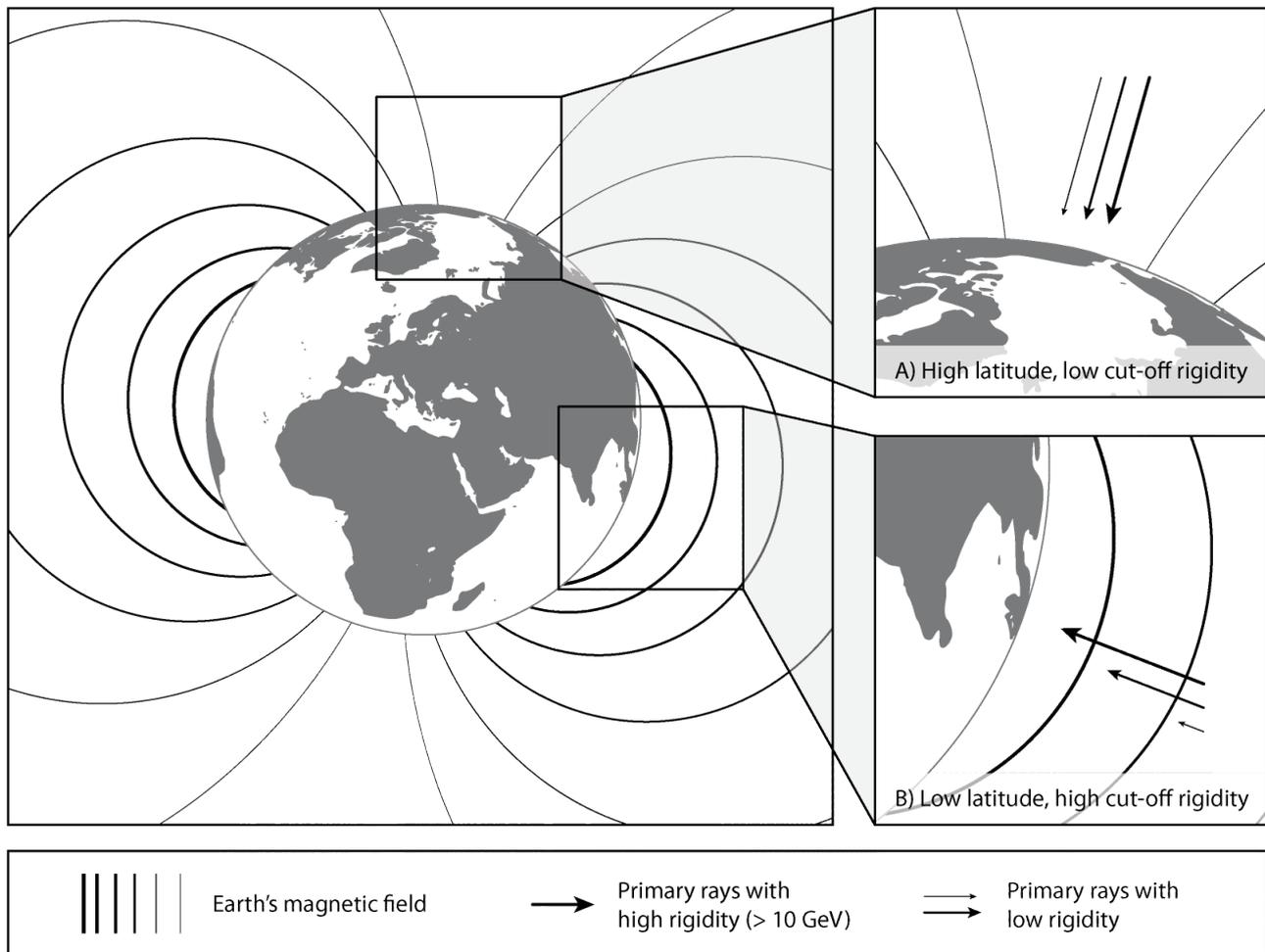


Figure 1: An illustration of the Earth's geomagnetic field and the effect it has on incoming primary cosmic rays. All rays must exceed the cut-off rigidity at a particular latitude (A and B), which is higher towards the equator (B) and lower towards the poles (A). Only rays with higher cut-off rigidities are permitted to enter the atmosphere at the equator, whilst at the poles the convergence of magnetic field lines means that most rays can enter. The consequence is higher production rates of cosmogenic nuclides at higher latitudes.

Cosmogenic nuclide production

The six most commonly used cosmogenic nuclides are ^3He and ^{21}Ne (stable, noble gas isotopes), and ^{10}Be , ^{26}Al , ^{36}Cl and ^{14}C (radioactive isotopes). Others, such as ^{38}Ar , ^{53}Mn and ^{41}Ca can be used but are not as ubiquitously applicable or require further development (Dunai, 2010). The use of any of these nuclides as a dating technique or for erosion/denudation rate studies relies on the accumulation of the isotopes within target minerals, moderated by surface erosion and, in the case of radioactive isotopes, radioactive decay. Their production involves three main stages: the production of (i) primary cosmic rays, (ii) secondary cosmic rays and (iii) nuclides via nucleonic spallation and muogenic reactions (Gosse and Phillips, 2001).

Primary cosmic rays and cut-off rigidity

The Earth is constantly bombarded by cosmic radiation in the form of primary cosmic rays. Most of these are galactic in origin, composed of high energy particles (0.1 – 10 GeV) in the form of protons (87%), α -particles (12%) and heavy nuclei (1%) (Masarik and Beer, 1999). The primary rays can also be of solar origin, with lower energies (< 100 MeV), but these only produce cosmogenic nuclides in the upper atmosphere or during intense solar activity (Masarik and Reedy, 1995).

Galactic primary rays are affected by the Earth's magnetic field and, to a lesser extent, solar activity (Lifton *et al.*, 2005). The *rigidity* of incoming rays relates to the degree to which their momentum is deflected by the Earth's magnetic field, and they must exceed a minimum rigidity (or *cut-off rigidity*) in order

to penetrate into the atmosphere. Rigidity is affected by the angle of incidence of the rays and their location relative to geomagnetic field lines, so the cut-off rigidity varies with latitude. For primary particles, this is highest at the equator (fewer rays can penetrate) and lowest at the poles (more rays can penetrate; Dunai, 2010; Figure 1).

Significant variations in long-term primary ray production (1000's – 100,000's of years) affect cosmic radiation flux at the Earth's surface and are caused by changes in the Earth's magnetic field (dominantly the *geocentric axial dipole*, but also *non-dipole components*; Dunai, 2001; Dunai, 2010). This can be reconstructed using proxy records and accounted-for in scaling models (Guyodo and Valet, 1999; Dunai, 2001; Desilets and Zreda, 2003; Lifton *et al.*, 2005). Short-term cyclical changes in solar activity (10's – 100's of years) also affect low energy galactic primary rays, though this is essentially constant over timescales of thousands to millions of years. This modulation generally only affects primary ray flux at high latitudes where cut-off

rigidity is low, but has implications for scaling models (see 'Production rates and scaling factors' section; and Lifton *et al.*, 2005).

Secondary cosmic rays and attenuation

Primary rays that penetrate the Earth's magnetic field trigger a reaction with atmospheric gas nuclei that results in a cascade of secondary cosmic rays, composed of high energy nucleons (e.g. protons and neutrons) and mesons (e.g. kaons, pions and muons; Figure 2; and Gosse and Phillips, 2001). Because the production of secondary rays triggers further collisions and interactions and results in a scattering and absorption of energy, the secondary ray intensity decreases down through the atmosphere (roughly exponentially with increasing atmospheric depth below 100 g cm^{-2} ; Gosse and Phillips, 2001). This decrease in intensity is called *attenuation* and varies with the density of material through which the secondary rays pass. The *attenuation length* changes according to the energy of the incoming

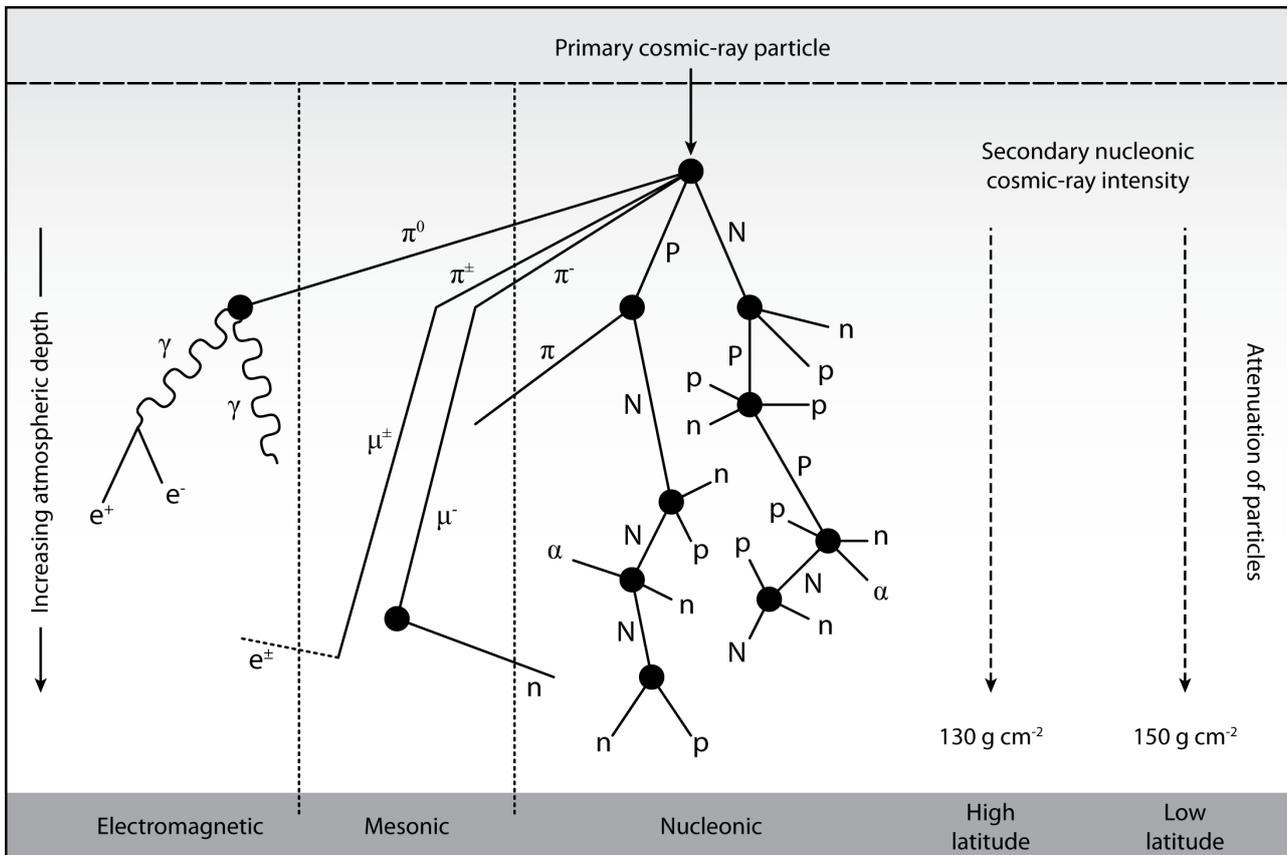


Figure 2: The cascade of secondary ray particles produced in the atmosphere by incoming primary cosmic rays (adapted from Dunai, 2010). Cosmic rays have greater attenuation lengths at the equator (150 g cm^{-2}) than at the poles (130 g cm^{-2} ; see discussion in the text).

Table 1: The six major cosmogenic nuclides used in geomorphological research, their target elements and the reaction pathways through which they are formed. From Ivy-Ochs & Kober (2008) and Dunai (2010).

Isotope	Main target elements	Reactions
^3He	Many, including Li	Spallation (100%) Thermal neutron capture (on Li via ^3H)
^{21}Ne	Mg, Al, Si	Spallation (>96.4%)
^{10}Be	O, Si	Spallation (96.4%) Muons (3.6%)
^{26}Al	Si	Spallation (95.4%) Muons (4.6%)
^{36}Cl	K, Ca, Cl (Fe, Ti)	Spallation (95.4% from K; 86.6% from Ca; 100% from Fe and Ti) Thermal neutron capture (from Cl and K) Muons (4.6% from K; 13.4% from Ca;)
^{14}C	O, Si	Spallation (82%) Muons (18%)

cosmic rays and thus varies with latitude. At lower latitudes, the rays have higher energies (because the cut-off rigidity is higher) so must pass through more of the atmosphere to reduce the cosmic ray flux. Consequently, attenuation length varies between roughly 130 g cm^{-2} at high latitudes and 150 g cm^{-2} at low latitudes (Dunai, 2000).

Thus, before cosmic rays interact with the Earth's surface, they have already been significantly influenced by the geomagnetic field (primary rays), latitude (primary and secondary rays), and altitude (secondary rays). This has important implications for calculations of cosmogenic nuclide production (see 'Analysis and age determination' section).

Spallation, capture and muonic reactions

On encountering the Earth's surface, secondary ray particles are also attenuated with depth according to rock density and their attenuation length. These particles trigger a number of reactions in target minerals that can result in the production of cosmogenic nuclides (see Table 1 and also Figure 5). Spallation reactions involve fast and high energy neutrons and produce most

cosmogenic nuclides (^3He , ^{21}Ne , ^{10}Be , ^{26}Al , ^{14}C , ^{36}Cl), but neutron flux attenuates to <1% below 3 m from the surface (Figure 3). Thermal neutrons result from a slowing-down of some neutrons during the atmospheric cascade, and trigger capture reactions which can be important for the production of some nuclides, particularly ^{36}Cl if natural Cl is available. Because these low energy neutrons can leak back out of a rock, thermal neutron flux peaks at roughly 20 cm from the surface; an important consideration if sampling for ^{36}Cl . Muons have a lower mass than neutrons and are not highly reactive, so can penetrate much deeper into rock. Consequently, they account for roughly 2.5% of cosmogenic reactions at the surface, but 100% of reactions below 1000 g cm^{-2} (normally around 3 m) as the spallogenic reactions decrease (Heisinger *et al.*, 2002a, 2002b; Braucher *et al.*, 2013).

Analysis and age determination

Procedure

With the preceding understanding of how cosmogenic nuclides are produced, this section explains how their analysis can be used in geomorphic studies. Rocks and sediments containing the required target mineral are sampled in the field using the protocol given in Table 2. Samples are then

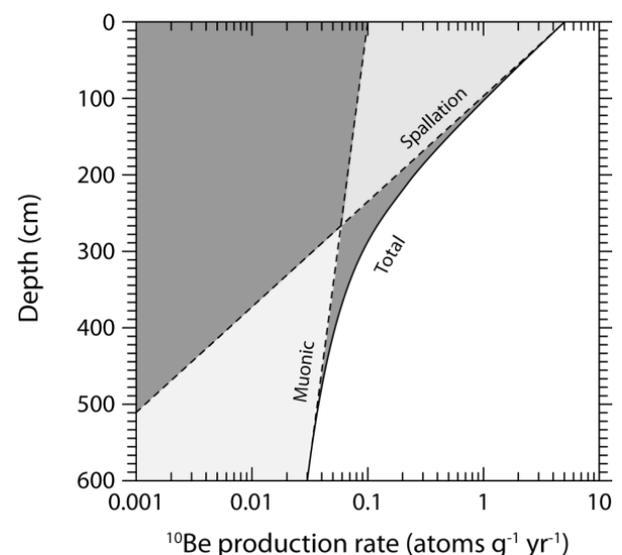


Figure 3: Production of ^{10}Be with depth in quartz arenite (adapted from Gosse and Phillips, 2001). Deeper into the rock, muonic reactions play a greater role in total production of cosmogenic nuclides than nucleonic spallation.

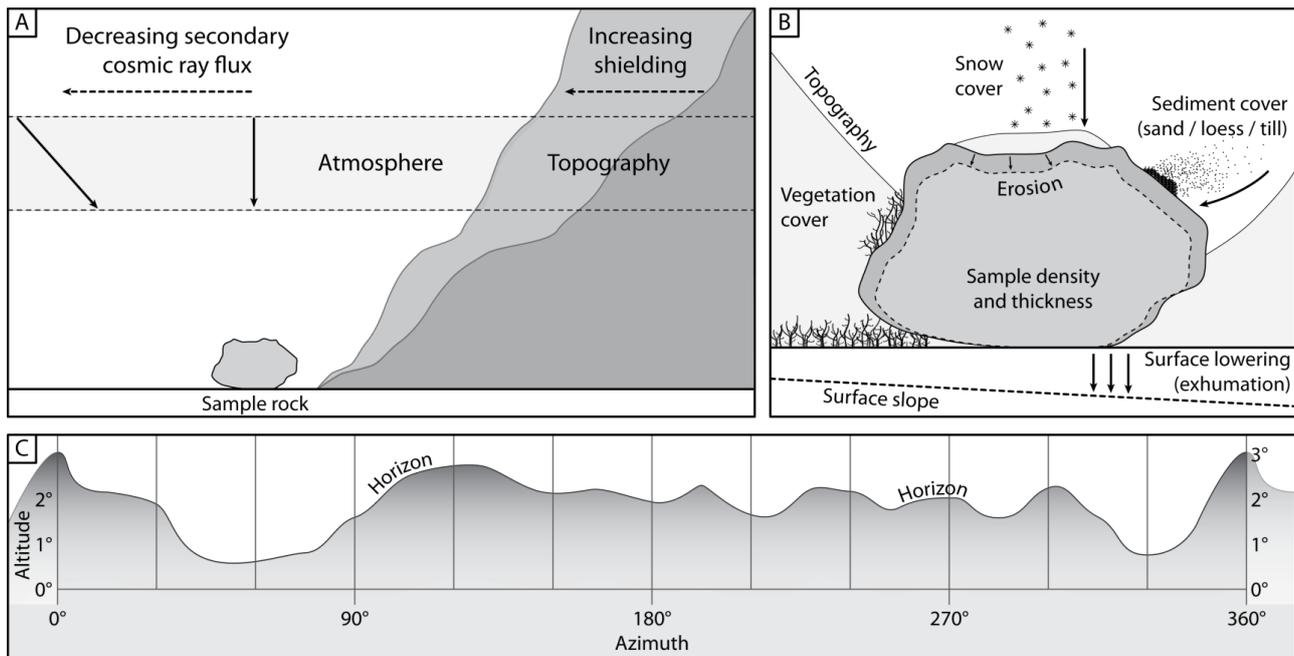


Figure 4: An illustration of factors affecting shielding. (A) Cosmic ray flux decreases if it passes through a greater amount of atmosphere (i.e. at a lower angle of incidence) and steeper local topography will decrease the incoming flux at a site. (B) Various local factors which will alter the shielding of a sample. It is also important to consider that these factors may change over time. (C) Topographic shielding is calculated by determining the degree to which a sampling site is exposed to a full hemisphere of open sky. One approach is to measure the altitude of the horizon at regular intervals (e.g. every 30°) for a full 360° and then calculate the exposure (Balco et al., 2008).

mechanically crushed and physically and chemically separated to isolate first the target mineral and then the target nuclide within that mineral (Wilson et al., 2008). This is a complex and time-consuming process, the cost of which is often a significant limiting factor on the number of analyses achievable. Nuclide concentration is then measured, in the form of a ratio (e.g. $^9\text{Be}/^{10}\text{Be}$) in an accelerator mass spectrometer to measure the amount of nuclide within that sample. This can be converted into a concentration which can be used to calculate an age (Gosse and Phillips, 2001; Dunai, 2010), which has been greatly aided by the development of calculator tools such as those of Vermeesch (2007) and Balco et al. (2008). However, it is still important to understand how both the geomorphological context of the sample (e.g. shielding, erosion rate and inheritance) and the selection of different production rate parameters can affect the age calculation and may limit its usefulness.

Geomorphological context

The choice of cosmogenic nuclide used is determined by the timescale of interest (nuclide half-life) and the rock types available (target minerals; Ivy-Ochs and Kober, 2008).

It is important to demonstrate that a sample of rock or sediment will relate unambiguously to the geomorphological process or event being studied, requiring an initial understanding of the environment and relative history, such as through geomorphological mapping. Changes to that environment over time will influence how nuclides have accumulated within the sample, so it is also necessary to consider the impact of shielding, erosion, inheritance and elevation-change over time (expanded in following sections; and Gosse and Phillips, 2001).

Shielding

The cosmic ray flux can be reduced at a particular location by local obstructions to the full horizon, such as the surrounding topography, surface slope and local rock formations (Gosse et al., 1995; Dunne et al., 1999). This topographic shielding blocks some of the incoming rays, altering the local production rate and, if uncorrected, can produce erroneously young ages. Incident cosmic radiation has a non-linear dependence on the angle from the horizon,

Table 2: A sampling checklist for cosmogenic nuclide analysis (adapted from author's own protocol and: Gosse and Phillips (2001), Hubbard and Glasser (2005), Hein et al. (2009) and Dunai (2010).

1. Equipment	<ul style="list-style-type: none"> • Marker pens • Hand-held GPS • Digital camera • Tape measure • Sample bags (cloth bags are more robust than plastic) • Hammer and chisel / core-drill / rock-cutting saw (boulders or bedrock) • Scales (for weighing samples, otherwise estimate) • Spade, pick-axe, hand-trowel (depth-profiles) • Spirit level and spray paint (depth profiles) <ul style="list-style-type: none"> • Maps/aerial photos of the area • Compass • Abney level or (compass) clinometer
------------------------	--

2. Considerations	Context	<p>The geomorphic context of the samples(s) must be understood. It is important to be clear about the link between the sample and geomorphic process or event being studied.</p> <p>For single samples, avoid any situations where the stability of the boulder or surface is in question (e.g. signs of rolling, shifting or slumping). For depth profiles, be sure that the geomorphic surface is stable and that it is the original surface being studied.</p>
	Sample history	<p>Are the following likely to have affected the sample?</p> <ul style="list-style-type: none"> • Soil cover • Snow cover • Ice/glacier cover • Vegetation • Ash cover • Water/lake cover <p>Could the history be complex? Avoid complexity if possible.</p>
	Ethics	Minimise the effects that sampling will have on the environment.

3a. Single samples overview	<ol style="list-style-type: none"> 1. Rock lithology needs to contain the target mineral required for the cosmogenic nuclide being used 2. Aim for at least three samples per landform (sequence) being dated 3. If sampling boulders, aim for the tallest boulders possible 4. Do not sample edges and corners of boulders and rock outcrops 5. If possible, avoid signs of post-depositional erosion 6. Aim to take a sample roughly 5cm deep, but be aware that for ^{36}Cl, a thicker sample will better record the nuclides produced by thermal neutron capture 7. Sample weight will depend on lithology and age, so size/weight of sample will be an estimate. It is consequently better to aim for a larger sample than is likely to be necessary. <p>Record features according to Part 4 below</p>
3b. Depth profile sampling overview	<ol style="list-style-type: none"> 1. Profile should be at least 1.5 - 2.5 m deep, measured from regional surface level 2. Excavate a clean, relatively flat surface and describe stratigraphy of the section 3. Measure soil depth from surface if necessary 4. Use spirit level to mark sampling depths from surface across width of the profile 5. Sample at least 6 depths every 10-50 cm, capturing closely-spaced samples towards the top of the profile and at least one deep sample (> 3 m if possible) 6. Aim to sample at least 1 kg of amalgamated pebbles or sand samples – use clasts < 3 cm to avoid single clasts dominating the nuclide signature 7. Collect clasts/sediment lithologies containing the target mineral required 8. Measure or estimate thickness of samples 9. If possible, measure average bulk density of each sedimentary unit. 10. If sampling surface clasts, aim for larger (20 - 30 cm) clasts of a suitable lithology that are embedded within the landscape surface and can be linked unambiguously to the depth profile. They will be treated individually as single samples. <p>Record features according to Part 4 below</p>

4. Record	Time	Note the date and time.
	Location	Record latitude, longitude and elevation using GPS.
	Context	Describe the sample site and how it relates to the surrounding geomorphic environment. Is the sample site (un)usual?
	Labelling	<p>Label samples with:</p> <ul style="list-style-type: none"> • Sample name • Top surface • Burial line if partially submerged. <p>If a single sample is taken in fragments, label in such a way as to be able to reconstruct the complete sample in the laboratory. For depth-profile amalgams, label all bags carefully, including depths. Label all samples with top surface, sample name and burial line if they are partially submerged.</p>
	Photos	<p>Take:</p> <ul style="list-style-type: none"> • Photos of the sample site from as many angles as possible, both near- and far-field, including a scale! • Photos of the rock before and after sampling (single samples) or sediment face (depth profile), including a scale! • A 360° panorama from sample site looking outward. <p>Include sample names in the photos for reference. Note all photograph numbers and what they are showing.</p>
	Physical characteristics	<p>Describe and photograph the following:</p> <ul style="list-style-type: none"> • Size • Shape • Lithology • Colour • Grain size • Lichen cover • Glacial polish • Striations • Visible cracks (width and depth) • Emergent veins (measure height) • Weathering pits • Rock-varnish • Ventifacts • Any other characteristics relevant to exposure history or geomorphic event/process being studied <p>For a depth profile, full, detailed stratigraphic log of the sedimentary exposure, noting imbrication and cementation</p>
	Topographic shielding	Record the angular elevation of the horizon from the sample at regular intervals for a full 360° See <i>Figure 4</i> .
	Sample strike/dip	Likely only an issue with single samples – measure the strike and dip of the surface from which the samples has been taken if it is not level
Thickness and fragments	<ul style="list-style-type: none"> • Record the thickness of a single sample, or the thicknesses of fragments to aid reconstruction in the laboratory. • Measure or estimate the thickness of depth profile amalgam samples – this will help inform the depth error of samples. 	

so it requires a high degree of shielding to significantly affect the cosmic ray flux. However, shielding is important for age calculation, and is measured in the field (Figure 4). In addition, samples from protruding objects can lose some high energy neutrons back to the atmosphere, resulting in a lower spallation rate (Masarik and Wieler, 2003). For this reason, the edges of rocks are avoided (Gosse and Phillips, 2001). Post-depositional shielding must also be considered, such as annual or semi-permanent snow, sand, loess, soil or vegetation cover. This is difficult to incorporate into an age calculation, so sampling should avoid signs of post-depositional shielding.

Erosion

Over time, the surface of a rock or sedimentary unit will be eroded. This has two consequences: first, the top layer in which cosmogenic nuclides are produced will be progressively removed, and secondly, the contemporary surface may have been exposed later than the event being dated (Gosse and Phillips, 2001). These two factors will combine to produce an erroneously young age for the sample. This can be accounted for, to a degree, by measuring the protrusion of more resistant mineral bands (such as quartz veins) or the depth of surface pitting on rocks and using this as a proxy for

erosion rate (Gosse and Phillips, 2001). Alternatively, a depth-profile through sediment can be used to model the erosion rate (see 'Single samples vs depth-profiles' section; and Hein *et al.*, 2009). Few independent measures of erosion rate exist, so as a compromise, two age estimates are often given for a single sample: one assuming no erosion has occurred and the other under a constant, *estimated* rate of erosion. The former provides a minimum age for the deposit, irrespective of the amount of erosion. Older exposure ages become increasingly susceptible to the erosion rate, hence this cautionary approach.

Inheritance

Another assumption of cosmogenic nuclide analysis is that the sample has not been previously exposed to cosmic rays prior to the event in question, and it thus contains no prior cosmogenic nuclide signature (Anderson *et al.*, 1996). However, pre-exposure to the atmosphere (e.g. reworking of clasts) and a lack of erosion of the rock surface to sufficient depth may fail to reset the cosmogenic 'clock' (Guido *et al.*, 2007). The presence of an inherited nuclide component will provide an anomalously old age for a sample. In single nuclide analysis, outliers can be identified using statistical analyses such as relative probability distributions and reduced chi-squared tests,

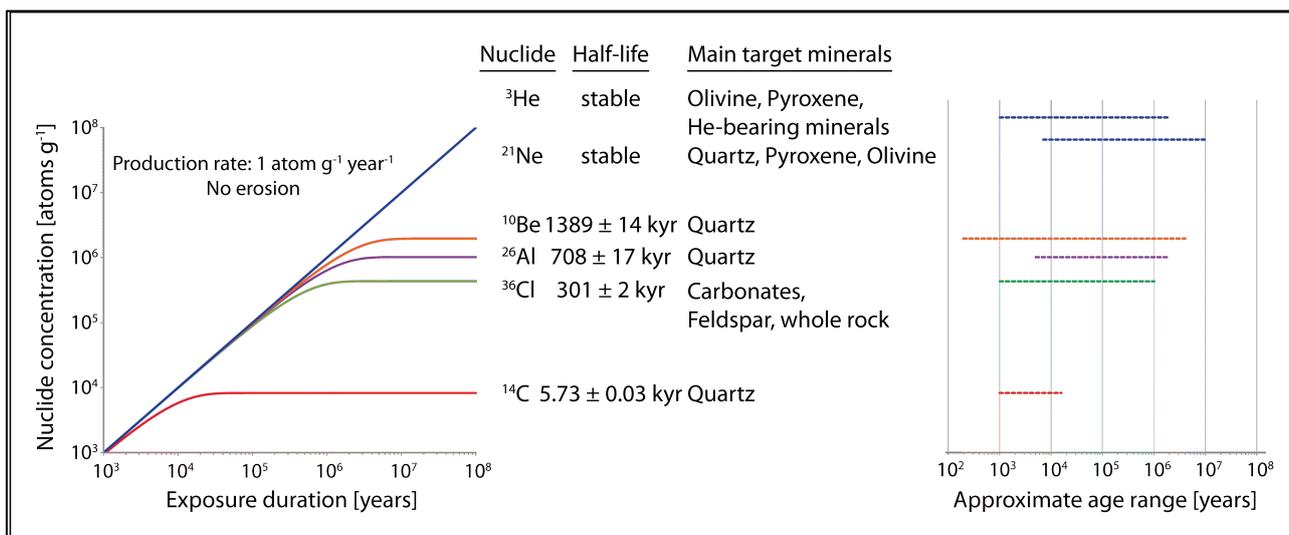


Figure 5: A summary of the properties of the six most commonly-used cosmogenic nuclides. Age ranges are approximate and are heavily controlled by laboratory methodology and local factors such as erosion. The plot shows that ^{14}C has the shortest useable age range because of its relatively rapid decay, whilst longer-lived radionuclides (e.g. ^{10}Be) can be used for much broader age ranges. Data from Ivy-Ochs and Kober (2008) and Dunai (2010). ^{10}Be half-life is from Chmeleff *et al.* (2010) and Korschinek *et al.* (2010).

but require numerous samples to have been taken (Barrows *et al.*, 2002). In depth-profiles, a nuclide concentration from a sufficiently deep sample (e.g. greater than roughly 3 m) can only have been created by inheritance, and this can be taken into account during age calculation (see 'Single samples vs depth profiles' section; Repka *et al.*, 1997; Hancock *et al.*, 1999). Some applications use multiple isotopes both to identify nuclide inheritance but also to quantify erosion rates.

Using multiple isotopes

Because radioactive cosmogenic nuclides decay over time, they have the potential to provide information on both the exposure time (the amount of accumulated isotope) and the degree of burial, shielding or erosion of the surface (the amount of the accumulated isotopes that have decayed; Dehnert and Schlüchter, 2008). Consequently, the technique is particularly powerful when multiple isotopes are used which have significantly different rates of production and different decay half-lives (Figure 5; Lal, 1991; Gosse and Phillips, 2001). For example, a combination of ^{10}Be ($\frac{1}{2}$ life: 1389 ± 14 kyr) and ^{26}Al ($\frac{1}{2}$ life: 708 ± 17 kyr) can show whether a rock has experienced a simple or complex exposure history because a period of shielding will allow almost twice as much ^{26}Al to decay than ^{10}Be , resulting in discordant age estimates (Lal, 1991; Balco *et al.*, 2005; Dehnert and Schlüchter, 2008).

Single samples vs depth-profiles

Single samples are more commonly used than depth-profiles for dating (Figure 6). Single samples can potentially yield a date per analysis and are often simpler to sample in the field. However, it can be difficult to quantify the effects of shielding, erosion and inheritance, creating scatter in data sets from single geomorphological features that increases with landform age (Balco, 2011). Larger datasets can statistically reduce this uncertainty, but require more analyses to be performed (Rinterknecht *et al.*, 2006).

Conversely, depth-profiles can provide a wealth of information in addition to an age. Several samples are analysed from various depths through a sedimentary unit, whereby the surface concentration relates to the age



Figure 6: Two different approaches to sampling for cosmogenic nuclide dating. (A) Single sample, such as a large, intact erratic boulder. (B) Depth profile, such as through glaciofluvial outwash. (Photographs taken by the author, southern South America).

of the unit, and samples at increasing depth will contain progressively decreasing nuclide concentrations according to the normal attenuation of cosmic radiation (Anderson *et al.*, 1996; Repka *et al.*, 1997). Modelling of the nuclide concentration data, such as using a chi-squared best-fit, can produce an age for the unit, but also information on average inheritance, surface erosion/deflation, and changes in deposition, thereby resulting in a more robust age estimate. The use of multiple isotopes in a profile can also yield both exposure and burial ages (Granger and Muzikar, 2001; Häuselmann *et al.*, 2007; Balco and Rovey, 2008; Hein *et al.*, 2009). However, a depth-profile requires a number of analyses (e.g. >5) to yield only one actual age estimate, and may prove problematic if using fractions greater than sand-size (Hidy *et al.*, 2010). It is also necessary to find or create a deep exposure within the sedimentary unit at a suitable geomorphological location, which may not always be possible. Consequently, depth profiles only tend to be used instead of single

samples where the sediment being dated lends itself to this form of analysis (e.g. fluvial/glaciofluvial/alluvial terraces) and/or where post-depositional processes are thought to undermine ages from single samples (Hein *et al.*, 2009, 2011a).

Production rates and scaling factors

In order to establish an age from the accumulation of a particular cosmogenic nuclide, the rate at which that nuclide is produced at the sampling site must be estimated, given that the *in situ* production of cosmogenic nuclides varies with latitude, altitude and the thickness and density of a sample (Lal, 1991; Dunai, 2000; Stone, 2000), and will be affected by shielding. Production rates of different nuclides at different locations are most commonly established using well-dated event features, where the dates are independently established using an alternative technique such as radiocarbon-dated moraines (Putnam *et al.*, 2010b; Kaplan *et al.*, 2011), argon-dated volcanic lavas (Dunai and Wijbrans, 2000; Foeken *et al.*, 2012), varve records (Balco and Schaefer, 2006) or tree-ring chronologies (Kubik *et al.*, 1998); as well as inter-comparison between different nuclides at such sites (Balco and Shuster, 2009; Schimmelpfennig *et al.*, 2011).

These production rates are usually published in a normalised form to Sea-Level and High Latitude (SLHL). Then, either the most appropriate regional production rate can be applied directly or a reference production rate can be used, which consists of an average of these different regional studies (Balco *et al.*, 2008). Whichever approach is adopted, *scaling factors* must be carefully considered in order to appropriately scale the production rate from the regional or global rate (given at SLHL) to the local site from where the sample was taken, taking into account the variable production at different latitudes and altitudes (Lal, 1991). Six main scaling models exist, which deal with variations in production rates in different ways (Lal, 1991; Stone, 2000; Dunai, 2001; Desilets and Zreda, 2003; Lifton *et al.*, 2005; Desilets *et al.*, 2006). There is not a great deal of difference between the models; they mainly manipulate the effects of atmospheric pressure, the geomagnetic field and solar variability in different ways (Balco *et al.*, 2008; Dunai, 2010), and age calculators

often produce a range of ages to show the effects of using each model. However, there must be consistency between the scaling factors used, including when comparing ages between different studies that may have used different production rates and scaling factors (Balco, 2011).

Age range and applications

Stable nuclides accumulate over time according to their production rate and will eventually reach equilibrium with respect to the erosion rate of the rock surface. By contrast, the production of a radioactive nuclide will eventually equal both the erosion rate and the amount of nuclide lost through radioactive decay. The state of equilibration is called *saturation* and limits the maximum timescales over which the nuclides can be used to date (Figure 5). The minimum age limit is determined by the measurement procedure (Ivy-Ochs and Kober, 2008).

A summary of the uses of cosmogenic nuclide analysis in different settings is given in Table 3. The applications are grouped into different geomorphological environments: glacial, hillslope, fluvial, coastal/lacustrine, volcanic, dryland, karst, extra-terrestrial and tectonic. Broadly, these constitute three different types of analysis: events (exposure and burial dating), rates (erosion/denudation, uplift and soil dynamic rates) and direct palaeo-altimetry. A limited number of exemplar publications are given as a starting point for further environment-specific investigation.

Recent advances

Cosmogenic nuclide analysis is an exciting and developing tool in geochronology and there is scope to improve the technique in coming years. Recent work has focussed on improving production rate estimates, critically analysing data in relation to geomorphological processes and extending the age range of the technique.

Uncertainties in reference production rates are being reduced by improving the number, quality and global distribution of regional calibration sites. This has been greatly aided by the CRONUS-Earth (2005-2010) and CRONUS-EU (2004-2008) initiatives and forthcoming projects such as SPICE (Fenton

Table 3: Summary of the range of applications for which cosmogenic nuclide dating have been used. This is an update to the table produced by Cockburn and Summerfield (2004) but is still a limited selection given the huge increase in literature over the last decade. It should therefore be used as the starting point for further reading.

Environment	Study	Type of dating	Example studies
Glacial	Moraines	Exposure Event	Phillips <i>et al.</i> (1990); Brown <i>et al.</i> (1991); Brook <i>et al.</i> (1993); Brook <i>et al.</i> (1995); Gosse <i>et al.</i> (1995); Phillips <i>et al.</i> (1996); Jackson <i>et al.</i> (1997); Barrows <i>et al.</i> (2001); Owen <i>et al.</i> (2001); Barrows <i>et al.</i> (2002); Briner <i>et al.</i> (2002); Owen <i>et al.</i> (2002); Schäfer <i>et al.</i> (2002); Kaplan <i>et al.</i> (2004); Schaefer <i>et al.</i> (2006); Barrows <i>et al.</i> (2007); Putnam <i>et al.</i> (2010a); Ballantyne (2012); Rinterknecht <i>et al.</i> (2012b); Rinterknecht <i>et al.</i> (2012a)
	Bedrock features	Exposure Event	Nishiizumi <i>et al.</i> (1989); Nishiizumi <i>et al.</i> (1991a); Stone <i>et al.</i> (1998); Bierman <i>et al.</i> (1999); Fabel <i>et al.</i> (2002); Stroeven <i>et al.</i> (2002); Marquette <i>et al.</i> (2004); Harbor <i>et al.</i> (2006); Phillips <i>et al.</i> (2006); McCormack <i>et al.</i> (2011); Briner <i>et al.</i> (2012b); Hippe <i>et al.</i> (2013)
	Ice-sheet thinning	Exposure Event	Stone <i>et al.</i> (2003); Bentley <i>et al.</i> (2006); Bentley <i>et al.</i> (2010); Todd <i>et al.</i> (2010); Hein <i>et al.</i> (2011b); Mackintosh <i>et al.</i> (2011)
	Glaciofluvial features	Exposure Event	Hein <i>et al.</i> (2009); Hein <i>et al.</i> (2011a)
	Buried ice	Burial Event	Schäfer <i>et al.</i> (2000); Marchant <i>et al.</i> (2002)
	Ablation	Exposure Rate	Lal <i>et al.</i> (1987); Lal & Jull (1990); Lal <i>et al.</i> (1990); Lal & Jull (1992); Jull <i>et al.</i> (1994)
Hillslope	Mass movement	Exposure Event	Ballantyne <i>et al.</i> (1998); Kubik <i>et al.</i> (1998); Barnard <i>et al.</i> (2001); Sanchez <i>et al.</i> (2010); Akçar <i>et al.</i> (2012a)
	Slope translocation	Exposure/Burial Rate	Small <i>et al.</i> (1999); Heimsath <i>et al.</i> (2002); Heimsath <i>et al.</i> (2005); Nichols <i>et al.</i> (2007)
	Surface burial	Burial Event	Granger & Smith (2000); Granger & Muzikar (2001); Kong <i>et al.</i> (2009); Hu <i>et al.</i> (2011); Matmon <i>et al.</i> (2012)
	Surface denudation (local)	Exposure Rate	Nishiizumi <i>et al.</i> (1986); Brown <i>et al.</i> (1995); Heimsath <i>et al.</i> (1997); Small <i>et al.</i> (1997); Heimsath <i>et al.</i> (1999); Cockburn <i>et al.</i> (2000); Heimsath <i>et al.</i> (2000;2001b); Heimsath <i>et al.</i> (2001a); Norton <i>et al.</i> (2011); Hippe <i>et al.</i> (2012)
	Surface denudation (regional)	Exposure Rate	Brown <i>et al.</i> (1995); Riebe <i>et al.</i> (2000); Bierman & Caffee (2001); Kirchner <i>et al.</i> (2001); Riebe <i>et al.</i> (2001a;b); Schaller <i>et al.</i> (2001); Nichols <i>et al.</i> (2005); Norton <i>et al.</i> (2008); Aguilar <i>et al.</i> (2013)

	Soil translocation	Exposure/Burial Rate	Brown <i>et al.</i> (1994); Heimsath <i>et al.</i> (1997); Heimsath <i>et al.</i> (1999); Heimsath <i>et al.</i> (2000;2001b); Heimsath <i>et al.</i> (2001a); Heimsath <i>et al.</i> (2005); Burke <i>et al.</i> (2007); Schaller <i>et al.</i> (2009); Riebe & Granger (2013)
Fluvial	Alluvial river terraces	Exposure Event	Molnar <i>et al.</i> (1994); Repka <i>et al.</i> (1997); Hancock <i>et al.</i> (1999); Hetzel <i>et al.</i> (2002); Schildgen <i>et al.</i> (2002); Darling <i>et al.</i> (2012)
	Alluvial fans / debris flows	Exposure/Burial Event	Siame <i>et al.</i> (1997); Brown <i>et al.</i> (1998); Nishiizumi <i>et al.</i> (2005); Dühnforth <i>et al.</i> (2007); Frankel <i>et al.</i> (2007); Blisniuk <i>et al.</i> (2012)
	Floods	Exposure/Burial Event	Cerling <i>et al.</i> (1994); Margerison <i>et al.</i> (2005); Reuther <i>et al.</i> (2006); Amidon & Farley (2011)
	Strath terraces	Exposure Event	Burbank <i>et al.</i> (1996); Leland <i>et al.</i> (1998); Pratt <i>et al.</i> (2002); Ruzsckiczay-Rüdiger <i>et al.</i> (2005); Schaller <i>et al.</i> (2005)
Coastal/ lacustrine	Shoreline deposits	Exposure Event	Trull <i>et al.</i> (1995); Perg <i>et al.</i> (2001); Matmon <i>et al.</i> (2003); Owen <i>et al.</i> (2007); Kurth <i>et al.</i> (2011)
	Wave-cut platforms	Exposure Event	Stone <i>et al.</i> (1996); Alvarez-Marrón <i>et al.</i> (2008)
	Lacustrine sediments	Burial Event	Kong <i>et al.</i> (2009); Davis <i>et al.</i> (2011)
Volcanic	Volcanic landforms	Exposure/Burial Event	Craig & Poreda (1986); Kurz (1986); Kurz <i>et al.</i> (1990); Poreda & Cerling (1992); Laughlin <i>et al.</i> (1994); Licciardi <i>et al.</i> (1999); Goethals <i>et al.</i> (2009); Schimmelpfennig <i>et al.</i> (2009); Schimmelpfennig <i>et al.</i> (2011); Medynski <i>et al.</i> (2013)
Dryland/ aeolian	Aeolian erosion and denudation	Exposure/Burial Rate	Bierman & Caffee (2001); Vermeesch <i>et al.</i> (2010); Fujioka & Chappell (2011); Ruzsckiczay-Rüdiger <i>et al.</i> (2011)
	Desert pavements	Exposure Event	Wells <i>et al.</i> (1995); Marchetti & Cerling (2005); Matmon <i>et al.</i> (2009)
Karst	Sediment deposition in caves	Burial Event	Granger <i>et al.</i> (1997); Granger <i>et al.</i> (2001); Stock <i>et al.</i> (2004); Stock <i>et al.</i> (2006); Anthony & Granger (2007); Matmon <i>et al.</i> (2012)
Extra-terrestrial	Meteorite impacts	Exposure Event	Nishiizumi <i>et al.</i> (1991b); Phillips <i>et al.</i> (1991)
Tectonic	Fault scarp development	Exposure Event	Zreda & Noller (1998); Mitchell <i>et al.</i> (2001); Daëron <i>et al.</i> (2004); Palumbo <i>et al.</i> (2004); Schlagenhaut <i>et al.</i> (2010); Akçar <i>et al.</i> (2012b)
	Change in elevation	Palaeo-altimetry	Brook <i>et al.</i> (1995); Van der Wateren <i>et al.</i> (1999); Gosse & Stone (2001); Riihimaki & Libarkin (2007)

et al., in press). For example, recent ^{10}Be calibration studies from northeastern North America (Balco *et al.*, 2009), New Zealand (Putnam *et al.*, 2010b) Patagonia (Kaplan *et al.*, 2011), Norway (Fenton *et al.*, 2011; Goehring *et al.*, 2012), Greenland (Briner *et al.*, 2012a) and western Greenland/Baffin Island (Young *et al.*, 2013) gave more accurate results than previous work (Young *et al.*, 2013). They also yielded production rates that are lower than those produced in previous studies, which is particularly important given that some studies are now using ^{10}Be to attempt to resolve sub-millennial glacial events and their inter-hemispheric relationships.

There has been a recent drive in the critical assessment of the way in which cosmogenic nuclide data is handled and statistically manipulated. Given the particular rise of cosmogenic nuclide exposure dating in glacial geomorphological studies, Balco (2011) cautioned against extending the technique beyond its limitations, particularly with regards to resolving shorter timescale events. More specifically, Heyman *et al.* (2011) compiled a vast number of published exposure ages from around the world to demonstrate that shielding, erosion and exhumation of boulders often cause incorrect ages – to a greater degree than inheritance, so that ages are more likely to be underestimates. This has been developed further using statistical techniques to manipulate cosmogenic nuclide data to examine geomorphological processes (Putkonen and Swanson, 2003; Putkonen and O'Neal, 2006; Applegate *et al.*, 2012; Mackey and Lamb, 2013).

Studies are continuing to extend the age range of cosmogenic nuclide analysis. Younger age constraints remain limited to hundreds of years at best by methodological constraints (Davis *et al.*, 1999; Akçar *et al.*, 2012a; Putnam *et al.*, 2012; Kaplan *et al.*, 2013), but the upper age range is being extended into the multiple millions of years using different cosmogenic isotopes (Schäfer *et al.*, 1999; Balco *et al.*, 2005; Dunai *et al.*, 2005; Stock *et al.*, 2005; Häuselmann *et al.*, 2007; Alvarez-Marrón *et al.*, 2008; Balco and Rovey, 2010; Hein *et al.*, 2011a).

In conclusion, cosmogenic nuclide analysis is likely to remain one of the primary tools used in geomorphological studies. However, the generation of meaningful data requires clear understanding of the geomorphological history of samples and the limitations of the technique.

Acknowledgments

This chapter was written whilst the author was in receipt of a NERC Ph.D. studentship at Durham University. An earlier draft benefitted from the constructive comments of Mike Bentley and Chris Stokes and the final paper was improved by comments from two anonymous reviewers.

References

- Aguilar G, Carretier S, Regard V, Vassallo R, Riquelme R, Martinod J. 2013. Grain size-dependent ^{10}Be concentrations in alluvial stream sediment of the Huasco Valley, a semi-arid Andes region. *Quaternary Geochronology* <http://dx.doi.org/10.1016/j.quageo.2013.01.011>
- Akçar N, Deline P, Ivy-Ochs S, Alfimov V, Hajdas I, Kubik P, Christl M, Schlüchter C. 2012a. The AD 1717 rock avalanche deposits in the upper Ferret Valley (Italy): a dating approach with cosmogenic ^{10}Be . *Journal of Quaternary Science* **27**: 383-392 <http://dx.doi.org/10.1002/jqs.1558>
- Akçar N, Tikhomirov D, Özkaymak Ç, Ivy-Ochs S, Alfimov V, Sözbilir H, Uzel B, Schlüchter C. 2012b. ^{36}Cl exposure dating of paleoearthquakes in the Eastern Mediterranean: First results from the western Anatolian Extensional Province, Manisa fault zone, Turkey. *Geological Society of America Bulletin* **124**: 1724-1735 <http://dx.doi.org/10.1130/b30614.1>
- Alvarez-Marrón J, Hetzel R, Niedermann S, Menéndez R, Marquínez J. 2008. Origin, structure and exposure history of a wave-cut platform more than 1 Ma in age at the coast of northern Spain: A multiple cosmogenic nuclide approach. *Geomorphology* **93**: 316-334 <http://dx.doi.org/10.1016/j.geomorph.2007.03.005>
- Amidon W, Farley K. 2011. Cosmogenic ^3He production rates in apatite, zircon and pyroxene inferred from Bonneville flood erosional surfaces. *Quaternary*

Geochronology **6**: 10-21
<http://dx.doi.org/10.1016/j.quageo.2010.03.005>

Anderson R, Repka J, Dick G. 1996. Explicit treatment of inheritance in dating depositional surfaces using in situ ^{10}Be and ^{26}Al . *Geology* **24**: 47-51 [http://dx.doi.org/10.1130/0091-7613\(1996\)024<0047:etoiid>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1996)024<0047:etoiid>2.3.co;2)

Anthony D, Granger D. 2007. A new chronology for the age of Appalachian erosional surfaces determined by cosmogenic nuclides in cave sediments. *Earth Surface Processes and Landforms* **32**: 874-887 <http://dx.doi.org/10.1002/esp.1446>

Applegate P, Urban N, Keller K, Lowell T, Laabs B, Kelly M, Alley R. 2012. Improved moraine age interpretations through explicit matching of geomorphic process models to cosmogenic nuclide measurements from single landforms. *Quaternary Research* **77**: 293-304 <http://dx.doi.org/10.1016/j.yqres.2011.12.002>

Balco G. 2011. Contributions and unrealized potential contributions of cosmogenic-nuclide exposure dating to glacier chronology, 1990–2010. *Quat. Sci. Rev.* **30**: 3-27 <http://dx.doi.org/10.1016/j.quascirev.2010.11.003>

Balco G, Rovey C. 2010. Absolute chronology for major Pleistocene advances of the Laurentide Ice Sheet. *Geology* **38**: 795-798 <http://dx.doi.org/10.1130/g30946.1>

Balco G, Shuster D. 2009. Production rate of cosmogenic ^{21}Ne in quartz estimated from ^{10}Be , ^{26}Al , and ^{21}Ne concentrations in slowly eroding Antarctic bedrock surfaces. *Earth Planet. Sci. Lett.* **281**: 48-58 <http://dx.doi.org/10.1016/j.epsl.2009.02.006>

Balco G, Rovey C. 2008. An isochron method for cosmogenic-nuclide dating of buried soils and sediments. *American Journal of Science* **308**: 1083-1114 <http://dx.doi.org/10.2475/10.2008.02>

Balco G, Schaefer J. 2006. Cosmogenic-nuclide and varve chronologies for the deglaciation of southern New England. *Quaternary Geochronology* **1**: 15-28 <http://dx.doi.org/10.1016/j.quageo.2006.06.014>

Balco G, Briner J, Finkel R, Rayburn J, Ridge J, Schaefer J. 2009. Regional beryllium-10 production rate calibration for late-glacial northeastern North America. *Quaternary Geochronology* **4**: 93-107 <http://dx.doi.org/10.1016/j.quageo.2008.09.001>

Balco G, Stone J, Lifton N, Dunai T. 2008. A complete and easily accessible means of calculating surface exposure ages or erosion

rates from ^{10}Be and ^{26}Al measurements. *Quaternary Geochronology* **3**: 174-195 <http://dx.doi.org/10.1016/j.quageo.2007.12.001>

Balco G, Stone J, Mason J. 2005. Numerical ages for Plio-Pleistocene glacial sediment sequences by $^{26}\text{Al}/^{10}\text{Be}$ dating of quartz in buried paleosols. *Earth Planet. Sci. Lett.* **232**: 179-191 <http://dx.doi.org/10.1016/j.epsl.2004.12.013>

Ballantyne C. 2012. Chronology of glaciation and deglaciation during the Loch Lomond (Younger Dryas) Stade in the Scottish Highlands: implications of recalibrated ^{10}Be exposure ages. *Boreas* **41**: 513-526 <http://dx.doi.org/10.1111/j.1502-3885.2012.00253.x>

Ballantyne C, Stone J, Fifield L. 1998. Cosmogenic Cl-36 dating of postglacial landsliding at The Storr, Isle of Skye, Scotland. *The Holocene* **8**: 347-351 <http://dx.doi.org/10.1191/095968398666797200>

Barnard P, Owen L, Sharma M, Finkel R. 2001. Natural and human-induced landsliding in the Garhwal Himalaya of northern India. *Geomorphology* **40**: 21-35 [http://dx.doi.org/10.1016/S0169-555X\(01\)00035-6](http://dx.doi.org/10.1016/S0169-555X(01)00035-6)

Barrows T, Lehman S, Fifield L, De Deckker P. 2007. Absence of Cooling in New Zealand and the Adjacent Ocean During the Younger Dryas Chronozone. *Science* **318**: 86-89 <http://dx.doi.org/10.1126/science.1145873>

Barrows T, Stone J, Fifield L, Cresswell RG. 2002. The timing of the Last Glacial Maximum in Australia. *Quat. Sci. Rev.* **21**: 159-173 [http://dx.doi.org/10.1016/S0277-3791\(01\)00109-3](http://dx.doi.org/10.1016/S0277-3791(01)00109-3)

Barrows T, Stone J, Fifield L, Cresswell RG. 2001. Late Pleistocene Glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quaternary Research* **55**: 179-189 <http://dx.doi.org/10.1006/qres.2001.2216>

Bentley M, Fogwill C, Le Brocq A, Hubbard A, Sugden D, Dunai T, Freeman S. 2010. Deglacial history of the West Antarctic Ice Sheet in the Weddell Sea embayment: Constraints on past ice volume change. *Geology* **38**: 411-414 <http://dx.doi.org/10.1130/g30754.1>

Bentley M, Fogwill C, Kubik P, Sugden D. 2006. Geomorphological evidence and cosmogenic $^{10}\text{Be}/^{26}\text{Al}$ exposure ages for the Last Glacial Maximum and deglaciation of the Antarctic Peninsula Ice Sheet. *Geological Society of America Bulletin* **118**: 1149-1159 <http://dx.doi.org/10.1130/b25735.1>

- Bierman P, Caffee M. 2001. Slow Rates of Rock Surface Erosion and Sediment Production across the Namib Desert and Escarpment, Southern Africa. *American Journal of Science* **301**: 326-358 <http://dx.doi.org/10.2475/ajs.301.4-5.326>
- Bierman P, Marsella K, Patterson C, Davis P, Caffee M. 1999. Mid-Pleistocene cosmogenic minimum-age limits for pre-Wisconsinan glacial surfaces in southwestern Minnesota and southern Baffin Island: a multiple nuclide approach. *Geomorphology* **27**: 25-39 [http://dx.doi.org/10.1016/S0169-555X\(98\)00088-9](http://dx.doi.org/10.1016/S0169-555X(98)00088-9)
- Blisniuk K, Oskin M, Fletcher K, Rockwell T, Sharp W. 2012. Assessing the reliability of U-series and ¹⁰Be dating techniques on alluvial fans in the Anza Borrego Desert, California. *Quaternary Geochronology* **13**: 26-41 <http://dx.doi.org/10.1016/j.quageo.2012.08.004>
- Braucher R, Bourlès D, Merchel S, Vidal Romani J, Fernandez-Mosquera D, Marti K, Léanni L, Chauvet F, Arnold M, Aumaitre G, Keddadouche K. 2013. Determination of muon attenuation lengths in depth profiles from in situ produced cosmogenic nuclides. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **294**: 484-490 <http://dx.doi.org/10.1016/j.nimb.2012.05.023>
- Briner J, Young N, Goehring B, Schaefer J. 2012a. Constraining Holocene ¹⁰Be production rates in Greenland. *Journal of Quaternary Science* **27**: 2-6 <http://dx.doi.org/10.1002/jqs.1562>
- Briner J, Lifton N, Miller G, Refsnider K, Anderson R, Finkel R. 2012b. Using in situ cosmogenic ¹⁰Be, ¹⁴C, and ²⁶Al to decipher the history of polythermal ice sheets on Baffin Island, Arctic Canada. *Quaternary Geochronology* <http://dx.doi.org/10.1016/j.quageo.2012.11.005>
- Briner J, Kaufman D, Werner A, Caffee M, Levy L, Manley W, Kaplan M, Finkel R. 2002. Glacier readvance during the late glacial (Younger Dryas?) in the Ahklun Mountains, southwestern Alaska. *Geology* **30**: 679-682 [http://dx.doi.org/10.1130/0091-7613\(2002\)030<0679:grdtlg>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2002)030<0679:grdtlg>2.0.co;2)
- Brook E, Brown E, Kurz M, Ackert R, Raisbeck G, Yiou F. 1995. Constraints on age, erosion, and uplift of Neogene glacial deposits in the Transantarctic Mountains determined from in situ cosmogenic ¹⁰Be and ²⁶Al. *Geology* **23**: 1063-1066 [http://dx.doi.org/10.1130/0091-7613\(1995\)023<1063:coaeau>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1995)023<1063:coaeau>2.3.co;2)
- Brook E, Kurz M, Ackert R, Denton G, Brown E, Raisbeck G, Yiou F. 1993. Chronology of Taylor Glacier Advances in Arena Valley, Antarctica, Using in Situ Cosmogenic ³He and ¹⁰Be. *Quaternary Research* **39**: 11-23 <http://dx.doi.org/10.1006/qres.1993.1002>
- Brown E, Bourlès D, Raisbeck G, Yiou F, Clark Burchfiel B, Molnar P, Qidong D, Jun L. 1998. Estimation of slip rates in the southern Tien Shan using cosmic ray exposure dates of abandoned alluvial fans. *Geological Society of America Bulletin* **110**: 377-386 [http://dx.doi.org/10.1130/0016-7606\(1998\)110<0377:eosrit>2.3.co;2](http://dx.doi.org/10.1130/0016-7606(1998)110<0377:eosrit>2.3.co;2)
- Brown E, Stallard R, Larsen M, Raisbeck G, Yiou F. 1995. Denudation rates determined from the accumulation of in situ-produced ¹⁰Be in the luquillo experimental forest, Puerto Rico. *Earth Planet. Sci. Lett.* **129**: 193-202 [http://dx.doi.org/10.1016/0012-821X\(94\)00249-X](http://dx.doi.org/10.1016/0012-821X(94)00249-X)
- Brown E, Bourlès D, Colin F, Sanfo Z, Raisbeck G, Yiou F. 1994. The development of iron crust lateritic systems in Burkina Faso, West Africa examined with in-situ-produced cosmogenic nuclides. *Earth Planet. Sci. Lett.* **124**: 19-33 [http://dx.doi.org/10.1016/0012-821X\(94\)00087-5](http://dx.doi.org/10.1016/0012-821X(94)00087-5)
- Brown E, Edmond J, Raisbeck G, Yiou F, Kurz M, Brook E. 1991. Examination of surface exposure ages of Antarctic moraines using in situ produced ¹⁰Be and ²⁶Al. *Geochimica Et Cosmochimica Acta* **55**: 2269-2283 [http://dx.doi.org/10.1016/0016-7037\(91\)90103-C](http://dx.doi.org/10.1016/0016-7037(91)90103-C)
- Burbank D, Leland J, Fielding E, Anderson R, Brozovic N, Reid M, Duncan C. 1996. Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas. *Nature* **379**: 505-510 <http://dx.doi.org/10.1038/379505a0>
- Burke B, Heimsath A, White A. 2007. Coupling chemical weathering with soil production across soil-mantled landscapes. *Earth Surface Processes and Landforms* **32**: 853-873 <http://dx.doi.org/10.1002/esp.1443>
- Cerling T, Poreda R, Rathburn S. 1994. Cosmogenic ³He and ²¹Ne age of the Big Lost River flood, Snake River Plain, Idaho. *Geology* **22**: 227-230 [http://dx.doi.org/10.1130/0091-7613\(1994\)022<0227:chanao>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1994)022<0227:chanao>2.3.co;2)
- Chmeleff J, von Blanckenburg F, Kossert K, Jakob D. 2010. Determination of the ¹⁰Be half-life by multicollector ICP-MS and liquid scintillation counting. *Nuclear Instruments and Methods in Physics Research Section B*

- Beam Interactions with Materials and Atoms* **268**: 192-199
<http://dx.doi.org/10.1016/j.nimb.2009.09.012>
- Cockburn H, Summerfield M. 2004. Geomorphological applications of cosmogenic isotope analysis. *Progress in Physical Geography* **28**: 1-42
<http://dx.doi.org/10.1191/0309133304pp395oa>
- Cockburn H, Brown R, Summerfield M, Seidl M. 2000. Quantifying passive margin denudation and landscape development using a combined fission-track thermochronology and cosmogenic isotope analysis approach. *Earth Planet. Sci. Lett.* **179**: 429-435
[http://dx.doi.org/10.1016/S0012-821X\(00\)00144-8](http://dx.doi.org/10.1016/S0012-821X(00)00144-8)
- Craig H, Poreda R. 1986. Cosmogenic ³He in terrestrial rocks: The summit lavas of Maui. *Proceedings of the National Academy of Sciences* **83**: 1970-1974
- Daëron M, Benedetti L, Tapponnier P, Sursock A, Finkel R. 2004. Constraints on the post 25-ka slip rate of the Yammoûneh fault (Lebanon) using in situ cosmogenic ³⁶Cl dating of offset limestone-clast fans. *Earth Planet. Sci. Lett.* **227**: 105-119
<http://dx.doi.org/10.1016/j.epsl.2004.07.014>
- Darling A, Karlstrom K, Granger D, Aslan A, Kirby E, Ouimet W, Lazear G, Coblenz D, Cole R. 2012. New incision rates along the Colorado River system based on cosmogenic burial dating of terraces: Implications for regional controls on Quaternary incision. *Geosphere* **8**: 1020-1041
<http://dx.doi.org/10.1130/ges00724.1>
- Davis M, Matmon A, Fink D, Ron H, Niedermann S. 2011. Dating Pliocene lacustrine sediments in the central Jordan Valley, Israel — Implications for cosmogenic burial dating. *Earth Planet. Sci. Lett.* **305**: 317-327
<http://dx.doi.org/10.1016/j.epsl.2011.03.003>
- Davis P, Bierman P, Marsella K, Caffee M, Southon J. 1999. Cosmogenic analysis of glacial terrains in the eastern Canadian Arctic: a test for inherited nuclides and the effectiveness of glacial erosion. *Annals of Glaciology* **28**: 181-188
<http://dx.doi.org/10.3189/172756499781821805>
- Dehnert A, Schlüchter C. 2008. Sediment burial dating using terrestrial cosmogenic nuclides. *Eiszeitalter und Gegenwart - Quaternary Science Journal* **57**: 210-225
<http://dx.doi.org/10.3285/eg.57.1-2.8>
- Desilets D, Zreda M. 2003. Spatial and temporal distribution of secondary cosmic-ray nucleon intensities and applications to in situ cosmogenic dating. *Earth Planet. Sci. Lett.* **206**: 21-42
[http://dx.doi.org/10.1016/S0012-821X\(02\)01088-9](http://dx.doi.org/10.1016/S0012-821X(02)01088-9)
- Desilets D, Zreda M, Prabu T. 2006. Extended scaling factors for in situ cosmogenic nuclides: New measurements at low latitude. *Earth Planet. Sci. Lett.* **246**: 265-276
<http://dx.doi.org/10.1016/j.epsl.2006.03.051>
- Dühnforth M, Densmore A, Ivy-Ochs S, Allen P, Kubik P. 2007. Timing and patterns of debris flow deposition on Shepherd and Symmes creek fans, Owens Valley, California, deduced from cosmogenic ¹⁰Be. *Journal of Geophysical Research: Earth Surface* **112**: F03S15
<http://dx.doi.org/10.1029/2006jf000562>
- Dunai T. 2010. *Cosmogenic Nuclides: Principles, concepts and applications in the earth surface sciences*. Cambridge University Press, Cambridge.
- Dunai T. 2001. Influence of secular variation of the geomagnetic field on production rates of in situ produced cosmogenic nuclides. *Earth Planet. Sci. Lett.* **193**: 197-212
[http://dx.doi.org/10.1016/S0012-821X\(01\)00503-9](http://dx.doi.org/10.1016/S0012-821X(01)00503-9)
- Dunai T. 2000. Scaling factors for production rates of in situ produced cosmogenic nuclides: a critical reevaluation. *Earth Planet. Sci. Lett.* **176**: 157-169
[http://dx.doi.org/10.1016/S0012-821X\(99\)00310-6](http://dx.doi.org/10.1016/S0012-821X(99)00310-6)
- Dunai T, Wijbrans J. 2000. Long-term cosmogenic ³He production rates (152 ka–1.35 Ma) from ⁴⁰Ar/³⁹Ar dated basalt flows at 29°N latitude. *Earth Planet. Sci. Lett.* **176**: 147-156
[http://dx.doi.org/10.1016/S0012-821X\(99\)00308-8](http://dx.doi.org/10.1016/S0012-821X(99)00308-8)
- Dunai T, López G, Juez-Larré J. 2005. Oligocene–Miocene age of aridity in the Atacama Desert revealed by exposure dating of erosion-sensitive landforms. *Geology* **33**: 321-324
<http://dx.doi.org/10.1130/G21184.1>
- Dunne J, Elmore D, Muzikar P. 1999. Scaling factors for the rates of production of cosmogenic nuclides for geometric shielding and attenuation at depth on sloped surfaces. *Geomorphology* **27**: 3-11
[http://dx.doi.org/10.1016/S0169-555X\(98\)00086-5](http://dx.doi.org/10.1016/S0169-555X(98)00086-5)
- Fabel D, Stroeven A, Harbor J, Kleman J, Elmore D, Fink D. 2002. Landscape preservation under Fennoscandian ice sheets determined from in situ produced ¹⁰Be and

- 26Al. *Earth Planet. Sci. Lett.* **201**: 397-406
[http://dx.doi.org/10.1016/S0012-821X\(02\)00714-8](http://dx.doi.org/10.1016/S0012-821X(02)00714-8)
- Fenton C, Mark D, Barfod D, Niedermann S, Goethals M, Stuart F. In Press. 40Ar/39Ar dating of the SP and Bar Ten lava flows AZ, USA: Laying the foundation for the SPICE cosmogenic nuclide production-rate calibration project. *Quaternary Geochronology*
<http://dx.doi.org/10.1016/j.quageo.2013.01.007>
- Fenton C, Hermanns R, Blikra L, Kubik P, Bryant C, Niedermann S, Meixner A, Goethals M. 2011. Regional 10Be production rate calibration for the past 12 ka deduced from the radiocarbon-dated Grøtlandsura and Russenes rock avalanches at 69° N, Norway. *Quaternary Geochronology* **6**: 437-452
<http://dx.doi.org/10.1016/j.quageo.2011.04.005>
- Foeken J, Stuart F, Mark D. 2012. Long-term low latitude cosmogenic 3He production rate determined from a 126 ka basalt from Fogo, Cape Verdes. *Earth Planet. Sci. Lett.* **359–360**: 14-25
<http://dx.doi.org/10.1016/j.epsl.2012.10.005>
- Frankel K, Brantley K, Dolan J, Finkel R, Klinger R, Knott J, Machette M, Owen L, Phillips F, Slate J, Wernicke B. 2007. Cosmogenic 10Be and 36Cl geochronology of offset alluvial fans along the northern Death Valley fault zone: Implications for transient strain in the eastern California shear zone. *Journal of Geophysical Research: Solid Earth* **112**: B06407
<http://dx.doi.org/10.1029/2006jb004350>
- Fujioka T, Chappell J. 2011. Desert landscape processes on a timescale of millions of years, probed by cosmogenic nuclides. *Aeolian Research* **3**: 157-164
<http://dx.doi.org/10.1016/j.aeolia.2011.03.003>
- Goehring B, Lohne Ø, Mangerud J, Svendsen J, Gyllencreutz R, Schaefer J, Finkel R. 2012. Late glacial and holocene 10Be production rates for western Norway. *Journal of Quaternary Science* **27**: 89-96
<http://dx.doi.org/10.1002/jqs.1517>
- Goethals M, Hetzel R, Niedermann S, Wittmann H, Fenton C, Kubik P, Christl M, von Blanckenburg F. 2009. An improved experimental determination of cosmogenic 10Be/21Ne and 26Al/21Ne production ratios in quartz. *Earth Planet. Sci. Lett.* **284**: 187-198
<http://dx.doi.org/10.1016/j.epsl.2009.04.027>
- Gosse J, Phillips F. 2001. Terrestrial in situ cosmogenic nuclides: theory and application. *Quat. Sci. Rev.* **20**: 1475-1560
[http://dx.doi.org/10.1016/s0277-3791\(00\)00171-2](http://dx.doi.org/10.1016/s0277-3791(00)00171-2)
- Gosse J, Stone J. 2001. Terrestrial cosmogenic nuclide methods passing milestones toward paleo-altimetry. *Eos, Transactions American Geophysical Union* **82**: 82-89
<http://dx.doi.org/10.1029/01eo00045>
- Gosse J, Evenson E, Klein J, Lawn B, Middleton R. 1995. Precise cosmogenic 10Be measurements in western North America: Support for a global Younger Dryas cooling event. *Geology* **23**: 877-880
[http://dx.doi.org/10.1130/00917613\(1995\)023<0877:pcbmiw>2.3.co;2](http://dx.doi.org/10.1130/00917613(1995)023<0877:pcbmiw>2.3.co;2)
- Granger D. 2007. Cosmogenic Nuclide Dating: Landscape Evolution, In: Editor-in-Chief: Scott, A.E. (Ed.), *Encyclopedia of Quaternary Science*. Elsevier, Oxford, pp. 445-452.
- Granger D. 2006. A review of burial dating methods using ²⁶Al and ¹⁰Be. In *Situ-Produced Cosmogenic Nuclides and Quantification of Geological Processes*, *Geological Society of America Special Paper* **415**: 1-16
- Granger D, Muzikar P. 2001. Dating sediment burial with in situ-produced cosmogenic nuclides: theory, techniques, and limitations. *Earth Planet. Sci. Lett.* **188**: 269-281
[http://dx.doi.org/10.1016/S0012-821X\(01\)00309-0](http://dx.doi.org/10.1016/S0012-821X(01)00309-0)
- Granger D, Smith A. 2000. Dating buried sediments using radioactive decay and muogenic production of 26Al and 10Be. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **172**: 822-826
[http://dx.doi.org/10.1016/S0168-583X\(00\)00087-2](http://dx.doi.org/10.1016/S0168-583X(00)00087-2)
- Granger D, Lifton N, Willenbring J. 2013. A cosmic trip: 25 years of cosmogenic nuclides in geology. *Geological Society of America Bulletin* **125**: 1379-1402
<http://dx.doi.org/10.1130/b30774.1>
- Granger D, Fabel D, Palmer A. 2001. Pliocene–Pleistocene incision of the Green River, Kentucky, determined from radioactive decay of cosmogenic 26Al and 10Be in Mammoth Cave sediments. *Geological Society of America Bulletin* **113**: 825-836
[http://dx.doi.org/10.1130/0016-7606\(2001\)113<0825:ppiotg>2.0.co;2](http://dx.doi.org/10.1130/0016-7606(2001)113<0825:ppiotg>2.0.co;2)
- Granger D, Kirchner J, Finkel R. 1997. Quaternary downcutting rate of the New River, Virginia, measured from differential decay of cosmogenic 26Al and 10Be in cave-

- deposited alluvium. *Geology* **25**: 107-110
[http://dx.doi.org/10.1130/0091-7613\(1997\)025<0107:qdrotn>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1997)025<0107:qdrotn>2.3.co;2)
- Guido Z, Ward D, Anderson R. 2007. Pacing the post–Last Glacial Maximum demise of the Animas Valley glacier and the San Juan Mountain ice cap, Colorado. *Geology* **35**: 739-742 <http://dx.doi.org/10.1130/g23596a.1>
- Guyodo Y, Valet J-P. 1999. Global changes in intensity of the Earth's magnetic field during the past 800 kyr. *Nature* **399**: 249-252
<http://dx.doi.org/10.1038/20420>
- Hancock G, Anderson R, Chadwick O, Finkel R. 1999. Dating fluvial terraces with ¹⁰Be and ²⁶Al profiles: application to the Wind River, Wyoming. *Geomorphology* **27**: 41-60
[http://dx.doi.org/10.1016/S0169-555X\(98\)00089-0](http://dx.doi.org/10.1016/S0169-555X(98)00089-0)
- Harbor J, Stroeven A, Fabel D, Clarhäll A, Kleman J, Li Y, Elmore D, Fink D. 2006. Cosmogenic nuclide evidence for minimal erosion across two subglacial sliding boundaries of the late glacial Fennoscandian ice sheet. *Geomorphology* **75**: 90-99
<http://dx.doi.org/10.1016/j.geomorph.2004.09.036>
- Häuselmann P, Fiebig M, Kubik P, Adrian H. 2007. A first attempt to date the original “Deckenschotter” of Penck and Brückner with cosmogenic nuclides. *Quaternary International* **164–165**: 33-42
<http://dx.doi.org/10.1016/j.quaint.2006.12.013>
- Heimsath A, Furbish D, Dietrich W. 2005. The illusion of diffusion: Field evidence for depth-dependent sediment transport. *Geology* **33**: 949-952 <http://dx.doi.org/10.1130/g21868.1>
- Heimsath A, Chappell J, Spooner N, Questiaux D. 2002. Creeping soil. *Geology* **30**: 111-114 [http://dx.doi.org/10.1130/0091-7613\(2002\)030<0111:cs>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2002)030<0111:cs>2.0.co;2)
- Heimsath A, Dietrich W, Nishiizumi K, Finkel R. 2001a. Stochastic processes of soil production and transport: erosion rates, topographic variation and cosmogenic nuclides in the Oregon Coast Range. *Earth Surface Processes and Landforms* **26**: 531-552 <http://dx.doi.org/10.1002/esp.209>
- Heimsath A, Chappell J, Dietrich W, Nishiizumi K, Finkel R. 2001b. Late Quaternary erosion in southeastern Australia: a field example using cosmogenic nuclides. *Quaternary International* **83–85**: 169-185
[http://dx.doi.org/10.1016/S1040-6182\(01\)00038-6](http://dx.doi.org/10.1016/S1040-6182(01)00038-6)
- Heimsath A, Chappell J, Dietrich W, Nishiizumi K, Finkel R. 2000. Soil production on a retreating escarpment in southeastern Australia. *Geology* **28**: 787-790
[http://dx.doi.org/10.1130/0091-7613\(2000\)28<787:spoare>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2000)28<787:spoare>2.0.co;2)
- Heimsath A, E. Dietrich W, Nishiizumi K, Finkel R. 1999. Cosmogenic nuclides, topography, and the spatial variation of soil depth. *Geomorphology* **27**: 151-172
[http://dx.doi.org/10.1016/S0169-555X\(98\)00095-6](http://dx.doi.org/10.1016/S0169-555X(98)00095-6)
- Heimsath A, Dietrich W, Nishiizumi K, Finkel R. 1997. The soil production function and landscape equilibrium. *Nature* **388**: 358-361
- Hein A, Dunai T, Hulton N, Xu S. 2011a. Exposure dating outwash gravels to determine the age of the greatest Patagonian glaciations. *Geology* **39**: 103-106
<http://dx.doi.org/10.1130/g31215.1>
- Hein A, Fogwill C, Sugden D, Xu S. 2011b. Glacial/interglacial ice-stream stability in the Weddell Sea embayment, Antarctica. *Earth Planet. Sci. Lett.* **307**: 211-221
<http://dx.doi.org/10.1016/j.epsl.2011.04.037>
- Hein A, Hulton N, Dunai T, Schnabel C, Kaplan M, Naylor M, Xu S. 2009. Middle Pleistocene glaciation in Patagonia dated by cosmogenic-nuclide measurements on outwash gravels. *Earth Planet. Sci. Lett.* **286**: 184-197 <http://dx.doi.org/10.1016/j.epsl.2009.06.026>
- Heisinger B, Lal D, Jull A, Kubik P, Ivy-Ochs S, Knie K, Nolte E. 2002a. Production of selected cosmogenic radionuclides by muons: 2. Capture of negative muons. *Earth Planet. Sci. Lett.* **200**: 357-369
[http://dx.doi.org/10.1016/S0012-821X\(02\)00641-6](http://dx.doi.org/10.1016/S0012-821X(02)00641-6)
- Heisinger B, Lal D, Jull A, Kubik P, Ivy-Ochs S, Neumaier S, Knie K, Lazarev V, Nolte E. 2002b. Production of selected cosmogenic radionuclides by muons: 1. Fast muons. *Earth Planet. Sci. Lett.* **200**: 345-355
[http://dx.doi.org/10.1016/S0012-821X\(02\)00640-4](http://dx.doi.org/10.1016/S0012-821X(02)00640-4)
- Hetzl R, Niedermann S, Tao M, Kubik P, Ivy-Ochs S, Gao B, Strecker M. 2002. Low slip rates and long-term preservation of geomorphic features in Central Asia. *Nature* **417**: 428-432 <http://dx.doi.org/10.1038/417428a>
- Heyman J, Stroeven A, Harbor J, Caffee M. 2011. Too young or too old: Evaluating cosmogenic exposure dating based on an analysis of compiled boulder exposure ages. *Earth Planet. Sci. Lett.* **302**: 71-80
<http://dx.doi.org/10.1016/j.epsl.2010.11.040>
- Hidy A, Gosse J, Pederson J, Mattern J, Finkel R. 2010. A geologically constrained Monte Carlo approach to modeling exposure ages from profiles of cosmogenic nuclides:

- An example from Lees Ferry, Arizona. *Geochemistry, Geophysics, Geosystems* **11**: Q0AA10 <http://dx.doi.org/10.1029/2010gc003084>
- Hippe K, Ivy-Ochs S, Kober F, Zasadni J, Wieler R, Wacker L, Kubik P, Schlüchter C. 2013. Chronology of Lateglacial ice flow reorganization and deglaciation in the Gotthard Pass area, Central Swiss Alps, based on cosmogenic ^{10}Be and in situ ^{14}C . *Quaternary Geochronology* <http://dx.doi.org/10.1016/j.quageo.2013.03.003>
- Hippe K, Kober F, Zeilinger G, Ivy-Ochs S, Maden C, Wacker L, Kubik P, Wieler R. 2012. Quantifying denudation rates and sediment storage on the eastern Altiplano, Bolivia, using cosmogenic ^{10}Be , ^{26}Al , and in situ ^{14}C . *Geomorphology* **179**: 58-70 <http://dx.doi.org/10.1016/j.geomorph.2012.07.031>
- Hu X, Kirby E, Pan B, Granger D, Su H. 2011. Cosmogenic burial ages reveal sediment reservoir dynamics along the Yellow River, China. *Geology* **39**: 839-842 <http://dx.doi.org/10.1130/g32030.1>
- Hubbard B, Glasser G. 2005. Field techniques in glaciology and glacial geomorphology. Wiley, Chichester.
- Ivy-Ochs S, Kober F. 2008. Surface exposure dating with cosmogenic nuclides. *Eiszeitalter und Gegenwart - Quaternary Science Journal* **57**: 179-209 <http://dx.doi.org/10.3285/eg.57.1-2.7>
- Jackson L, Phillips F, Shimamura K, Little E. 1997. Cosmogenic ^{36}Cl dating of the Foothills erratics train, Alberta, Canada. *Geology* **25**: 195-198 [http://dx.doi.org/10.1130/0091-7613\(1997\)025<0195:ccdof>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1997)025<0195:ccdof>2.3.co;2)
- Jull A, Lal D, Donahue D, Mayewski P, Lorus C, Raynaud D, Petit J. 1994. Measurements of cosmic-ray-produced ^{14}C in firn and ice from antarctica. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **92**: 326-330 [http://dx.doi.org/10.1016/0168-583X\(94\)96028-3](http://dx.doi.org/10.1016/0168-583X(94)96028-3)
- Kaplan M, Schaefer J, Denton G, Doughty A, Barrell D, Chinn T, Putnam A, Andersen B, Mackintosh A, Finkel R, Schwartz R, Anderson B. 2013. The anatomy of long-term warming since 15 ka in New Zealand based on net glacier snowline rise. *Geology* **41**: 887-890 <http://dx.doi.org/10.1130/g34288.1>
- Kaplan M, Strelin J, Schaefer J, Denton G, Finkel R, Schwartz R, Putnam A, Vandergoes M, Goehring B, Travis S. 2011. In-situ cosmogenic ^{10}Be production rate at Lago Argentino, Patagonia: Implications for late-glacial climate chronology. *Earth Planet. Sci. Lett.* **309**: 21-32 <http://dx.doi.org/10.1016/j.epsl.2011.06.018>
- Kaplan M, Ackert R, Singer B, Douglass D, Kurz M. 2004. Cosmogenic nuclide chronology of millennial-scale glacial advances during O-isotope stage 2 in Patagonia. *Geological Society of America Bulletin* **116**: 308-321 <http://dx.doi.org/10.1130/b25178.1>
- Kirchner J, Finkel R, Riebe C, Granger D, Clayton J, King J, Megahan W. 2001. Mountain erosion over 10 yr, 10 k.y., and 10 m.y. time scales. *Geology* **29**: 591-594 [http://dx.doi.org/10.1130/0091-7613\(2001\)029<0591:meoyky>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2001)029<0591:meoyky>2.0.co;2)
- Kong P, Granger D, Wu F-Y, Caffee M, Wang Y-J, Zhao X-T, Zheng Y. 2009. Cosmogenic nuclide burial ages and provenance of the Xigeda paleo-lake: Implications for evolution of the Middle Yangtze River. *Earth Planet. Sci. Lett.* **278**: 131-141 <http://dx.doi.org/10.1016/j.epsl.2008.12.003>
- Korschinek G, Bergmaier A, Faestermann T, Gerstmann U, Knie K, Rugel G, Wallner A, Dillmann I, Dollinger G, von Gostomski C, Kossert K, Maiti M, Poutivtsev M, Remmert A. 2010. A new value for the half-life of ^{10}Be by Heavy-Ion Elastic Recoil Detection and liquid scintillation counting. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **268**: 187-191 <http://dx.doi.org/10.1016/j.nimb.2009.09.020>
- Kubik P, Ivy-Ochs S, Masarik J, Frank M, Schlüchter C. 1998. ^{10}Be and ^{26}Al production rates deduced from an instantaneous event within the dendro-calibration curve, the landslide of Köfels, Ötztal Valley, Austria. *Earth Planet. Sci. Lett.* **161**: 231-241 [http://dx.doi.org/10.1016/S0012-821X\(98\)00153-8](http://dx.doi.org/10.1016/S0012-821X(98)00153-8)
- Kurth G, Phillips F, Reheis M, Redwine J, Paces J. 2011. Cosmogenic nuclide and uranium-series dating of old, high shorelines in the western Great Basin, USA. *Geological Society of America Bulletin* **123**: 744-768 <http://dx.doi.org/10.1130/b30010.1>
- Kurz M, Colodner D, Trull T, Moore R, O'Brien K. 1990. Cosmic ray exposure dating with in situ produced cosmogenic ^3He : Results from young Hawaiian lava flows.

- Earth Planet. Sci. Lett.* **97**: 177-189
[http://dx.doi.org/10.1016/0012-821X\(90\)90107-9](http://dx.doi.org/10.1016/0012-821X(90)90107-9)
- Kurz M. 1986. In situ production of terrestrial cosmogenic helium and some applications to geochronology. *Geochimica Et Cosmochimica Acta* **50**: 2855-2862
[http://dx.doi.org/10.1016/0016-7037\(86\)90232-2](http://dx.doi.org/10.1016/0016-7037(86)90232-2)
- Lal D. 1991. Cosmic ray labeling of erosion surfaces: in situ nuclide production rates and erosion models. *Earth Planet. Sci. Lett.* **104**: 424-439
[http://dx.doi.org/10.1016/0012-821X\(91\)90220-C](http://dx.doi.org/10.1016/0012-821X(91)90220-C)
- Lal D, Jull A. 1992. Cosmogenic nuclides in ice sheets. *Radiocarbon*: 227-233
- Lal D, Jull A. 1990. On determining ice accumulation rates in the past 40,000 years using in situ cosmogenic ¹⁴C. *Geophysical Research Letters* **17**: 1303-1306
<http://dx.doi.org/10.1029/GL017i009p01303>
- Lal D, Jull A, Donahue D, Burtner D, Nishiizumi K. 1990. Polar ice ablation rates measured using in-situ cosmogenic C-14. *Nature Geosci* **346**: 350-352
- Lal D, Nishiizumi K, Arnold J. 1987. In situ cosmogenic ³H, ¹⁴C, and ¹⁰Be for determining the net accumulation and ablation rates of ice sheets. *Journal of Geophysical Research: Solid Earth* **92**: 4947-4952
<http://dx.doi.org/10.1029/JB092iB06p04947>
- Laughlin A, Poths J, Healey H, Reneau S, WoldeGabriel G. 1994. Dating of Quaternary basalts using the cosmogenic ³He and ¹⁴C methods with implications for excess ⁴⁰Ar. *Geology* **22**: 135-138
[http://dx.doi.org/10.1130/0091-7613\(1994\)022<0135:doqbut>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1994)022<0135:doqbut>2.3.co;2)
- Leland J, Reid M, Burbank D, Finkel R, Caffee M. 1998. Incision and differential bedrock uplift along the Indus River near Nanga Parbat, Pakistan Himalaya, from ¹⁰Be and ²⁶Al exposure age dating of bedrock straths. *Earth Planet. Sci. Lett.* **154**: 93-107
[http://dx.doi.org/10.1016/S0012-821X\(97\)00171-4](http://dx.doi.org/10.1016/S0012-821X(97)00171-4)
- Licciardi J, Kurz M, Clark P, Brook E. 1999. Calibration of cosmogenic ³He production rates from Holocene lava flows in Oregon, USA, and effects of the Earth's magnetic field. *Earth Planet. Sci. Lett.* **172**: 261-271
[http://dx.doi.org/10.1016/S0012-821X\(99\)00204-6](http://dx.doi.org/10.1016/S0012-821X(99)00204-6)
- Lifton N, Bieber J, Clem J, Duldig M, Evenson P, Humble J, Pyle R. 2005. Addressing solar modulation and long-term uncertainties in scaling secondary cosmic rays for in situ cosmogenic nuclide applications. *Earth Planet. Sci. Lett.* **239**: 140-161
<http://dx.doi.org/10.1016/j.epsl.2005.07.001>
- Mackey B, Lamb M. 2013. Deciphering boulder mobility and erosion from cosmogenic nuclide exposure dating. *Journal of Geophysical Research: Earth Surface* **118**: 184-197
<http://dx.doi.org/10.1002/jgrf.20035>
- Mackintosh A, White D, Fink D, Gore D, Pickard J, Fanning P. 2007. Exposure ages from mountain dipsticks in Mac. Robertson Land, East Antarctica, indicate little change in ice-sheet thickness since the Last Glacial Maximum. *Geology* **35**: 551-554
<http://dx.doi.org/10.1130/g23503a.1>
- Marchant D, Lewis A, Phillips W, Moore E, Souchez R, Denton G, Sugden D, Potter N, Landis G. 2002. Formation of patterned ground and sublimation till over Miocene glacier ice in Beacon Valley, southern Victoria Land, Antarctica. *Geological Society of America Bulletin* **114**: 718-730
[http://dx.doi.org/10.1130/0016-7606\(2002\)114<0718:fopgas>2.0.co;2](http://dx.doi.org/10.1130/0016-7606(2002)114<0718:fopgas>2.0.co;2)
- Marchetti D, Cerling T. 2005. Cosmogenic ³He exposure ages of Pleistocene debris flows and desert pavements in Capitol Reef National Park, Utah. *Geomorphology* **67**: 423-435
<http://dx.doi.org/10.1016/j.geomorph.2004.11.004>
- Margerison H, Phillips W, Stuart F, Sugden D. 2005. Cosmogenic ³He concentrations in ancient flood deposits from the Coombs Hills, northern Dry Valleys, East Antarctica: interpreting exposure ages and erosion rates. *Earth Planet. Sci. Lett.* **230**: 163-175
<http://dx.doi.org/10.1016/j.epsl.2004.11.007>
- Marquette G, Gray J, Gosse J, Courchesne F, Stockli L, Macpherson G, Finkel R. 2004. Felsenmeer persistence under non-erosive ice in the Torngat and Kaumajet mountains, Quebec and Labrador, as determined by soil weathering and cosmogenic nuclide exposure dating. *Canadian Journal of Earth Sciences* **41**: 19-38
<http://dx.doi.org/10.1139/e03-072>
- Masarik J, Wieler R. 2003. Production rates of cosmogenic nuclides in boulders. *Earth Planet. Sci. Lett.* **216**: 201-208
[http://dx.doi.org/10.1016/S0012-821X\(03\)00476-X](http://dx.doi.org/10.1016/S0012-821X(03)00476-X)
- Masarik J, Beer J. 1999. Simulation of particle fluxes and cosmogenic nuclide production in the Earth's atmosphere. *Journal of Geophysical Research: Atmospheres* **104**: 12099-12111
<http://dx.doi.org/10.1029/1998jd200091>

- Masarik J, Reedy R. 1995. Terrestrial cosmogenic-nuclide production systematics calculated from numerical simulations. *Earth Planet. Sci. Lett.* **136**: 381-395 [http://dx.doi.org/10.1016/0012-821X\(95\)00169-D](http://dx.doi.org/10.1016/0012-821X(95)00169-D)
- Matmon A, Ron H, Chazan M, Porat N, Horwitz L. 2012. Reconstructing the history of sediment deposition in caves: A case study from Wonderwerk Cave, South Africa. *Geological Society of America Bulletin* **124**: 611-625 <http://dx.doi.org/10.1130/b30410.1>
- Matmon A, Simhai O, Amit R, Haviv I, Porat N, McDonald E, Benedetti L, Finkel R. 2009. Desert pavement-coated surfaces in extreme deserts present the longest-lived landforms on Earth. *Geological Society of America Bulletin* **121**: 688-697 <http://dx.doi.org/10.1130/b26422.1>
- Matmon A, Crouvi O, Enzel Y, Bierman P, Larsen J, Porat N, Amit R, Caffee M. 2003. Complex exposure histories of chert clasts in the late Pleistocene shorelines of Lake Lisan, southern Israel. *Earth Surface Processes and Landforms* **28**: 493-506 <http://dx.doi.org/10.1002/esp.454>
- McCormack D, Brocklehurst S, Irving D, Fabel D. 2011. Cosmogenic ^{10}Be insights into the extent and chronology of the last deglaciation in Wester Ross, northwest Scotland. *Journal of Quaternary Science* **26**: 97-108 <http://dx.doi.org/10.1002/jqs.1437>
- Medynski S, Pik R, Burnard P, Williams A, Vye-Brown C, Ferguson D, Blard P, France L, Yirgu G, Seid J, Ayalew D, Calvert A. 2013. Controls on magmatic cycles and development of rift topography of the Manda Hararo segment (Afar, Ethiopia): Insights from cosmogenic ^3He investigation of landscape evolution. *Earth Planet. Sci. Lett.* **367**: 133-145 <http://dx.doi.org/10.1016/j.epsl.2013.02.006>
- Mitchell S, Matmon A, Bierman P, Enzel Y, Caffee M, Rizzo D. 2001. Displacement history of a limestone normal fault scarp, northern Israel, from cosmogenic ^{36}Cl . *Journal of Geophysical Research: Solid Earth* **106**: 4247-4264 <http://dx.doi.org/10.1029/2000jb900373>
- Molnar P, Erik Thorson B, Burchfiel B, Deng Q, Feng X, Li J, Raisbeck G, Shi J, Zhangming W, Yiou F, You H. 1994. Quaternary Climate Change and the Formation of River Terraces across Growing Anticlines on the North Flank of the Tien Shan, China. *The Journal of Geology* **102**: 583-602 <http://dx.doi.org/10.2307/30068558>
- Nichols K, Bierman P, Eppes M, Caffee M, Finkel R, Larsen J. 2007. Timing of surficial process changes down a Mojave Desert piedmont. *Quaternary Research* **68**: 151-161 <http://dx.doi.org/10.1016/j.yqres.2007.02.001>
- Nichols K, Bierman P, Caffee M, Finkel R, Larsen J. 2005. Cosmogenically enabled sediment budgeting. *Geology* **33**: 133-136 <http://dx.doi.org/10.1130/g21006.1>
- Nishiizumi K, Caffee M, Finkel R, Brimhall G, Mote T. 2005. Remnants of a fossil alluvial fan landscape of Miocene age in the Atacama Desert of northern Chile using cosmogenic nuclide exposure age dating. *Earth Planet. Sci. Lett.* **237**: 499-507 <http://dx.doi.org/10.1016/j.epsl.2005.05.032>
- Nishiizumi K, Kohl C, Arnold J, Klein J, Fink D, Middleton R. 1991a. Cosmic ray produced ^{10}Be and ^{26}Al in Antarctic rocks: exposure and erosion history. *Earth Planet. Sci. Lett.* **104**: 440-454 [http://dx.doi.org/10.1016/0012-821X\(91\)90221-3](http://dx.doi.org/10.1016/0012-821X(91)90221-3)
- Nishiizumi K, Kohl C, Shoemaker E, Arnold J, Klein J, Fink D, Middleton R. 1991b. In situ ^{10}Be - ^{26}Al exposure ages at Meteor Crater, Arizona. *Geochimica Et Cosmochimica Acta* **55**: 2699-2703 [http://dx.doi.org/10.1016/0016-7037\(91\)90388-L](http://dx.doi.org/10.1016/0016-7037(91)90388-L)
- Nishiizumi K, Winterer E, Kohl C, Klein J, Middleton R, Lal D, Arnold J. 1989. Cosmic ray production rates of ^{10}Be and ^{26}Al in quartz from glacially polished rocks. *Journal of Geophysical Research: Solid Earth* **94**: 17907-17915 <http://dx.doi.org/10.1029/JB094iB12p17907>
- Nishiizumi K, Lal D, Klein J, Middleton R, Arnold J. 1986. Production of ^{10}Be and ^{26}Al by cosmic rays in terrestrial quartz in situ and implications for erosion rates. *Nature* **319**: 134-136 <http://dx.doi.org/10.1038/319134a0>
- Norton K, Blanckenburg F, DiBiase R, Schlunegger F, Kubik P. 2011. Cosmogenic ^{10}Be -derived denudation rates of the Eastern and Southern European Alps. *Int J Earth Sci (Geol Rundsch)* **100**: 1163-1179 <http://dx.doi.org/10.1007/s00531-010-0626-y>
- Norton K, von Blanckenburg F, Schlunegger F, Schwab M, Kubik P. 2008. Cosmogenic nuclide-based investigation of spatial erosion and hillslope channel coupling in the transient foreland of the Swiss Alps. *Geomorphology*

- 95:** 474-486
<http://dx.doi.org/10.1016/j.geomorph.2007.07.013>
- Owen L, Bright J, Finkel R, Jaiswal M, Kaufman D, Mahan S, Radtke U, Schneider J, Sharp W, Singhvi A, Warren C. 2007. Numerical dating of a Late Quaternary spit-shoreline complex at the northern end of Silver Lake playa, Mojave Desert, California: A comparison of the applicability of radiocarbon, luminescence, terrestrial cosmogenic nuclide, electron spin resonance, U-series and amino acid racemization methods. *Quaternary International* **166**: 87-110 <http://dx.doi.org/10.1016/j.quaint.2007.01.001>
- Owen L, Finkel R, Caffee M, Gualtieri L. 2002. Timing of multiple late Quaternary glaciations in the Hunza Valley, Karakoram Mountains, northern Pakistan: Defined by cosmogenic radionuclide dating of moraines. *Geological Society of America Bulletin* **114**: 593-604 [http://dx.doi.org/10.1130/0016-7606\(2002\)114<0593:tomlqg>2.0.co;2](http://dx.doi.org/10.1130/0016-7606(2002)114<0593:tomlqg>2.0.co;2)
- Owen L, Gualtieri L, Finkel R, Caffee M, Benn D, Sharma M. 2001. Cosmogenic radionuclide dating of glacial landforms in the Lahul Himalaya, northern India: defining the timing of Late Quaternary glaciation. *Journal of Quaternary Science* **16**: 555-563 <http://dx.doi.org/10.1002/jqs.621>
- Palumbo L, Benedetti L, Bourlès D, Cinque A, Finkel R. 2004. Slip history of the Magnola fault (Apennines, Central Italy) from ³⁶Cl surface exposure dating: evidence for strong earthquakes over the Holocene. *Earth Planet. Sci. Lett.* **225**: 163-176 <http://dx.doi.org/10.1016/j.epsl.2004.06.012>
- Perg L, Anderson R, Finkel R. 2001. Use of a new ¹⁰Be and ²⁶Al inventory method to date marine terraces, Santa Cruz, California, USA. *Geology* **29**: 879-882 [http://dx.doi.org/10.1130/0091-7613\(2001\)029<0879:uoanba>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2001)029<0879:uoanba>2.0.co;2)
- Phillips F, Zreda M, Benson L, Plummer M, Elmore D, Sharma P. 1996. Chronology for Fluctuations in Late Pleistocene Sierra Nevada Glaciers and Lakes. *Science* **274**: 749-751 <http://dx.doi.org/10.1126/science.274.5288.749>
- Phillips F, Zreda M, Smith S, Elmore D, Kubik P, Dorn R, Roddy D. 1991. Age and geomorphic history of Meteor Crater, Arizona, from cosmogenic ³⁶Cl and ¹⁴C in rock varnish. *Geochimica Et Cosmochimica Acta* **55**: 2695-2698 [http://dx.doi.org/10.1016/0016-7037\(91\)90387-K](http://dx.doi.org/10.1016/0016-7037(91)90387-K)
- Phillips F, Zreda M, Smith S, Elmore D, Kubik P, Sharma P. 1990. Cosmogenic Chlorine-36 Chronology for Glacial Deposits at Bloody Canyon, Eastern Sierra Nevada. *Science* **248**: 1529-1532 <http://dx.doi.org/10.1126/science.248.4962.1529>
- Phillips W, Hall A, Mottram R, Fifield L, Sugden D. 2006. Cosmogenic ¹⁰Be and ²⁶Al exposure ages of tors and erratics, Cairngorm Mountains, Scotland: Timescales for the development of a classic landscape of selective linear glacial erosion. *Geomorphology* **73**: 222-245 <http://dx.doi.org/10.1016/j.geomorph.2005.06.009>
- Poreda R, Cerling T. 1992. Cosmogenic neon in recent lavas from the western United States. *Geophysical Research Letters* **19**: 1863-1866 <http://dx.doi.org/10.1029/92gl01998>
- Pratt B, Burbank D, Heimsath A, Ojha T. 2002. Impulsive alluviation during early Holocene strengthened monsoons, central Nepal Himalaya. *Geology* **30**: 911-914 [http://dx.doi.org/10.1130/0091-7613\(2002\)030<0911:iadehs>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2002)030<0911:iadehs>2.0.co;2)
- Putkonen J, O'Neal M. 2006. Degradation of unconsolidated Quaternary landforms in the western North America. *Geomorphology* **75**: 408-419 <http://dx.doi.org/10.1016/j.geomorph.2005.07.024>
- Putkonen J, Swanson T. 2003. Accuracy of cosmogenic ages for moraines. *Quaternary Research* **59**: 255-261 [http://dx.doi.org/10.1016/S0033-5894\(03\)00006-1](http://dx.doi.org/10.1016/S0033-5894(03)00006-1)
- Putnam A, Schaefer J, Denton G, Barrell D, Finkel R, Andersen B, Schwartz R, Chinn T, Doughty A. 2012. Regional climate control of glaciers in New Zealand and Europe during the pre-industrial Holocene. *Nature Geosci* **5**: 627-630 <http://dx.doi.org/10.1038/ngeo1548>
- Putnam A, Denton G, Schaefer J, Barrell D, Andersen B, Finkel R, Schwartz R, Doughty A, Kaplan M, Schluchter C. 2010a. Glacier advance in southern middle-latitudes during the Antarctic Cold Reversal. *Nature Geosci* **3**: 700-704 <http://dx.doi.org/10.1038/ngeo962>
- Putnam A, Schaefer J, Barrell D, Vandergoes M, Denton G, Kaplan M, Finkel R, Schwartz R, Goehring B, Kelley S. 2010b. In situ cosmogenic ¹⁰Be production-rate calibration from the Southern Alps, New Zealand. *Quaternary Geochronology* **5**: 392-409 <http://dx.doi.org/10.1016/j.quageo.2009.12.001>
- Repka J, Anderson R, Finkel R. 1997. Cosmogenic dating of fluvial terraces,

- Fremont River, Utah. *Earth Planet. Sci. Lett.* **152**: 59-73 [http://dx.doi.org/10.1016/S0012-821X\(97\)00149-0](http://dx.doi.org/10.1016/S0012-821X(97)00149-0)
- Reuther A, Herget J, Ivy-Ochs S, Borodavko P, Kubik P, Heine K. 2006. Constraining the timing of the most recent cataclysmic flood event from ice-dammed lakes in the Russian Altai Mountains, Siberia, using cosmogenic in situ ¹⁰Be. *Geology* **34**: 913-916 <http://dx.doi.org/10.1130/g22755a.1>
- Riebe C, Granger D. 2013. Quantifying effects of deep and near-surface chemical erosion on cosmogenic nuclides in soils, saprolite, and sediment. *Earth Surface Processes and Landforms* **38**: 523-533 <http://dx.doi.org/10.1002/esp.3339>
- Riebe C, Kirchner J, Granger D, Finkel R. 2001a. Minimal climatic control on erosion rates in the Sierra Nevada, California. *Geology* **29**: 447-450 [http://dx.doi.org/10.1130/0091-7613\(2001\)029<0447:mccoer>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2001)029<0447:mccoer>2.0.co;2)
- Riebe C, Kirchner J, Granger D, Finkel R. 2001b. Strong tectonic and weak climatic control of long-term chemical weathering rates. *Geology* **29**: 511-514 [http://dx.doi.org/10.1130/0091-7613\(2001\)029<0511:stawcc>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2001)029<0511:stawcc>2.0.co;2)
- Riebe C, Kirchner J, Granger D, Finkel R. 2000. Erosional equilibrium and disequilibrium in the Sierra Nevada, inferred from cosmogenic ²⁶Al and ¹⁰Be in alluvial sediment. *Geology* **28**: 803-806 [http://dx.doi.org/10.1130/0091-7613\(2000\)28<803:eedit>2.0.co;2](http://dx.doi.org/10.1130/0091-7613(2000)28<803:eedit>2.0.co;2)
- Riihimäki C, Libarkin J. 2007. Terrestrial Cosmogenic Nuclides as Paleoaltimetric Proxies. *Reviews in Mineralogy and Geochemistry* **66**: 269-278 <http://dx.doi.org/10.2138/rmg.2007.66.11>
- Rinterknecht V, Matoshko A, Gorokhovich Y, Fabel D, Xu S. 2012a. Expression of the Younger Dryas cold event in the Carpathian Mountains, Ukraine? *Quat. Sci. Rev.* **39**: 106-114 <http://dx.doi.org/10.1016/j.quascirev.2012.02.005>
- Rinterknecht V, Braucher R, Böse M, Bourlès D, Mercier J. 2012b. Late Quaternary ice sheet extents in northeastern Germany inferred from surface exposure dating. *Quat. Sci. Rev.* **44**: 89-95 <http://dx.doi.org/10.1016/j.quascirev.2010.07.026>
- Rinterknecht V, Clark P, Raisbeck G, Yiou F, Bitinas A, Brook E, Marks L, Zelčs V, Lunkka J-P, Pavlovskaya I, Piotrowski J, Raukas A. 2006. The Last Deglaciation of the Southeastern Sector of the Scandinavian Ice Sheet. *Science* **311**: 1449-1452 <http://dx.doi.org/10.1126/science.1120702>
- Ruszkiczay-Rüdiger Z, Braucher R, Csillag G, Fodor L, Dunai T, Bada G, Bourlès D, Müller P. 2011. Dating Pleistocene aeolian landforms in Hungary, Central Europe, using in situ produced cosmogenic ¹⁰Be. *Quaternary Geochronology* **6**: 515-529 <http://dx.doi.org/10.1016/j.quageo.2011.06.001>
- Ruszkiczay-Rüdiger Z, Dunai T, Bada G, Fodor L, Horváth E. 2005. Middle to late Pleistocene uplift rate of the Hungarian Mountain Range at the Danube Bend, (Pannonian Basin) using in situ produced ³He. *Tectonophysics* **410**: 173-187 <http://dx.doi.org/10.1016/j.tecto.2005.02.017>
- Sanchez G, Rolland Y, Corsini M, Braucher R, Bourlès D, Arnold M, Aumaître G. 2010. Relationships between tectonics, slope instability and climate change: Cosmic ray exposure dating of active faults, landslides and glacial surfaces in the SW Alps. *Geomorphology* **117**: 1-13 <http://dx.doi.org/10.1016/j.geomorph.2009.10.019>
- Schaefer J, Denton G, Barrell D, Ivy-Ochs S, Kubik P, Andersen B, Phillips F, Lowell T, Schlüchter C. 2006. Near-Synchronous Interhemispheric Termination of the Last Glacial Maximum in Mid-Latitudes. *Science* **312**: 1510-1513 <http://dx.doi.org/10.1126/science.1122872>
- Schäfer J, Tschudi S, Zhao Z, Wu X, Ivy-Ochs S, Wieler R, Baur H, Kubik P, Schlüchter C. 2002. The limited influence of glaciations in Tibet on global climate over the past 170 000 yr. *Earth Planet. Sci. Lett.* **194**: 287-297 [http://dx.doi.org/10.1016/S0012-821X\(01\)00573-8](http://dx.doi.org/10.1016/S0012-821X(01)00573-8)
- Schäfer J, Baur H, Denton G, Ivy-Ochs S, Marchant D, Schlüchter C, Wieler R. 2000. The oldest ice on Earth in Beacon Valley, Antarctica: new evidence from surface exposure dating. *Earth Planet. Sci. Lett.* **179**: 91-99 [http://dx.doi.org/10.1016/S0012-821X\(00\)00095-9](http://dx.doi.org/10.1016/S0012-821X(00)00095-9)
- Schäfer J, Ivy-Ochs S, Wieler R, Leya I, Baur H, Denton G, Schlüchter C. 1999. Cosmogenic noble gas studies in the oldest landscape on earth: surface exposure ages of the Dry Valleys, Antarctica. *Earth Planet. Sci. Lett.* **167**: 215-226 [http://dx.doi.org/10.1016/S0012-821X\(99\)00029-1](http://dx.doi.org/10.1016/S0012-821X(99)00029-1)
- Schaller M, Ehlers T, Blum J, Kallenberg M. 2009. Quantifying glacial moraine age,

- denudation, and soil mixing with cosmogenic nuclide depth profiles. *Journal of Geophysical Research: Earth Surface* **114**: F01012 <http://dx.doi.org/10.1029/2007jf000921>
- Schaller M, Hovius N, Willett S, Ivy-Ochs S, Synal H, Chen M. 2005. Fluvial bedrock incision in the active mountain belt of Taiwan from in situ-produced cosmogenic nuclides. *Earth Surface Processes and Landforms* **30**: 955-971 <http://dx.doi.org/10.1002/esp.1256>
- Schaller M, von Blanckenburg F, Hovius N, Kubik P. 2001. Large-scale erosion rates from in situ-produced cosmogenic nuclides in European river sediments. *Earth Planet. Sci. Lett.* **188**: 441-458 [http://dx.doi.org/10.1016/S0012-821X\(01\)00320-X](http://dx.doi.org/10.1016/S0012-821X(01)00320-X)
- Schildgen T, Dethier D, Bierman P, Caffee M. 2002. ²⁶Al and ¹⁰Be dating of late pleistocene and holocene fill terraces: a record of fluvial deposition and incision, Colorado front range. *Earth Surface Processes and Landforms* **27**: 773-787 <http://dx.doi.org/10.1002/esp.352>
- Schimmelpfennig I, Williams A, Pik R, Burnard P, Niedermann S, Finkel R, Schneider B, Benedetti L. 2011. Inter-comparison of cosmogenic in-situ ³He, ²¹Ne and ³⁶Cl at low latitude along an altitude transect on the SE slope of Kilimanjaro volcano (3°S, Tanzania). *Quaternary Geochronology* **6**: 425-436 <http://dx.doi.org/10.1016/j.quageo.2011.05.002>
- Schimmelpfennig I, Benedetti L, Finkel R, Pik R, Blard P-H, Bourlès D, Burnard P, Williams A. 2009. Sources of in-situ ³⁶Cl in basaltic rocks. Implications for calibration of production rates. *Quaternary Geochronology* **4**: 441-461 <http://dx.doi.org/10.1016/j.quageo.2009.06.003>
- Schlagenhauf A, Gaudemer Y, Benedetti L, Manighetti I, Palumbo L, Schimmelpfennig I, Finkel R, Pou K. 2010. Using in situ Chlorine-36 cosmonuclide to recover past earthquake histories on limestone normal fault scarps: a reappraisal of methodology and interpretations. *Geophysical Journal International* **182**: 36-72 <http://dx.doi.org/10.1111/j.1365-246X.2010.04622.x>
- Siame L, Bourlès D, Sébrier M, Bellier O, Carlos Castano J, Araujo M, Perez M, Raisbeck G, Yiou F. 1997. Cosmogenic dating ranging from 20 to 700 ka of a series of alluvial fan surfaces affected by the El Tigre fault, Argentina. *Geology* **25**: 975-978 [http://dx.doi.org/10.1130/0091-7613\(1997\)025<0975:cdfftk>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1997)025<0975:cdfftk>2.3.co;2)
- Small E, Anderson R, Hancock G. 1999. Estimates of the rate of regolith production using ¹⁰Be and ²⁶Al from an alpine hillslope. *Geomorphology* **27**: 131-150 [http://dx.doi.org/10.1016/S0169-555X\(98\)00094-4](http://dx.doi.org/10.1016/S0169-555X(98)00094-4)
- Small E, Anderson R, Repka J, Finkel R. 1997. Erosion rates of alpine bedrock summit surfaces deduced from in situ ¹⁰Be and ²⁶Al. *Earth Planet. Sci. Lett.* **150**: 413-425 [http://dx.doi.org/10.1016/S0012-821X\(97\)00092-7](http://dx.doi.org/10.1016/S0012-821X(97)00092-7)
- Stock G, Riihimaki C, Anderson R. 2006. Age constraints on cave development and landscape evolution in the Bighorn Basin of Wyoming, USA. *Journal of Cave and Karst Studies* **68**: 76-84
- Stock G, Anderson R, Finkel R. 2005. Rates of erosion and topographic evolution of the Sierra Nevada, California, inferred from cosmogenic ²⁶Al and ¹⁰Be concentrations. *Earth Surface Processes and Landforms* **30**: 985-1006 <http://dx.doi.org/10.1002/esp.1258>
- Stock G, Anderson R, Finkel R. 2004. Pace of landscape evolution in the Sierra Nevada, California, revealed by cosmogenic dating of cave sediments. *Geology* **32**: 193-196 <http://dx.doi.org/10.1130/g20197.1>
- Stone J. 2000. Air pressure and cosmogenic isotope production. *J. Geophys. Res.* **105**: 23753-23759 <http://dx.doi.org/10.1029/2000jb900181>
- Stone J, Balco G, Sugden D, Caffee M, Sass L, Cowdery S, Siddoway C. 2003. Holocene Deglaciation of Marie Byrd Land, West Antarctica. *Science* **299**: 99-102 <http://dx.doi.org/10.1126/science.1077998>
- Stone J, Ballantyne C, Fifield L. 1998. Exposure dating and validation of periglacial weathering limits, northwest Scotland. *Geology* **26**: 587-590 [http://dx.doi.org/10.1130/0091-7613\(1998\)026<0587:edavop>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1998)026<0587:edavop>2.3.co;2)
- Stone J, Lambeck K, Fifield L, Evans J, Cresswell R. 1996. A Lateglacial age for the Main Rock Platform, western Scotland. *Geology* **24**: 707-710 [http://dx.doi.org/10.1130/0091-7613\(1996\)024<0707:alafm>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1996)024<0707:alafm>2.3.co;2)
- Stroeven A, Fabel D, Hättestrand C, Harbor J. 2002. A relict landscape in the centre of Fennoscandian glaciation: cosmogenic radionuclide evidence of tors preserved through multiple glacial cycles.

Geomorphology **44**: 145-154
[http://dx.doi.org/10.1016/S0169-555X\(01\)00150-7](http://dx.doi.org/10.1016/S0169-555X(01)00150-7)

Todd C, Stone J, Conway H, Hall B, Bromley G. 2010. Late Quaternary evolution of Reedy Glacier, Antarctica. *Quat. Sci. Rev.* **29**: 1328-1341 <http://dx.doi.org/10.1016/j.quascirev.2010.02.001>

Trull T, Brown E, Marty B, Raisbeck G, Yiou F. 1995. Cosmogenic ^{10}Be and ^3He accumulation in Pleistocene beach terraces in Death Valley, California, U.S.A.: Implications for cosmic-ray exposure dating of young surfaces in hot climates. *Chemical Geology* **119**: 191-207
[http://dx.doi.org/10.1016/0009-2541\(94\)00092-M](http://dx.doi.org/10.1016/0009-2541(94)00092-M)

Van der Wateren F, Dunai T, Van Balen R, Klas W, Verbers A, Passchier S, Herpers U. 1999. Contrasting Neogene denudation histories of different structural regions in the Transantarctic Mountains rift flank constrained by cosmogenic isotope measurements. *Global and Planetary Change* **23**: 145-172 [http://dx.doi.org/10.1016/S0921-8181\(99\)00055-7](http://dx.doi.org/10.1016/S0921-8181(99)00055-7)

Vermeesch P. 2007. CosmoCalc: An Excel add-in for cosmogenic nuclide calculations. *Geochemistry, Geophysics, Geosystems* **8**: Q08003 <http://dx.doi.org/10.1029/2006gc001530>

Vermeesch P, Fenton C, Kober F, Wiggs G, Bristow C, Xu S. 2010. Sand residence times of one million years in the Namib Sand Sea from cosmogenic nuclides. *Nature Geosci* **3**: 862-865 <http://dx.doi.org/10.1038/ngeo985>

Wells S, McFadden L, Poths J, Olinger C. 1995. Cosmogenic ^3He surface-exposure dating of stone pavements: Implications for landscape evolution in deserts. *Geology* **23**: 613-616 [http://dx.doi.org/10.1130/0091-7613\(1995\)023<0613:chsedo>2.3.co;2](http://dx.doi.org/10.1130/0091-7613(1995)023<0613:chsedo>2.3.co;2)

Wilson P, Bentley M, Schnabel C, Clark R, Xu S. 2008. Stone run (block stream) formation in the Falkland Islands over several cold stages, deduced from cosmogenic isotope (^{10}Be and ^{26}Al) surface exposure dating. *Journal of Quaternary Science* **23**: 461-473 <http://dx.doi.org/10.1002/jqs.1156>

Young N, Schaefer J, Briner J, Goehring B. 2013. A ^{10}Be production-rate calibration for the Arctic. *Journal of Quaternary Science* **28**: 515-526 <http://dx.doi.org/10.1002/jqs.2642>

Zreda M, Noller J. 1998. Ages of Prehistoric Earthquakes Revealed by Cosmogenic Chlorine-36 in a Bedrock Fault Scarp at Hebgen Lake. *Science* **282**: 1097-1099
<http://dx.doi.org/10.1126/science.282.5391.1097>

Modelling Geomorphic Systems: Numerical Modelling

Christopher J. Hutton¹

¹Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK (c.j.hutton@ex.ac.uk)



Numerical models seek to represent the interaction between landscape forms and processes through mathematical equations. By integrating these equations over space and time, numerical models have allowed geomorphologists to extend enquiry beyond observation alone, and explore landscape dynamics over a range of temporal and spatial scales. Choosing the correct temporal and spatial scale of investigation, the correct processes that control landscape form at these scales, and then converting this conceptual model to a mathematical representation of these process-form interactions is not straightforward. The decision requires careful consideration of process dominance and scale, the ability of equations to parameterise these processes, computational resources, and data availability to constrain model parameters and evaluate model performance. These issues shall be considered in general terms, and illustrated mainly with reference to catchment systems. Finally, numerical modelling of geomorphic systems is considered from a Bayesian perspective to provide a conceptual grounding for the development and application of numerical models, and therefore for their role in geomorphic enquiry.

KEYWORDS: Numerical modelling; Scale; Resolution; Evaluation; Data; Uncertainty.

Introduction

The interaction of landscape form and process may be represented mathematically in the form of a numerical model. Coupling such models with observations provides a formal framework to assemble scientific understanding, and a powerful tool to investigate landscape change. Numerical models provide some liberation from the temporal and spatial shackles imposed by enquiry through observation alone; models have allowed exploration of landscape processes and evolution over spatial scales ranging from particles to plate tectonics and temporal scales ranging from milliseconds to millennia (Figure 1; Bishop, 2007; Hardy, 2005). Numerical models have shown potential as powerful tools for understanding reductionist process-form interactions (e.g. Schmeeckle and Nelson, 2003; Wainwright et al., 2008a), and also the relative importance of autogenic versus allogenic controls for larger scale system behaviour (Coulthard et al., 2005; Nicholas and Quine, 2007b; Wainwright and Parsons, 2002).

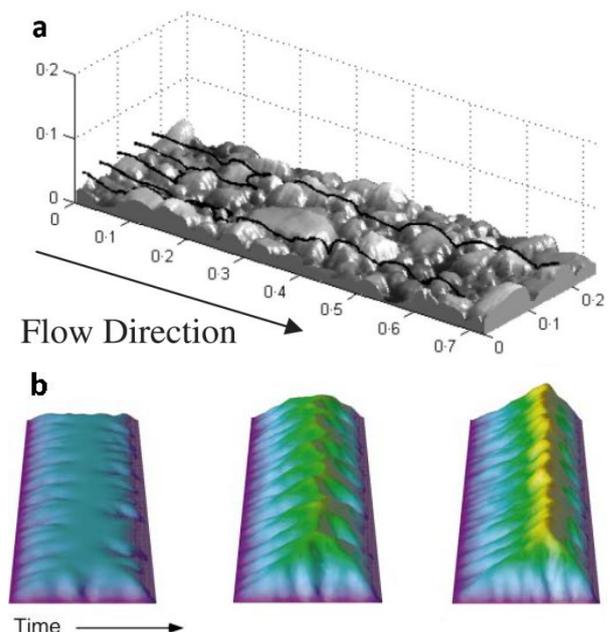


Figure 1. Contrasting scales of model application: grain scale predictions of particle paths (black lines) over water worked gravel (modified from Hardy, 2005); convergent orogen formation modelling, considering tectonic uplift and surface erosion (modified from Willett and Brandon, 2002).

In order to fully exploit the potential for numerical models to elicit understanding of form-process interactions, and inform landscape management, numerical models must be developed and applied carefully considering three key questions:

1. What are the relevant form-process interactions at the scale of enquiry?
2. What are the correct mathematical representations of these processes?
3. Are there appropriate data to constrain model parameters and evaluate model predictions?

The main modelling issues that need to be considered in order to address these questions are presented here to provide a basis for more domain specific sections in Chapter 5, and illustrated mainly with reference to catchment systems. The strength of the assumptions made in developing and applying a model will determine the validity of model predictions, the strength of conclusions derived from model application, and therefore the ability of models to inform us of real world process-form phenomena. A Bayesian approach emphasising iterative dialogue between model development and data collection is recommended as a robust means to appropriately develop numerical models, and therefore geomorphic understanding.

Publications reviewing specific areas of geomorphic modelling (e.g. Bishop, 2007; Coulthard, 2001; Livingstone et al., 2007; Merritt et al., 2003; Morgan and Nearing, 2010; Pelletier, 2011; Reinhardt et al., 2010; Tucker and Hancock, 2010; Van de Wiel et al., 2011; Wainwright et al., 2008a), and publications expanding on the more general issues of model application to natural systems considered here (e.g. Beven, 2002; Bloschl and Sivapalan, 1995; Brazier et al., 2011; Church, 1996; Krueger et al., 2009; Nicholas, 2005; Refsgaard et al., 2006; Van de Wiel et al., 2011; Wainwright and Mulligan, 2004; Wilcock and Iverson, 2003) are additionally recommended.

Model structure

A model M contains equations with associated parameters θ that represent the functional relationship between a vector of

driving conditions \mathbf{D} (e.g. rainfall), a vector of initial system states \mathbf{X}_0 (e.g. landscape elevation), and vectors (with length t , the length of the simulation) representing future system states \mathbf{X} , and outputs \mathbf{Y} (e.g. catchment runoff/ sediment flux):

$$\mathbf{Y}, \mathbf{X} = M(\theta, \mathbf{X}_0, \mathbf{D}) \quad (1)$$

Geomorphic models are generally concerned with the action of a number of processes which locally transport mass (e.g. sediment, including organics and nutrients) that lead to changes in landscape form (\mathbf{X}) at a specific point over time (t). Though models generally differ in the processes evoked to move sediment, a fundamental approach governing most geomorphic models is to divide the landscape into units called control volumes (in 1, 2 or 3 dimensions). In most models a quasi-2D conservation of mass is applied by calculating the change in elevation $d\eta$ in response to sediment flux into and out of a control, where dx indicates the size of the control in one dimension. The control may be divided into three stores, and M partly specified by (Figure 2; Tucker and Hancock, 2010; Wainwright et al., 2008a):

$$\frac{\partial H_r}{\partial t} = T_u - S_c \quad (2)$$

$$\frac{\partial H_s}{\partial t} = d - \varepsilon + S_c \quad (3)$$

$$\frac{\partial H_t}{\partial t} = - \frac{\partial q_s}{\partial x} + \varepsilon - d \quad (4)$$

where H_r is depth of bedrock (m), T_u is tectonic uplift (m); S_c (ms^{-1}) represents the rate of conversion of rock to soil/surface regolith, H_s (m); d is sediment deposition rate (ms^{-1}) and ε is sediment entrainment rate (ms^{-1}) from and into the equivalent depth of sediment in transport H_t (m), and q_s (m^2s^{-1}) is sediment discharge across the surface, dx . Up to specifying the source terms, accounting for density/particle size differences between the stores, and developing an appropriate numerical solution, Equations 2-4 can generally be used to simulate the evolution of any point in the landscape, though some specific exceptions apply (Tucker and Hancock, 2010).

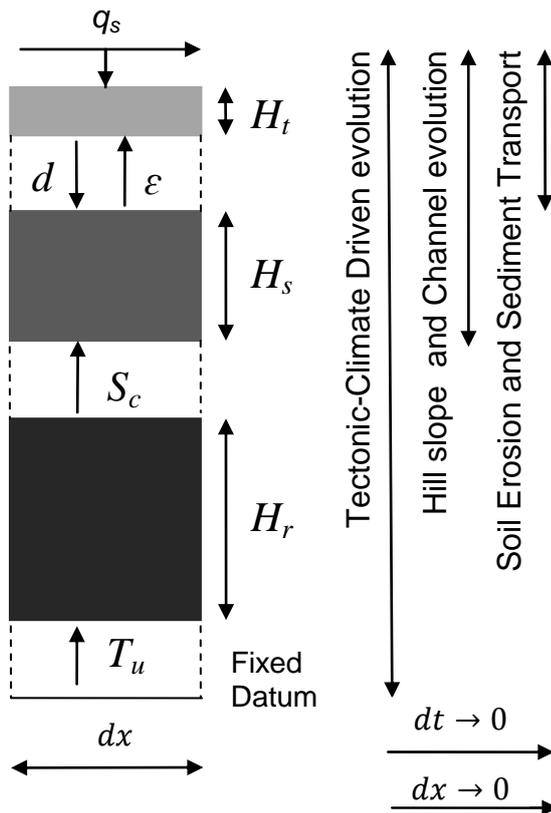


Figure 2. Key fluxes governing sediment movement within a one-dimensional control volume. The temporal scale of model application determines the relevant equations and methods of parameterisation.

To completely specify M, first a conceptual understanding of the relevant source terms on the right hand side of equations 2-4 and the processes that control them are required at the scale of enquiry. Second, the conceptual model needs to be codified into a set of equations, and an appropriate analytical/numerical solution sought. As with all geomorphic enquiry, an understanding of scale underpins the specification of these source terms, and the answers to the questions posed in the introduction.

Process dominance and scale

The processes that govern changes in H_r operate over larger timescales (e.g. are relatively slower) than the processes governing changes in H_s which in turn are slower than processes governing changes in H_t . Therefore as the timescale over which a model needs to be applied tends to zero, so the number of relevant stores and processes that control $d\eta$ also reduce (Figure 2).

In catchment systems over millennia climate-induced fluctuations in sediment transport and rock breakdown, alongside tectonic uplift, govern the evolution of plate tectonics. Therefore, changes in H_r are important and need to be included in the conceptual model of landscape change (Bishop, 2007). Over decadal and centennial timescales, catchments predominantly evolve in response to climate-induced fluctuations in sediment transport, and therefore the stores of sediment in the landscape evolve through transport between control volumes, and H_r may be considered fixed and the processes that control them (T_u and S_c) relaxed. At this scale the state vector (\mathbf{X}) needs to consider not only the evolution of sediment and mass, but also controlling and interacting factors such as vegetation (Istanbulluoglu and Bras, 2005; Reinhardt et al., 2010) and potentially anthropogenic influence (Wainwright and Millington, 2010). Alongside exerting control on sediment flux directly, these factors will also respond independently to climatic changes, creating potentially complex, and non-linear landscape feedbacks (Corenblit and Steiger, 2009). At even smaller temporal scales, sediment flux is controlled by current weather conditions and the effect of previous events operating at all scales (Schumm and Lichty, 1965), which manifest their effects through the model initial conditions. Therefore many system states are fixed (e.g. H_r and vegetation cover) and only need to be specified in the initial conditions (\mathbf{X}_0), with no additional equations required to simulate their evolution.

Similarly, over different spatial scales different processes will become important in controlling landscape behaviour. At small spatial scales on hill slopes instantaneous fluxes of water (raindrops and overland flow) control grain scale movements of sediment (Brazier et al., 2011). As catchment size increases overland flow concentrates to form rills, gullies and channels which, alongside mass movements, are increasingly important in controlling catchment sediment flux (Nichols, 2006). Therefore as the spatial and temporal scales of interest reduce, the range of processes that must be considered also reduces, and other, larger scale processes are manifest through the model boundary conditions (\mathbf{X}_0).

Developing a model to address a specific problem therefore requires a sound

conceptual understanding of the time and space scales over which processes operate to control the landforms in question. The extent to which a process is represented in a model, however, depends on its mathematical formulation.

Process representation

Experimental work has been conducted, both in the laboratory and in the field, to investigate surface process, such as rain-splash and overland flow driven erosion (Furbish et al., 2009; Wainwright et al., 2000). Alongside fundamental, physical equations (e.g. Navier-Stokes equations of fluid motion), many such studies have provided us with predictive equations from which process-form interactions can be simulated (Tucker and Hancock, 2010; Wilcock and Crowe, 2003). When combined with conservation of mass equations and integrated over time, such models represent the fundamental mechanisms by which climatic fluctuations manifest in catchment-scale landscape evolution.

Given the temporal and spatial constraints on observation, the majority of experimental work has attempted to parameterise the processes in equation 4, through what may be determined process-based models (Wainwright et al., 2008a; Wilcock and Crowe, 2003). However, even at small spatial and temporal (reductionist/observational) scales, different process parameterisations have been developed.

In catchment systems erosion, transport and deposition of sediment by water is controlled by both transport limited processes (T_L ; e.g. presence and power/stress imparted by water at the surface) and supply/detachment limited processes (S_L ; e.g. the resistive forces at the sediment bed that impede sediment movement). The most widely applied predictive equations have calculated a sediment transport rate (q_s) as a function of T_L and/or S_L , which implicitly assume sediment transport is in equilibrium (and potentially at some capacity), and evolve H_s according to sediment flux into and out of the cell support (Wainwright et al., 2008a; Wilcock and Crowe, 2003). Though such methods may provide useful predictions, changes in sediment transport and soil depth are inherently in disequilibrium. Alternative parameterisations have been developed that

explicitly calculate an entrainment and deposition flux into, and out of transport H_t (Figure 3; Hairsine and Rose, 1992; Wainwright et al., 2008a). Developing better parameterisations of equation 4 have been limited by the difficulty of measuring sediment in transport. Therefore, even at scales where monitoring can take place to parameterise system processes, competing process representations may be derived reflecting process uncertainty, and also different experimental setups (Wainwright et al., 2000).

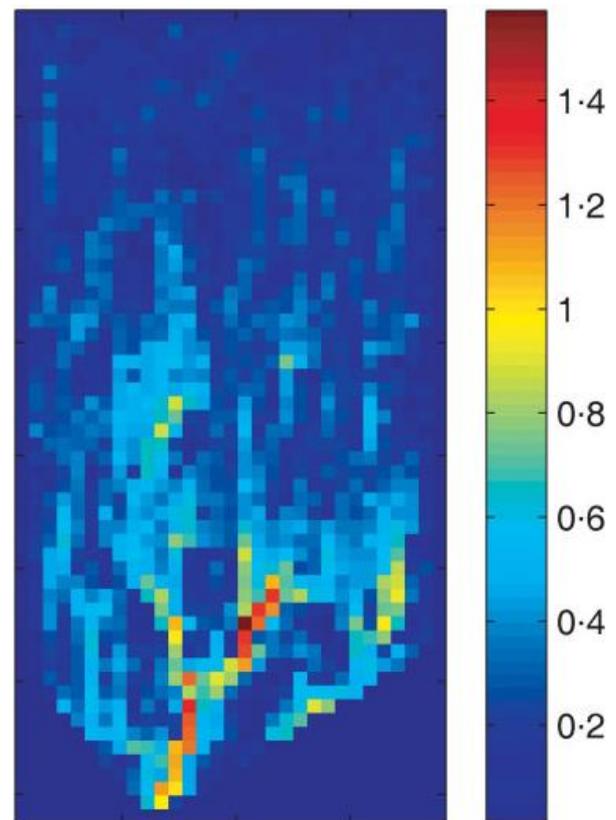


Figure 3. Spatial pattern of at-a-point total sediment movement (kg) predicted by MAHLERAN soil erosion model when applied to an 18x35m runoff plot (Wainwright et al.2008b).

Specifying the source terms in equation 2-4 may also require separate models that simulate the behaviour of phenomena that control landscape change. In catchment systems transport limiting factors controlling sediment transport are primarily derived from free surface flow. The fundamental equations that simulate free surface flow are the 3D Navier-Stokes equations (Lane, 1998). Although it has been argued that flow

sediment interactions should be modelled using these equations (Hardy, 2008), various simplifications are typically made because of computational limitations involved in deriving accurate numerical solutions – a class of models often termed “reduced complexity” models.

In the case of hydraulic/hydrologic modelling, such simplifications include depth averaging to two dimensions (Lane et al., 1999) and the fully dynamic 1D St Venant equations with further simplifications thereof that neglect potentially unimportant terms to derive the diffusive- and kinematic-wave models (Hunter et al., 2007; Tucker and Hancock, 2010). Additional flow simplifications have led to a number of cellular approaches for flow routing, which employ simplified rules to route flow in river channels and hillslopes (Favis-Mortlock, 1998; Nicholas, 2009; Thomas and Nicholas, 2002). Furthermore, some approaches have sought to employ simpler cellular and physically based rules to simulate morphological change, and have simulated sand dune formation (Figure 4; Nield and Baas, 2008), meander migration (Coulthard and Van De Wiel, 2006), braided river evolution (Thomas et al., 2007), and floodplain evolution (Figure 5; Karssenberg and Bridge, 2008).

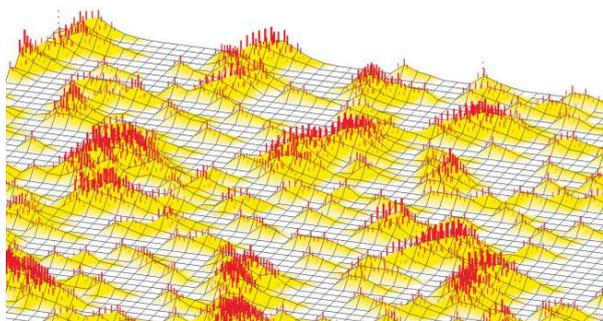


Figure 4. Nebkha dunes and vegetation (dark bars) simulated with a cellular approach (Modified from Nield and Baas, 2008).

The specific conditions under which such equations are ‘valid’ should be carefully considered (see Lane (1998) and Cao and Carling (2002) for a consideration of hydraulic equations in a geomorphic context). The scale and method by which transport limited conditions are modelled may be more important than the equations that link properties of flow to actual sediment

movement, given the potential for non-linear error propagation.

The increased availability of high resolution topographic data has facilitated the application of small scale model parameterisations over increasingly large domains. For computational reasons and availability of other distributed data, however, it is often necessary to model at coarser resolutions in both space and time. One approach to deal with the problem of process parameterisation at coarser scales is simply to apply the same equations developed at smaller scales. However, a key problem with this approach is that many if not all geomorphic laws are scale dependent. For example many geomorphic laws are slope dependent, thus increasing model cell size reduces slope, and can lead to inaccurate predictions of erosion (Kalin et al., 2003). Such inaccuracies occur because of how changing scale of resolution affects both geomorphic laws and the laws governing flow (Brazier et al., 2011). Inferring the validity of a model equation independently of the grid within which it is applied may be difficult (Nicholas, 2005). Thus in model development changing model scale will affect the validity of all processes included in the model.

Although all spatial parameterisation is lumped to some degree, critical scales in the landscape that govern larger scale behaviour (e.g. the scale of interest) should be considered. For example, on hillslopes and in channels coarser scale models will fail to account for the spatial heterogeneity of flow. Given the relationship between sediment transport and flow is strong and non-linear, neglecting this heterogeneity will lead to an under-prediction of erosion and sediment transport (Ferguson, 2003; Nicholas, 2000). Relying on equilibrium concepts to model sub-grid scale channel features is a popular approach to deal with this problem in landscape evolution models, but relies on equilibrium concepts to model potentially non-equilibrium behaviour (Nicholas and Quine, 2007a; Tucker and Hancock, 2010). Therefore the specific scale at which a model equation is a valid representation of sub grid-scale processes is an important consideration when developing a numerical model.

At larger temporal scales of enquiry models have to deal with the disparity between the timescales of individual events (e.g. rainfall runoff) and the evolution of landscapes

(Tucker and Hancock, 2010). As a result of computational limitations, many landscape evolution models applied over larger temporal scales relate at-a-point discharge to upslope contributing area, which implies runoff is in equilibrium with uniform rainfall. Such approximations subsequently used to simulate sediment transport fail to account for spatial and temporal variability. Sub-scale events may be important in controlling runoff production and therefore the sediment transport that actually governs landscape behaviour. Simple averaging to an effective event may miss that some events are more important in controlling erosion and sediment transport than others (Nichols, 2006). Furthermore, modification of the landscape by continued operation of smaller events may be (more) important in controlling morphological change (Goodrich et al., 2008; Sambrook-Smith et al., 2010).

The issue of what scales to consider and therefore what processes to resolve explicitly in a given model structure points towards a fundamental issue for geomorphologists: given natural systems often display non-linear, threshold responses, it is uncertain - and debated in the literature - to what extent small scale processes (in both space and time) control larger scale system behaviour (Lane and Richards, 1997). It is therefore

uncertain to what extent fine scale processes need to be resolved explicitly in geomorphic models, or whether simpler treatments, relying on for example regime theory, are applicable (Nicholas and Quine, 2007). Choosing a specific process representation therefore reflects a specific modelling hypothesis regarding the relevant processes governing a different problem. Multiple representations, and therefore hypotheses of the same processes may require investigation (Krueger et al., 2009). To help overcome this issue, data are required to constrain model parameters and evaluate model hypothesis (Kleinhans et al., 2012).

Data and model evaluation

Data availability is an essential factor governing model development, as data provides the modeller with the ability to constrain model parameters and evaluate the quality of model predictions. The evaluation of model process representation is an essential step as it often occurs prior to application of models to investigate so called "what if" questions (Michaelides and Wilson, 2007; Nicholas and Quine, 2010). Such model application is often at space and time scales over which data are insufficient to differentiate between competing model

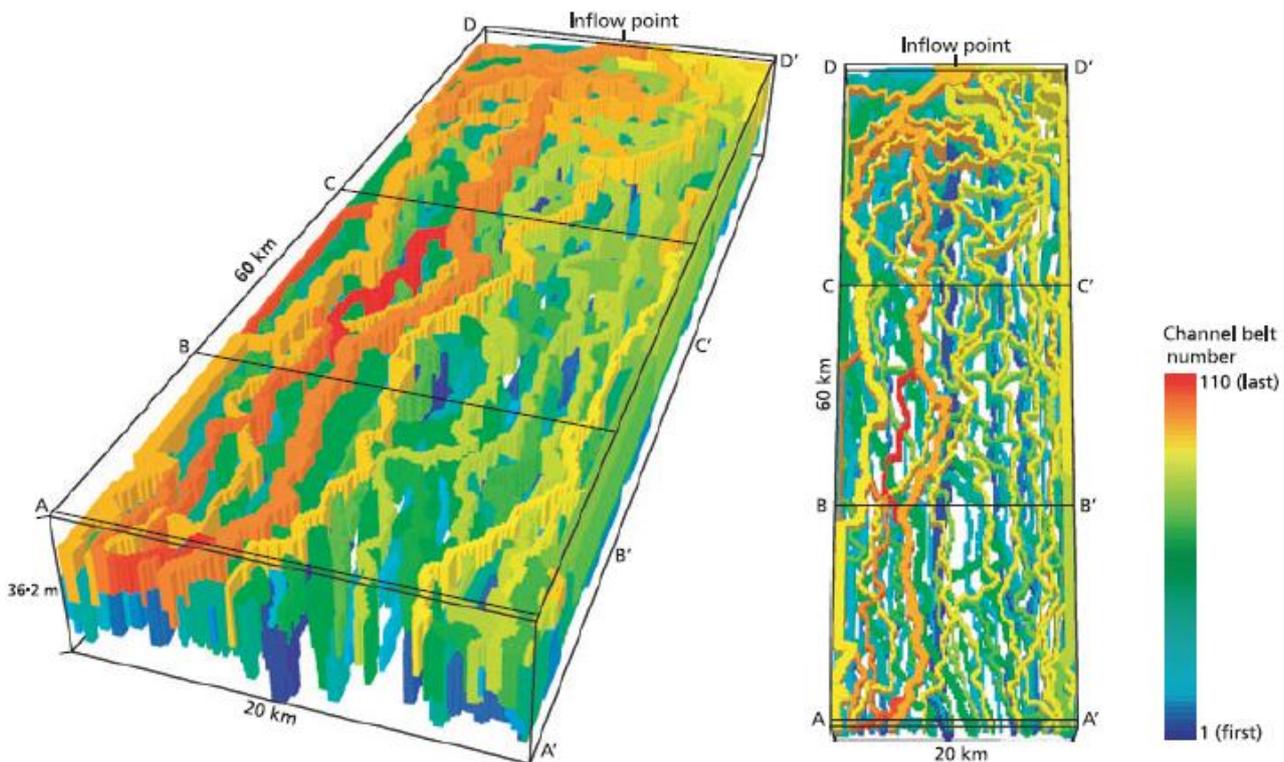


Figure 5. Simulated channel belt and floodplain evolution (e.g. by bifurcation, avulsion and aggradation) in response to base level rise (Karssenbergh and Bridge, 2008).

hypotheses (Pelletier, 2011).

In the ideal case, all process parameterisations of equations 2-4 will be known and data will be available at the scale of the model cell size to constrain model parameters and the initial model conditions. For example in soil erosion modelling distributed information on particle size may be required to parameterise both roughness for overland flow modelling and the supply limiting factors controlling sediment entrainment (Wainwright et al., 2008b). However, often such data are unavailable, or inconsistent with the scale of model application (Brazier et al., 2011).

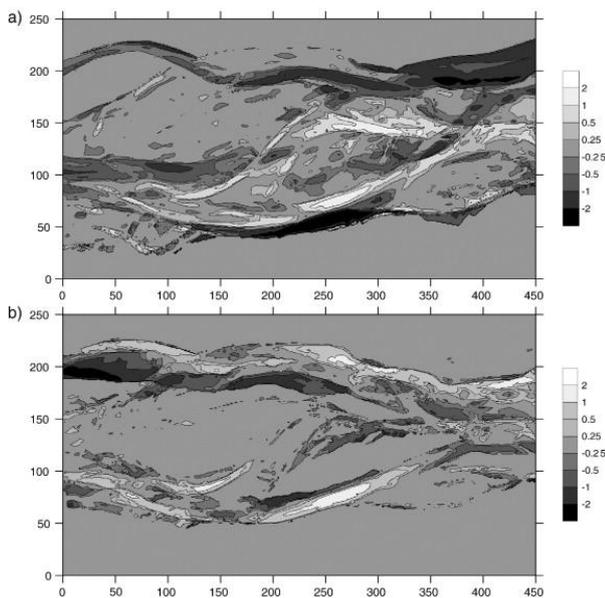


Figure 6. Comparison of measured (a) and modelled (b) elevation change in the braided Avoca River, New Zealand (Nicholas and Quine, 2007a).

As a consequence, models may be calibrated by comparing model outputs to observations. In the case of catchment modelling, models are typically calibrated by adjusting internal model parameters to derive the best fit between model outputs (e.g. sediment/water flux at a catchment outlet) and the equivalent observations at specific locations (Canfield and Goodrich, 2006; Nearing, 2000). In morphological modelling, distributed model predictions may be compared to: observations or morphological change (Figure 6; Nicholas and Quine, 2007a); to results derived from models with a stronger physical basis for prediction (Nicholas, 2009);

and also to results derived from physical model experiments (Nicholas et al., 2009).

A number of implicit assumptions made in model calibration may potentially undermine model application; First, it is often assumed that parameter uncertainty is the only form of modelling uncertainty; Second, that the model is equal to reality; Third, the initial states are the true initial states (e.g. the DEM is error-free); and Fourth, that the input/driving conditions and output data to evaluate model performance are true. In most, if not all situations these conditions do not hold due to the problems of modelling an open system, where true model validation and verification is impossible (Oreskes et al., 1994).

As a result of these assumptions incorrect model parameters can be identified that reproduce catchment outlet conditions (i.e. larger scale measurements) with insufficient consideration of how well they reproduce the internal spatial patterns of process-form interaction that ultimately control larger (and longer) scale response. Furthermore, model parameters may be identified that are highly unique to specific settings (Nearing, 1999; Nearing, 2000), and inapplicable elsewhere because of the non-linear open nature of natural systems. Worse still, a number of parameter combinations within a specific model may provide equally good predictions - a form of model equifinality (Brazier et al., 2000). Similarly, another form of model equifinality may occur if the data are insufficient to differentiate between competing models. Equifinality has arisen at a range of scales, from using metrics of landscape form to differentiate between transport and supply limited models of landscape evolution (Pelletier, 2011), evaluating alluvial fan evolution (Nicholas and Quine, 2010), and at smaller scales when applying models with complex, and ill constrained parameters (Brazier et al., 2000). However, equifinality is not all bad if it avoids over confidence in the information content of data and therefore the potential rejection of good model structures. In the face of equifinality, simpler models may be preferred that are justified by the data: A model is only as good as the data available to constrain model structure, parameters, and therefore predictions.

Accounting for uncertainty when developing numerical models

The preceding sections have discussed the main issues to consider in developing and applying a numerical model to address a geomorphic problem. Even if we consider our conceptual model of the system and the dominant processes to be accurate, computational resources and data availability will limit our ability to apply the preferred model at the desired scale. Furthermore, uncertainty surrounding the relationship between process dominance and scale is a fundamental geomorphic question governing model development. Therefore there is no single answer to the three questions posed in the introduction, and nor can each be answered independently. As a result of these factors a model can and will only remain as a (working) hypotheses of how processes and landforms interact. In order to develop, evaluate and use models in geomorphic enquiry, uncertainty in both data and models needs to be dealt with in a robust manner.

Considering numerical modelling from a Bayesian perspective provides a suitable framework for robust model development of non-linear open systems. As a result of modelling uncertainties we should be interested in obtaining the probability $P(\cdot)$ that the model (M) and its parameters (θ) are a correct representation of reality, given the available data (Y), initial model conditions (X_0), and driving conditions (D). In order to do this we combine our prior beliefs about the model structure, $P(M)$ and associated parameters, $P(\theta|M)$ (which are dependent on the specific model structure) with some data using a Likelihood Function (e.g. a measure of model performance based on a given dataset) $P(Y|\theta, M, X_0, D)$, to obtain our posterior belief from Bayes' equation (Draper, 1995):

$$P(M, \theta|Y, X_0, D) \propto P(Y|\theta, M, X_0, D)P(\theta|M)P(M) \quad (5)$$

Thus, our confidence in the model is specifically dependent on the data to constrain initial conditions, driving conditions and that used to evaluate model performance.

In many applications only a single model structure is considered, and associated parameters are either derived from previous studies or are optimised to the specific data available. As a result the final two terms in Equation 5 collapse to a single set of structural assumptions. In such cases overconfidence in the data for the reasons discussed above – both in its accuracy and its general applicability to a wide range of settings – may lead to inappropriate model rejection and narrowing of the posterior probability of all possible models. This may lead to an entrenchment of modelling concepts that may prevent wider exploration of $P(M)$ and therefore potentially more appropriate models for particular circumstances (Nicholas and Quine, 2007a; Wainwright et al., 2008a).

A better position to develop models and therefore understanding of process-form interactions is to consider the uncertainty in different model parameters, and therefore consider a range of possible combinations of parameters $P(\theta|M)$ that may reproduce the data, according to a likelihood function that considers potential uncertainty in the data. Significant advances have been made in developing appropriate likelihood functions in the related discipline of hydrology that consider different forms of model uncertainty (Beven, 2006; Schoups and Vrugt, 2010). Such statistical treatments of model uncertainty require further adaptation to geomorphic problems, including potential uncertainty in model boundary conditions and elevation data (Hutton and Brazier, 2012; Nicholas and Quine, 2010; Wheaton et al., 2010). Calibration and the related sensitivity analysis conducted by exploring adequately different parameter combinations that constitute $P(\theta|M)$ can guide the modeller as to which parameters are most important in controlling system response (Hutton et al., 2012; Saltelli, 1999). Such information can be then used to guide further data collection targeted at constraining the most important parameters.

Furthermore, when different models $P(M)$ and therefore different hypotheses of form-process interactions are confronted with the same observations (Hancock et al., 2011; Krueger et al., 2009; Pelletier, 2011; Tatard et al., 2008) the strengths and weaknesses of different model structures may be identified. Such information can be used to guide further

data collection and understand the conditions (e.g. $\mathbf{Y}, \mathbf{X}_0, \mathbf{D}$) under which different process parameterisations (e.g. M, θ) are valid.

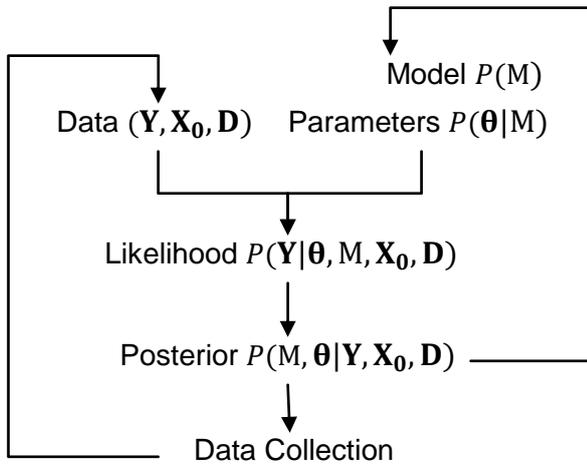


Figure 7. Schematic illustration of the iterative use of a Bayesian Approach for model development

Therefore posterior understanding derived from comparison to data can guide further data collection (Figure 7). Depending on the similarities between the conditions used to derive the posterior (e.g. $\mathbf{Y}, \mathbf{X}_0, \mathbf{D}$) and the newly collected data, the posterior in equation 5 (e.g. the left hand side) can then become the prior (e.g. move to the right hand side of Equation 5) for comparison to the newly collected data when combined with a likelihood function. When using model structures and parameters derived from previous applications it is up to the modeller to consider how appropriate such models are to the situation under current consideration, and whether such model application is supported by available data.

Bayes' equation therefore provides both a conceptual and probabilistic framework for model development that can explicitly consider uncertainty in models and data, and therefore appropriately frame the use of models in developing understanding of process-form interactions. Process-form understanding is developed iteratively through continuous dialogue between models and data. Furthermore such development provides a more robust grounding for subsequent model application to investigate "what if" type questions by considering a range of possible model structures supported by available data, which will prevent

overconfidence in the results of a single model prediction.

Conclusion

Numerical Models have, especially over recent years, become a central tool in geomorphic enquiry, and have allowed exploration of a range of system dynamics at a range of spatial and temporal scales. Appropriate use of numerical models should consider the scale of model application, the potential processes controlling landscape form at the scale of application, computational resources, data availability, and the validity of modelling concepts derived from previous modelling applications.

Given the many uncertainties governing model application, not least uncertainty regarding fundamental issues concerning process dominance and scale, an approach to model development considering such uncertainties from a Bayesian perspective is recommended. Such an approach provides a robust framework for model development, model rejection and therefore hypothesis testing that considers uncertainty in both data and models. Advances in data collection and specification of errors in available data will facilitate robust model development.

References

- Beven K. 2002. Towards a coherent philosophy for modelling the environment *Proceedings of the Royal Society of London A* **458**: (2026): 2465-2484.
- Beven K. 2006. A manifesto for the equifinality thesis *Journal of Hydrology* **320**: 18-36.
- Bishop P. 2007. Long-term landscape evolution: linking tectonics and surface processes *Earth Surface Processes And Landforms* **32**(3): 329-365.
- Bloschl G, Sivapalan M. 1995. Scale Issues In Hydrological Modeling - A Review. *Hydrological Processes* **9**(3-4): 251-290.
- Brazier RE, Beven KJ, Freer J, Rowan JS. 2000. Equifinality and uncertainty in physically based soil erosion models: Application of the glue methodology to WEPP-the water erosion prediction project-for sites in the UK and USA. *Earth Surface Processes And Landforms* **25**(8): 825-845.

- Brazier RE, Hutton CJ, Parsons AJ, Wainwright J. 2011. Scaling Soil Erosion Models in Space and Time. In *Handbook of Erosion Modelling*, RPC Morgan RPC and Nearing MA (eds). Blackwell: 98-117.
- Canfield HE, Goodrich DC. 2006. The impact of parameter lumping and geometric simplification in modelling runoff and erosion in the shrublands of southeast Arizona. *Hydrological Processes* **20**(1): 17-35.
- Cao Z, Carling PA. 2002. Mathematical modelling of alluvial rivers: reality and myth Part I: General review. *Proceedings of the Institution of Civil Engineers-Water and Maritime Engineering* **154**(3): 207-219.
- Church M. 1996. Space time and the mountain - how do we order what we see? In *The scientific nature of geomorphology*, Rhoads BL and Thorn CE (eds). Wiley: Chichester; 147-170.
- Corenblit D, and Steiger J. 2009. Vegetation as a major conductor of geomorphic changes on the Earth surface: toward evolutionary geomorphology. *Earth Surface Processes And Landforms* **34**(6): 891-896.
- Coulthard TJ. 2001. Landscape evolution models: a software review *Hydrological Processes* **15**: 165-173.
- Coulthard TJ, Lewin J, and Macklin MG. 2005. Modelling differential catchment response to environmental change *Geomorphology* **69**: 222-241.
- Coulthard TJ, Van De Wiel MJ. 2006. A cellular model of river meandering. *Earth Surface Processes And Landforms* **31**(1): 123-132.
- Draper D. 1995. Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society Series B-Methodological* **57**(1): 45-97.
- Favis-Mortlock D. 1998. A self-organizing dynamic systems approach to the simulation of rill initiation and development on hillslopes. *Computers & Geosciences* **24**(4): 353-372.
- Ferguson RI. 2003. The missing dimension: effects of lateral variation on 1-D calculations of fluvial bedload transport. *Geomorphology* **56**: 1-14.
- Furbish DJ, Childs EM, Haff PK, Schmeckle MW. 2009. Rain splash of soil grains as a stochastic advection-dispersion process with implications for desert plant-soil interactions and land-surface evolution. *Journal Of Geophysical Research-Earth Surface* **114**: F00A03. doi:10.1029/2009JF001265.
- Goodrich DC, Unkrich CL, Keefer TO, Nichols MH, Stone JJ, Levick LR, Scott RL. 2008. Event to multidecadal persistence in rainfall and runoff in southeast Arizona. *Water Resources Research* **44**: W05S14. doi:10.1029/2007WR006222.
- Hairsine PB, Rose CW. 1992. Modeling Soil Erosion due to overland flow using physical principles: 2 Sheet Flow. *Water Resource Research* **28**: 237-243.
- Hancock GR, Coulthard TJ, Martinez C, Kalma JD. 2011. An evaluation of landscape evolution models to simulate decadal and centennial scale soil erosion in grassland catchments. *Journal Of Hydrology* **398**(3-4): 171-183.
- Hardy RJ. 2005. Modelling granular sediment transport over water-worked gravels. *Earth Surface Processes And Landforms* **30**: 1069-1076.
- Hardy RJ. 2008. Geomorphology Fluid flow Modelling: Can Fluvial Flow Only Be Modelled Using a Three-Dimensional Approach? *Geography Compass* **2**(1): 215-234.
- Hunter NM, Bates PD, Horritt MS, Wilson MD. 2007. Simple spatially-distributed models for predicting flood inundation: A review. *Geomorphology* **90**(3-4): 208-225.
- Hutton C, Brazier R. 2012. Quantifying riparian zone structure from airborne LiDAR: Vegetation filtering anisotropic interpolation and uncertainty propagation. *Journal Of Hydrology* **442**: 36-45
- Hutton C, Brazier RE, Nicholas AP, Nearing MA. 2012. On the Effects of Improved Cross-section Representation in One-dimensional Flow Routing Models Applied to Ephemeral Rivers. *Water Resources Research* **48**: W04509. doi:10.1029/2011WR011298.
- Istanbulluoglu E, Bras RL. 2005. Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes drainage density and topography *Journal Of Geophysical Research-Earth Surface* **110**: F02012. doi:10.1029/2004JF000249.
- Kalin L, Govindaraju RS, Hantush MM. 2003. Effect of geomorphic resolution on modeling of runoff hydrograph and sedimentograph

- over small watersheds. *Journal Of Hydrology* **276**: 89-111.
- Karszenberg D, Bridge, JS. 2008. A three-dimensional numerical model of sediment transport, erosion and deposition within a network of channel belts, floodplain and hill slope: extrinsic and intrinsic controls on floodplain dynamics and alluvial architecture. *Sedimentology* **55**(6): 1717-1745.
- Kleinhans MG, de Haas T, Lavooi E, Makaske B. 2012. Evaluating competing hypotheses for the origin and dynamics of river anastomosis *Earth Surface Processes And Landforms* doi: 10.1002/esp3282.
- Krueger T, Quniton JN, Freer J, Macleod CJ, Bilotta GS, Brazier RE, Butler P, Haygarth PM. 2009. Uncertainties in Data and Models to Describe Dynamics of Agricultural Sediment and Phosphorus Transfer. *Journal of Environmental Quality* **38**: 1137-1148.
- Lane SN. 1998. Hydraulic modelling in hydrology and geomorphology: A review of high resolution approaches *Hydrological Processes* **12**(8): 1131-1150.
- Lane SN, Bradbrook KF, Richards KS, Biron PA, Roy AG. 1999. The application of computational fluid dynamics to natural river channels: three-dimensional versus two-dimensional approaches. *Geomorphology* **29**(1-2): 1-20.
- Lane SN, Richards KS. 1997. Linking river channel form and process: Time space and causality revisited. *Earth Surface Processes And Landforms* **22**(3): 249-260.
- Livingstone I, Wiggs GFS, Weaver CM. 2007. Geomorphology of desert sand dunes: A review of recent progress *Earth-Science Reviews* **80**(3-4): 239-257.
- Merritt WS, Letcher RA, Jakeman AJ. 2003. A review of erosion and sediment transport models. *Environmental Modelling and Software* **18**: 761-799
- Michaelides K, Wilson MD. 2007. Uncertainty in predicted runoff due to patterns of spatially variable infiltration. *Water Resources Research* **43**(2): W02415. doi:10.1029/2006WR005039.
- Morgan RPC, Nearing MA (eds) 2010 *Handbook of Erosion Modelling*. Wiley-Blackwell 416 pp.
- Nearing MA. 1999. Variability in soil erosion data from replicated plots. *Soil science society of America Journal* **63**(6): 1829-1835.
- Nearing MA. 2000. Evaluating soil erosion models using measured plot data: Accounting for variability in the data. *Earth Surface Processes And Landforms* **25**(9): 1035-1043.
- Nicholas AP. 2000. Modelling bedload yield in braided gravel bed rivers. *Geomorphology* **36**: 89-106.
- Nicholas AP. 2005. Cellular Modelling in Fluvial Geomorphology *Earth Surface Processes And Landforms* **30**: 645-649
- Nicholas AP. 2009. Reduced-complexity flow routing models for sinuous single-thread channels: intercomparison with a physically-based shallow-water equation model. *Earth Surface Processes And Landforms* **34**: 641-653
- Nicholas AP, Clarke L, Quine TA. 2009. A numerical modelling and experimental study of flow width dynamics on alluvial fans. *Earth Surface Processes And Landforms* **34**(15): 1985-1993
- Nicholas AP and Quine TA 2007a Crossing the divide: Representation of channels and processes in reduced-complexity river models at reach and landscape scales. *Geomorphology* **90**(3-4): 318-339
- Nicholas AP, Quine TA. 2007b. Modeling alluvial landform change in the absence of external environmental forcing *Geology* **35**(6): 527-530
- Nicholas AP, Quine TA. 2010. Quantitative assessment of landform equifinality and palaeoenvironmental reconstruction using geomorphic models. *Geomorphology* **121**(3-4): 167-183.
- Nichols MH. 2006. Measured sediment yield rates from semiarid rangeland watersheds *Rangeland Ecology & Management* **59**(1): 55-62
- Nield JM, Baas ACW. 2008. Investigating parabolic and nebkha dune formation using cellular automation modelling approach. *Earth Surface Processes And Landforms* **33**(724-740).
- Oreskes N, Shraderfrechette K, Belitz K. 1994. Verification Validation And Confirmation Of Numerical-Models In The Earth-Sciences. *Science* **263**(5147): 641-646

- Pelletier JD. 2011. Fluvial and slope-wash erosion of soil-mantled landscapes: detachment- or transport-limited? *Earth Surface Processes And Landforms* **37**(1): 37-51.
- Refsgaard JC, van der Sluijs JP, Brown J, van der Keur P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* **29**: 1586-15977.
- Reinhardt L, Jerolmack D, Cardinale BJ, Vanacker V, Wright J. 2010. Dynamic interactions of life and its landscape: feedbacks at the interface of geomorphology and ecology. *Earth Surface Processes And Landforms* **35**(1): 78-101.
- Saltelli A. 1999. The role of sensitivity analysis Transparency relevance robustness and parsimony in modelling *Safety and Reliability* Vols 1 & 2: 1181-1186.
- Sambrook-Smith GHS, Best JL, Ashworth PJ, Lane SN, Parker NO, Lunt IA, Thomas RE, Simpson CJ. 2010. Can we distinguish flood frequency and magnitude in the sedimentological record of rivers? *Geology* **38**(7): 579-582.
- Schmeeckle MW, Nelson JM. 2003. Direct numerical simulation of bedload transport using a local dynamic boundary condition. *Sedimentology* **50**: 279-301.
- Schoups G, Vrugt JA. 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated heteroscedastic and non-Gaussian errors. *Water Resources Research* **46**: W10531. doi:10.1029/2009WR008933.
- Schumm SA, Lichty RW. 1965. Time space and causality in geomorphology *American Journal of Science* **263**: 110-119.
- Tatard L, Planchon O, Wainwright J, Nord G, Favis-Mortlock D, Silvera N, Ribolzi O, Esteves M, Hua Huang C. 2008. Measurement and modelling of high-resolution flow-velocity data under simulated rainfall on a low-slope sandy soil *Journal Of Hydrology* **348**(1-2): 1-12
- Thomas R, Nicholas AP. 2002. Simulation of braided river flow using a new cellular routing scheme. *Geomorphology* **43**(3-4): 179-195.
- Thomas R, Nicholas AP, Quine TA. 2007. Cellular modelling as a tool for interpreting historic braided river evolution *Geomorphology* **90**: 302-217.
- Tucker GE, Hancock G. 2010. Modelling landscape evolution. *Earth Surface Processes and Landforms* **35**: 28-50
- Van de Wiel MJ, Coulthard TJ, Macklin MG, Lewin J. 2011. Modelling the response of river systems to environmental change: Progress problems and prospects for palaeo-environmental reconstructions. *Earth-Science Reviews* **104**(1-3): 167-185.
- Wainwright J, Millington JDA. 2010. Mind, the gap in landscape-evolution modelling. *Earth Surface Processes And Landforms* **35**(7): 842-855.
- Wainwright J, Mulligan M (eds). 2004. *Environmental Modelling: Finding Simplicity in Complexity*. John Wiley and Sons Chichester 430 pp.
- Wainwright J, Parsons AJ. 2002. The effect of temporal variations in rainfall on scale dependency in runoff coefficients. *Water Resources Research* **38**(12): 1271, doi:10.1029/2000WR000188.
- Wainwright J, Parsons AJ, Abrahams AD. 2000. Plot-scale studies of vegetation overland flow and erosion interactions: case studies from Arizona and New Mexico. *Hydrological Processes* **14**(16-17): 2921-2943.
- Wainwright J, Parsons AJ, Muller EN, Brazier RE, Powell DM, Fenti, B. 2008a. A transport-distance approach to scaling erosion rates: I Background and model development. *Earth Surface Processes And Landforms* **33**(5): 813-826.
- Wainwright J, Parsons AJ, Muller EN, Brazier RE, Powell DM, Fenti, B. 2008b. A transport-distance approach to scaling erosion rates: 2 Sensitivity and evaluation of MAHLERAN. *Earth Surface Processes And Landforms* **33**(6): 962-984.
- Wheaton JM, Brasington J, Darby SE, Sear DA. 2010. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. *Earth Surface Processes And Landforms* **35**(2): 136-156.
- Wilcock P, Crowe JC. 2003. Surface-based transport model for mixed sized sediment. *Journal of Hydraulic Engineering* **129**(2): 120-128.

Wilcock PR, Iverson RM (eds). 2003. *Prediction in Geomorphology*. Geophysical Monograph Series Vol 135, Washington DC 256 doi:101029/GM135 pp

Willett SD, Brandon MT. 2002. On steady states in mountain belts. *Geology* **30**(2): 175-178

Modelling Geomorphic Systems: Scaled Physical Models

Daniel L. Green¹

¹Centre for Hydrological and Ecosystem Science, Department of Geography, Loughborough University (D.Green@lboro.ac.uk)



ABSTRACT: Physical models are scaled representations of a full-scale physical system which can be applied to inform our understanding of geomorphic process-form interactions. Physical and experimental modelling has been used extensively and has been proven to be of critical importance to the geomorphological user. Physical models can be loosely divided into a number of categories: 1:1 replica models; Froude-scaled models; distorted scale models; and analogue ‘similarity of process’ models. The choice of physical model type is dependent on the researcher’s aims and objectives. Advantages include the ability to: (i) isolate variables within a controlled laboratory setting; (ii) incorporate actual physical processes rather than simplifications; (iii) study infrequent or hypothetical scenarios, and; (iv) extract qualitative and quantitative data. Users of physical models must be cautious of the potential shortcomings of using a physical model, such as scale and laboratory effects. Despite these shortcomings, physical models provide a useful technique to observe, visualise and measure process-form interactions. This permits an improved understanding of complex physical relationships which other modelling methodologies may not be able to simulate.

KEYWORDS: Physical modelling, experimental methods, laboratory techniques, scale, similitude.

Introduction

Physical models are scaled representations of a physical system (Hughes, 1993). The use of physical models is well established, offering an alternative or complementary approach to what can be simulated accurately using numerical models or observed and measured through field-based investigations (Peakall *et al.*, 1996; Frostick *et al.*, 2011). Physical models have been applied to understand, assess and inform stakeholder decisions in a number of disciplines, ranging from the biological and environmental sciences to aeronautical and infrastructural engineering. Physical models provide a reputable research technique allowing the reproduction of complex physical phenomena and an understanding of process interactions to be generated in a visual and informative manner (Sutherland and Barfuss, 2011).

Physical modelling has also been used extensively within the field of geomorphology

(Peakall *et al.*, 1996) including studies of alluvial fan dynamics (e.g. Clarke *et al.*, 2010; see Figure 1), tsunami waves, jökulhaups or catastrophic dam failure inundation (e.g. Rushmer, 2007; Soares-Frazão and Zech, 2008; Rossetto *et al.*, 2011), sediment and bedform dynamics (e.g. Guy *et al.*, 1966; Allen, 1982; Southard and Boguchwal, 1990; Warburton and Davies, 1998; Madej *et al.*, 2009) and erosion plot and rill development studies (e.g. Bryan and Poesen, 1989). These studies have emphasised the importance of physical models as a method of visualising, interpreting, observing and measuring physical processes, something which is potentially problematic in a model’s full-scale counterpart (Kamphuis, 1991). This permits intrinsic factors to be separated from extrinsic factors (Clarke *et al.*, 2010), allowing the isolation of variables within a controlled laboratory environment. Consequently, physical models provide a number of advantages to the geomorphological user, which will be outlined later.

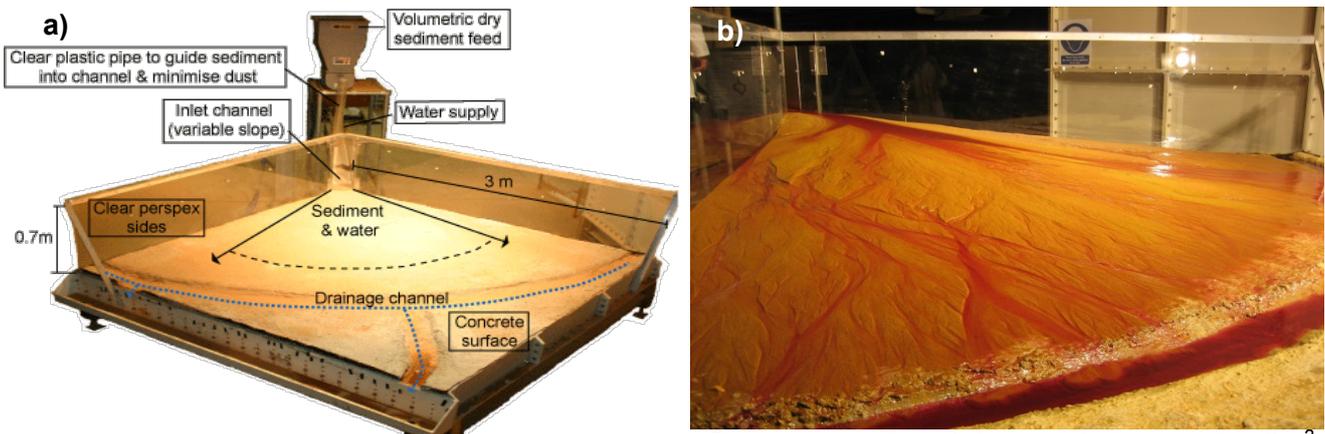


Figure 1: a) Annotated photograph of a 9m² experimental alluvial fan at the Sediment Research Facility, University of Exeter, b) experimental alluvial fan with dyed red water to assist in visualisation of surface flow paths. Source: Clarke (2013).

This paper presents: (i) a discussion on physical model typology; (ii) a brief introduction to the key physical modelling principles; (iii) an overview of the applications and importance of physical models for the geomorphological user, as well as (iv) a critical assessment of the strengths and weaknesses of using physical models in geomorphology.

Model Typology

The choice of physical model type is dependent on various factors including the project objectives and rationale, as well as cost and space limitations (Frostick *et al.*,

2011). Additionally, the purpose of the physical model will control the type of model that is used, with models generally being constructed for: (i) research purposes; (ii) communication and education purposes and/or (iii) informing decisions or providing foresight (Maynard, 2006; see Table 1).

Two types of boundary condition have been recognised within physical models: fixed-bed, where the model boundaries are non-erodible and no sediment transport can occur; and, moveable-bed, where substrate is free to move within a constrained or non-constrained channel (Hughes, 1993; Peakall *et al.*, 1996; Waldron, 2008).

Table 1: Purposes or aims of physical models

Project aims	Sub-discipline of geomorphology	Examples
Research tools to study process-form interactions	Fluvial	Influence of in-channel/floodplain vegetation on river morphology (e.g. Gran and Paola, 2010)
	Fluvial	Investigations into alluvial fan dynamics and evolution (e.g. Clarke <i>et al.</i> , 2010)
	Hillslope	Investigations into hillslope-channel coupling processes (e.g. Michaelides and Wainwright, 2002)
	Glacial	Investigations into jökulhaups with different hydrograph shapes and their subsequent impacts (e.g. Rushmer, 2007)
Education, demonstration and communication tools	Aeolian / dryland	Wind-tunnel tests on aeolian transport of different sized sand grains under varying wind velocities (e.g. Dong <i>et al.</i> , 2003)
	Fluvial	Micro-model flume to communicate channel avulsion processes to the public, students or stakeholders
Screening tools to seek alternative approaches / improve understanding	Glacial	Glacier dynamics under changing climate experiments using PVC piping valley and viscous flow medium
	Fluvial	Use of physical models to inform understanding of the downstream and upstream impacts of channel impoundment or dam removal (e.g. Einhellig <i>et al.</i> , 2010)
	Fluvial	Process understanding of ice jams at river confluences (Ettema and Muste, 2001)

Physical models can be loosely divided into a number of categories, including: (i) scaled models; (ii) Froude number scaled models; (iii) distorted scale models; (iv) analogue models; and, (v) 1:1 replica models. Despite this classification scheme, some overlap may exist, for example, an analogue model may exhibit characteristics associated with all of the other categories.

Scaled models

Scaled physical models that are built and function at reduced scale (or enlarged scale in some cases) are an important type of physical model for examining and measuring processes which are difficult to observe in reality (Michaelides and Wainwright, 2013; see Figures 2 and 3). Scaled models allow geomorphological users to overcome the inherent obstacles associated with investigating physical systems (Hughes, 1993), such as the long spatiotemporal timescales involved and problems associated with working in a naturally variable environment. Scaled physical models conform to scale ratios, shown by Eq. 1:

$$N_x = \frac{X_p}{X_m} = \frac{\text{Value of } X \text{ in Prototype}}{\text{Value of } X \text{ in Model}} \quad (\text{Eq. 1})$$

where N_x is the actual-to-model scale ratio of parameter x (which may represent width, depth, length, grain size, time, diameter etc.), and where p and m represent the actual/original system and model, respectively. Thus, if a river reach has a length of 300 m in reality but this is scaled to 3 m under modelled conditions, the model length is said to be scaled by 1:100.

Scaling occurs in all physical models to varying extents, however, modellers must be cautious when downscaling a model too much from the real world system. Maynard (2006) evaluated a 'micro-model' river system (1:14,000 horizontal, 1:1,200 vertical scale; see Figure 3) with a river channel width as small as 4 cm. It was demonstrated that using large scaling factors resulted in models becoming incomparable to the hydrodynamic processes occurring in reality. For example, a river channel width of 560 m in reality cannot be represented as 4 cm width in a physical model due to the significantly different hydrodynamic processes occurring. Additionally, when scaling particle sizes, users must be cautious of cohesive forces becoming a dominant factor in the model while being absent or negligible in reality. This may result in 'micro-model' systems

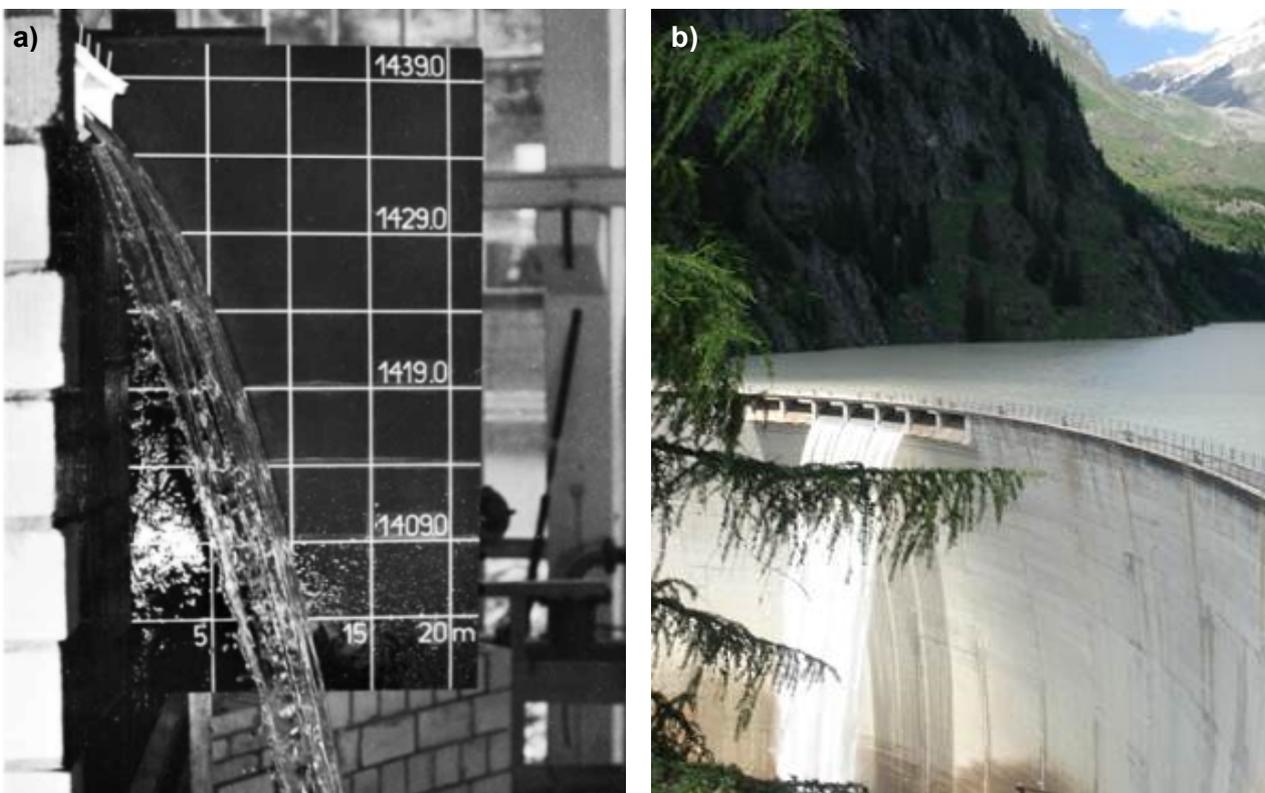


Figure 2: Overflow spillway of Gibidum Dam, Switzerland. a) 1:30 scaled physical hydraulic model, b) real-world, full scale photograph of Gibidum Dam, which the model is based on. Source: Heller (2011).

losing their predictive capabilities and becoming qualitative rather than quantitative. Whether this is disadvantageous or not depends on the researcher's aims; qualitative modelling may still be useful for demonstration, education and communication purposes, as well as a rapid, and visual screening tool to inform research direction (Maynard, 2006).



Figure 3: 'Fluvial geomorphology in a box'. Micro-model river system allowing the study of fluvial dynamics, similar to the one used by Maynard (2006). Source: <http://www.EMriver.com>.

Froude Number Scaled Models

A true scaled model requires perfect geometric, kinematic and dynamic similitude, something that cannot be achieved when using the same fluid as in the real world system due to equivalent gravitational and fluid motion forces. Therefore, one or more variables must be relaxed in order to achieve model-field similitude (Ashworth *et al.*, 1994; Heller, 2011; Michaelides and Wainwright, 2013). Froude number scaling can be applied, whereby the Reynolds number (Re), a dimensionless quantity used to quantify turbulence rate, is relaxed while correctly scaling the Froude number (Fr), a measurement of different flow states, e.g. subcritical, critical and supercritical. If this was not done, experimental models involving water would have a significantly lower Reynolds number than their counterpart full-scale system, resulting in a lack of similarity between model and reality (Paola, 2000). Instead of having to reduce the viscosity of fluid or to build a 1:1 replica model, Froude number scaling allows a smaller-scale

physical model involving fluid flow to produce similar characteristics to its real-world counterpart. For free surface flow, gravitational forces are dominant. Therefore, hydraulic similarity can be established by equating the ratio of gravitational forces to that of inertial forces (Waldron, 2008). Examples of the effectiveness of Froude number scaling include the work of Ashmore (1982, 1991, 1993) who classified the mechanisms of river braiding and controls on bar formation and related the internal generation of bedload pulses to channel avulsion (see Figure 4). A Froude number scaled model was applied to produce comparable results within the physical model to that of the field counterpart. In Ashmore's study, the use of a Froude number scaled model allowed an understanding of braided channel morphology, flow characteristics and bedload movement where field measurements and observations were challenging due to the large spatio-temporal scales involved.

Distorted Scale Models

Scaled physical models adhere to dimensional scaling of all axes to the same ratio, whereby all attributes within the model are geometrically similar to the original system. However, it is also common for scaled physical models to be geometrically distorted and skewed. Geometrically distorted scaled models, where the scaling of a model's vertical to horizontal scaling ratio differs, are especially important to the geomorphological user when large spatial scales that cannot correctly be replicated under laboratory conditions or fine sediment sizes are involved (Peakall *et al.*, 1996). Distorted scale models enable small physical models to be built or large physical systems to be modelled. Additionally, distorted scale models may be applied to avoid problem of water or fluids behaving viscid at rigid boundaries. Distorted scale model experiments may involve variables such as width, length, slope and/or grain size/density adhering to differing scaling factors. For example, McCollum (1988) used a distorted flume to understand sediment transport dynamics along a 7km river reach which experienced significant rates of sedimentation. Because of the impracticability of reproducing a 7 km flume under laboratory conditions and because

conducting field studies would not allow experimental control over system variables (e.g. slope and/or discharge), a distorted flume with horizontal and vertical scaling ratios of 1:120 and 1:80 respectively was used. Furthermore, crushed coal was used to avoid unrealistic cohesion within the scaled model to ensure the distorted scale model produced a similar response to the field system. The San Francisco Bay Model, a working hydraulic model of the San Francisco Bay and Sacramento-San Joaquin River Delta system is also an example of a geometrically skewed physical model, with horizontal and vertical scaling being 1:1,000 and 1:100, respectively. Furthermore, the model operates at a temporal scale of 1:100, with one diurnal cycle being represented in approximately 15 minutes.

Analogue 'Similarity of Process' Models

Analogue models are models that reproduce certain features of a natural system even though the processes, forms, dynamics,

behaviour, materials and/or geometries do not conform to scaling ratios of the actual system (Chorley, 1967; Hooke, 1968). These are useful when true similarity between model and original system is unachievable or unnecessary. Analogue models may appear to be considerably different from the original field system but are based upon Hooke's (1968) 'similarity of process' concept, whereby the laboratory setup is considered a small system in its own right, rather than a scaled down reality. Seen as models and not miniature reproductions, analogue models should be treated as real, albeit simple physical systems (Paola, 2000) relying on the premise that processes occurring within a natural system will be comparable to those within a laboratory environment (Clarke *et al.*, 2010). This allows analogue models to output detailed and transferable process understanding rather than an understanding that is case study specific. Advantages of using analogue models over other physical model types include their potentially rapid

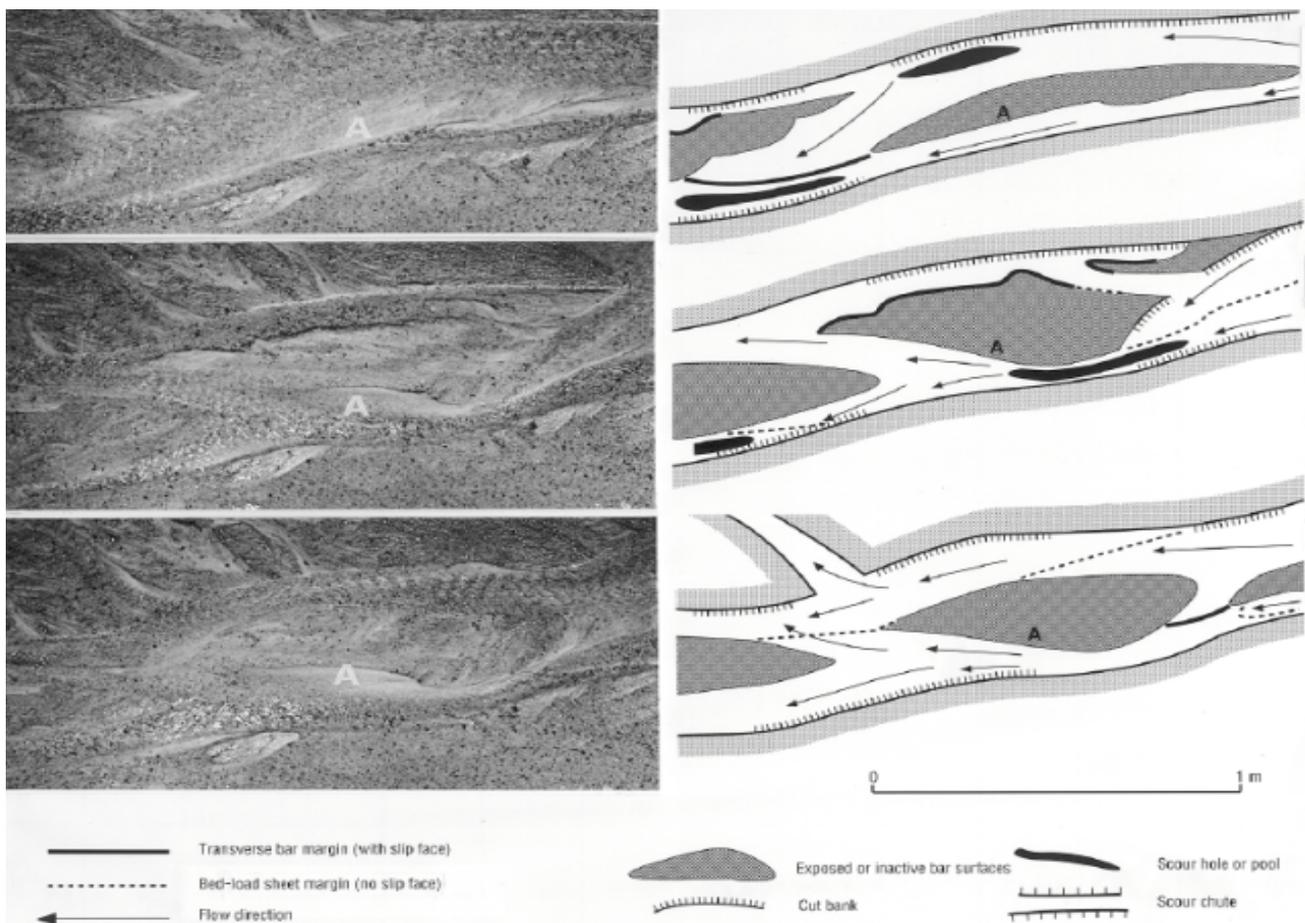


Figure 4: Flume study of a braided river system showing medial bar destruction caused by longitudinal translation and change in total discharge of an upstream confluence, demonstrating the influence that physical modelling has had upon braided river system understanding. Elapsed time of physical model simulation is 1 hour. Source: Ashmore (1991).

setup times, their ability to conduct prompt scenario testing and the reduced quantitative extrapolation required to make conclusions upon. However, analogue models may encounter difficulties with relating measurements and results obtained within the modelling environment to real world situations (Hooke, 1968; Isidoro *et al.*, 2012).

1:1 Replica Models

Some systems are small enough to be replicated in a laboratory and can be simulated at 1:1 scale (i.e. maintaining the exact dimensions of the studied physical system). This allows the studied system to be modelled under laboratory conditions with little or no difference (Peakall *et al.*, 1996). This has a number of advantages, such as a large degree of experimental control over model parameters. However, 1:1 replica models are not suitable for large-scale geomorphological systems due to space limitations within a laboratory setting.

Numerous 1:1 flume studies exist, such as Wilson *et al.* (2013) who used observations from an unscaled flume to understand fluvial bedload abrasion rates. Additionally, 1:1 experiments can be conducted in the field under natural settings but with controlled inputs and conditions. The Outdoor Stream Laboratory, Minnesota, is a field-size reproduction of a fluvial system which is part of the Saint Anthony Falls Laboratory. This allows an understanding of the underlying physical, biological and chemical mechanisms that govern stream and riparian processes and their response to natural and human disturbances under controlled conditions, e.g. steady and unsteady inlet hydrographs, to simulate overbank flood dynamics. Additionally, the Laboratory for Experimental Geomorphology in Leuven, as well as Moss and Walker's (1978) experiments, have conducted 1:1 laboratory experiments focusing on a number of in-situ surface erosion processes and their relationship to surface material properties. These include soil and rain splash erosion plot studies (e.g. De Ploey and Moeyerson, 1975; De Ploey *et al.*, 1976; De Ploey and Mucher, 1981) and tillage experiments and rill development (discussed in Slaymaker, 1991).

Principles of Physical Modelling

Despite the long history of physical modelling studies (Da Vinci used physical models to observe flow characteristics in the 1500s, Reynolds conducted moveable bed models of the River Mersey, UK in 1885 and the US Army Corps of Engineers commissioned multiple large-scale physical modelling experiments from the 1920s onwards, e.g. Coastal Engineering Research Centre and the Waterways Experiment Station; Markle, 1989) there is currently no established framework for conducting experimental and physical modelling studies. At present, laboratories using physical models adopt their own individual approaches based on institutional experience or communication with other similar projects (Frostick *et al.*, 2011). Because physical models may be used for a variety of geomorphological applications, procedures vary significantly between projects. Despite this, there are a number of unifying principles that all physical modelling projects should consider, including: (i) similitude requirements; (ii) dimensional analysis; and, (iii) the materials that are used.

Similitude

Similitude, also known as similarity, involves the model resembling and being correspondent to the system which the model is based upon (Hughes, 1993). Similitude can be divided into three types: geometric (*form*); kinematic (*motion*); and, dynamic (*force*) similitude (Yalin, 1971). Firstly, geometric similitude involves the physical model being similar to its real world counterpart in regards to its dimensions and measurements, involving similarity in form. Therefore, a reduced or enlarged reproduction of the studied physical system is needed to achieve geometric similitude. Secondly, kinematic similitude involves similarity in motion being achieved between the model and real world system, with the ratio of movement in both systems being directly proportional. As a result, true kinematic similitude produces model particle/flow pathways that are geometrically similar to the actual physical system. Finally, dynamic similitude involves the proportion of relevant forces acting upon fluid flows and boundary surfaces being comparable between the model and full scale systems. This produces length, mass and time measurements that are proportionate,

implying a constant ratio of forces between both systems. To achieve dynamic similitude, both geometric and kinematic similitude is required.

The degree to which similitude is satisfied is dependent upon: (i) the aims and objectives of the researcher; (ii) whether the physical model is generating qualitative or quantitative data and; (iii) whether the model can be calibrated and adjusted using existing data or models (Maynard, 2006).

Scale effects which arise due to force ratios being incomparable between a model and its real-world counterpart, and laboratory effects which arise due to the inability of a laboratory to simulate the correct forcing conditions and model boundaries (Chanson, 1999; Heller, 2011) may hinder dynamic similitude. Full dynamic similitude within a scaled physical model is often difficult, if not impossible, to achieve due to force vectors being required to be equal between both systems. Scaled models involving water are unable to achieve dynamic similitude as the model fluid would be needed to have a different viscosity to its real-world counterpart (Ettema, 2000; Frostick *et al.*, 2011). To avoid using a different liquid from the real-world system, users may either use a physical model that functions at full scale (1:1), or use Froude number scaled models to relax the similitude requirements. Because all geomorphic physical models are unique, achieving similarity between model and actual system can be relaxed as long as the similitude requirements are justified and reasonable.

Dimensional analysis

For scaled physical models to be representative of their full-scale system, quantities measured may adhere to scaling laws. Neglect of scaling considerations may render model results meaningless for scientific interpretation or prevent the model from correctly predicting process-form interactions at the actual system scale (Frostick *et al.*, 2011). Scale models are based on similitude theory (above). One method of achieving similitude is by producing a series of dimensionless parameters that are able to form relationships between physical processes (Peakall *et al.*, 1996). Yalin (1971) notes that the dimension of any physical system can be characterised

in terms of its fundamental dimensions; *length, time and mass*. Using this concept, dimensional analysis, involving the examination of the relationships between different physical parameters by identifying their fundamental dimensions (*time, mass, length*) to determine the derived quantities (e.g. *area, volume, force, velocity, frequency* which are a function of the fundamental dimensions) can be applied (Yalin, 1971; Hughes, 1993; see Table 2). A detailed overview of dimensional analysis is beyond the scope of this paper and readers should refer to Yalin (1971) and chapter 2 of Hughes (1993) for further discussions.

Table 2: Dimensions of physical entities using a mass system of units. Source: Adapted from Hughes (1993).

Physical property	Dimensions	Type of quantity
Fundamental quantities		
Time	$[T]$	-
Mass	$[M]$	-
Length	$[L]$	Geometric
Temperature	$[\theta]$	-
Angle	$[1]$	(Supplementary)
Derived quantities		
Area	$[L^2]$	Geometric
Volume	$[L^3]$	Geometric
Force	$[MLT^{-2}]$	Dynamic
Velocity	$[LT^{-1}]$	Kinematic
Acceleration	$[LT^{-2}]$	Kinematic
Volumetric Flow Rate	$[L^3T^{-1}]$	Kinematic
Strain	$[1]$	Dimensionless

Materials

The use of materials in geomorphological research is often highly project-specific. Because of this, a few case studies and the author's experience have been highlighted allowing users to make informed but not constrained decisions.

Physical models may use the exact materials as their real world counterpart, e.g. soil, gravel, flora and fauna in nature and in the physical model (Frostick *et al.*, 2011). Conversely, physical models may use surrogate / proxy materials that differ from the

actual materials present within the physical system. This includes the scaling down of materials within a physical modelling environment and using alternative materials which mimic or substitute the use of the actual materials present in nature. Examples include: (i) using sand instead of gravel; (ii) substituting live vegetation with a smaller species or using an artificial surrogate; (iii) using sponge to represent soil; (iv) using lighter bed materials like pumice or charcoal to represent larger clasts; or, (v) using of fluids with differing viscosities to that of the real world system.

Using similar sediment in a scaled physical model to that found in an actual system may lead to responses in model behaviour that are not comparable to the real world counterpart. This has been documented in scaled flume studies, where using similar sediments resulted in the formation of ripples that had no equivalent in the field (Peakall *et al.*, 1996). This affected the flow dynamics, leading to supercritical flow, hydraulic jumps and standing waves that influenced bed morphology and rates of erosion (Peakall *et al.*, 1996). To avoid this, lighter bed material such as pumice, charcoal or sand can be used (Hughes, 1993).

Vegetation is commonly used in physical models. When using vegetation in physical

models either: (i) artificial / surrogate plants; (ii) scaled, smaller species; or, (iii) natural vegetation can be used (Frostick *et al.*, 2011; see Table 3). When using artificially scaled substitutes, careful consideration must be taken to ensure that these are representative of the actual physical system. Using artificial vegetation has the benefit that it is inert, controllable and easy to use. However, the user must be aware of the limitations associated with using a proxy material to represent a highly variable component of a physical system. Limitations of using artificial vegetation include: (i) misrepresentation of a plant surrogate to replicate the behaviour of natural vegetation; and, (ii) the vegetation characteristics (e.g. flexibility and density) not being comparable between model and nature. Natural vegetation has the benefit that it is directly comparable to that of a natural system, however, it is also highly variable and potentially difficult to maintain in laboratory conditions (Frostick *et al.*, 2011; Frostick *et al.*, 2014). Scale is also important to consider. Modellers would need to substitute larger vegetation types, e.g. trees, with smaller saplings or shrubs due to the space limitations associated with using an experimental set-up. Therefore, the plant materials used depends upon number factors including the scale of the flume and the purpose of the experiments.

Table 3: Choice of plants in physical modelling. Source: modified from Frostick *et al.* (2011)

	Choice of plant	Purpose	Example publications
Artificial / surrogate	Rods or wooden dowels	Stem density effects on drag and flow resistance; Flow resistance on flood plains	Nepf (1999), Stone and Shen (2002), James <i>et al.</i> (2004), Gao <i>et al.</i> (2011)
	Rods with strips, plastic strips or strips with foliage attached	Flow structures; vegetation-flow interactions	Pashe and Rouvé (1985), Naot <i>et al.</i> (1996), Rameshwaran and Shiono (2007), Wilson <i>et al.</i> (2008)
	Plastic bushes or grasses	Floodplain roughness on flow structures, bedforms and sediment transport rates	Shiono <i>et al.</i> (2009)
Scaled	Smaller vegetation (e.g. <i>Medicago sativa</i>)	Flow resistance and controls on stream morphodynamics	Järvelä (2002), Coulthard (2005), Gran and Paola (2001), Clarke <i>et al.</i> (2014)
Natural	Natural, full-scale vegetation, such as grass, shrubs or trees	Flow resistance; plant-flow interactions	Stephan and Gutknecht (2002), Wilson and Horritt (2002), James <i>et al.</i> (2004), Carollo <i>et al.</i> (2005)

Cellulose sponge can be used as a proxy material for soil. Richardson and Siccama (2000) investigated the validity of the simile 'soils are like sponges', demonstrating through experimental methods that sponges store and release water in much the same ways that soils do, with cellulose sponge having intermediate hydrological characteristics to peat and topsoil (see Figure 5). Although it is identified that sponge has a higher water retention capacity than soil, over 2.5 times than peat soils, Richardson and Siccama (2000) do not address the fact that sponge could be scaled volumetrically by thickness/depth to account for this additional storage, which the author plans to apply within a physical model to simulate soil storage capacity during surface-water flood events. Scaling by storage capacity allows sponge to provide a clean and non-erodible medium to investigate runoff and infiltration processes. Sponge can also be compressed to remove stored water, allowing rapid repetition of experimental runs. This highlights the potential benefits of using sponge as a proxy material in physical models. Although sponge allows numerous benefits to the physical modeller, sponge would not be suitable at 1:1 scale; using soil would produce more realistic outputs and avoids using proxy/surrogate materials.

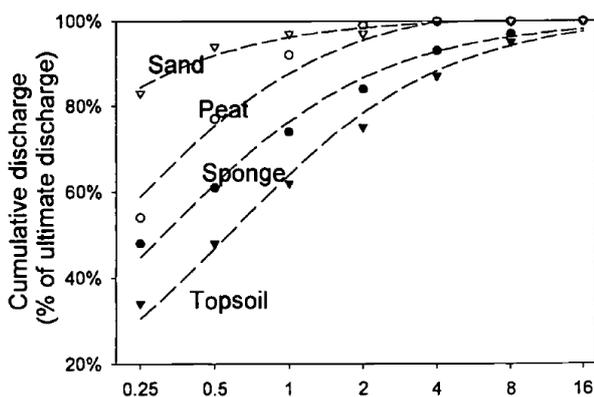


Figure 5: Results from laboratory testing of cumulative discharge of water from four media, demonstrating sponge's ability to act as a proxy material for soil. Other experimental tests including rates of flow, water potential curves and gravimetric water contents in cellulose sponge were also intermediate to that of topsoil and peat. Source: Richardson and Siccama (2000).

Applications of Physical Models

Physical models have an important role in geomorphological research. To demonstrate the scope and potential of physical modelling in geomorphological research, fluvial-, glacial-, aeolian- and bio- geomorphological case studies have been highlighted. Physical models have also been used within coastal- (e.g. Dalrymple, 1985; Markle, 1989; Hughes, 1993; Rossetto *et al.*, 2011) and hillslope-geomorphology / soil erosion studies (e.g. Giménez and Govers, 2001; Parsons and Wainwright, 2006; Michaelides and Wainwright, 2008; Cooper *et al.*, 2012; Turnball *et al.*, 2013). The reader is advised to consult relevant studies and references therein.

Fluvial Geomorphology

Physical models have been used extensively in fluvial geomorphology to understand a plethora of geomorphic processes, including sediment transport, river channel change and the influence of vegetation on channel adjustment. Fluvial geomorphological research using experimental methods is predominantly flume-based. Fluvial physical models were of crucial importance in the work of Hooke (1968), who used laboratory streams to develop the 'similarity of process' model concept, as well as a number of US Army Corps of Engineers projects, e.g. the SEDflume project, a 6m long mobile flume which can analyse fluvial sediment sorting, and the Ice Harbour Lock and Dam Physical Model Study, a 1:55 scale dam commissioned to understand the downstream impacts of river impoundment. Ashworth *et al.* (2007) applied an experimental basin model of an aggrading braided river channel to investigate the relationship between the frequency of channel avulsion, the duration of time that the braidplain is occupied by flow, the spatial pattern of sedimentation and how these respond to a change in sediment supply. Results obtained from the physical model demonstrated a strong positive relationship between sediment supply and channel avulsion rates. Results attained within the physical modelling environment were also able to be extrapolated to real-world examples to gain an understanding of braided river sedimentation. Furthermore, Schumm's (1987) book 'Experimental Fluvial Geomorphology', compiles research from the

fluvial physical modelling literature, comprising studies relating to drainage basin, rivers and fans and fluvial landform development.

Other examples of physical models in fluvial geomorphology include the work of Smith (1998), who applied flume studies to model the development of channel migration and avulsion in high sinuosity meandering channels, and the work of Ashmore (1982, 1991, 1993) which demonstrated how flumes may be applied to study channel morphodynamics. More recently, the work of Braudrick *et al.* (2009) used a scaled flume study to explore mechanisms controlling migration rate, sinuosity, floodplain formation and planform morphodynamics in meandering river channels (see Figure 6). Additionally, Johnson and Whipple (2010) used a scaled experimental flume to model bedrock incision rates by building a weak concrete streambed within a flume to understand rates of erosion relating to sediment flux.

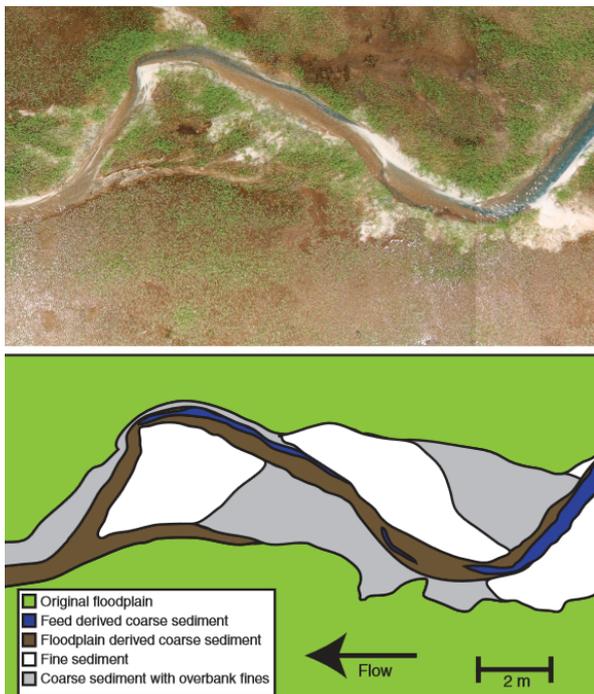


Figure 6: Sediment in second and third bars downstream from the flume inlet. Fine sediment is mapped where the majority of the floodplain thickness was fine sediment. Accumulation of organic matter from the dead alfalfa makes some of the bar appear brown where it is primarily fine sediment. Source: Braudrick *et al.* (2009).

Glacial Geomorphology

Published research on physical models in glacial geomorphology is sparse. The few studies that exist include Rushmer (2007), who applied experimental flume methods to study the impact of glacial outburst floods with differing hydrographs and Corti *et al.* (2008) who used physical modelling to investigate the influence of bedrock topography and ablation on ice flow direction and velocity using silicone gel (see Figure 7). This study confirmed current conceptual models of ice flow around obstacles, demonstrating that variations in bed topography and internal layers of the ice are strongly influenced by the presence and height of bedrock obstacles.

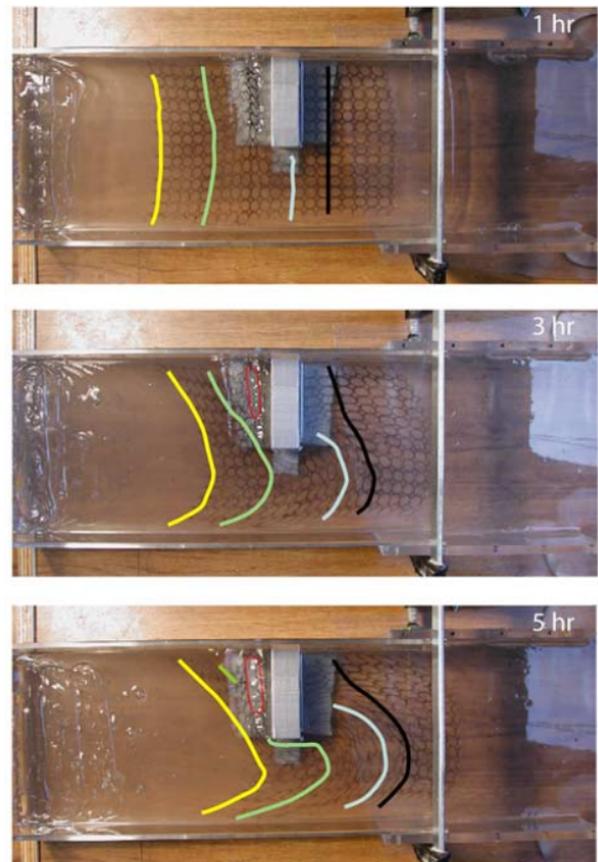


Figure 7: Physical model of a glacier, showing progressive deformation of silicone gel, an ice surrogate material, around an obstacle. Source: Corti *et al.* (2008).

Glacial geomorphology is generally investigated using numerical, rather than physical models to describe relationships between mass balance, ice dynamics and climate (Rowan, 2014), however, some aspects of glacial behaviour can be simulated using physical models, such as controls on

ice melting (e.g. Reznichenko *et al.*, 2010), ice flow (e.g. Glen, 1955) and sub-glacial erosion and sediment transport processes (e.g. Iverson, 1990). Despite this lack of research, Corti *et al.* (2008) express that physical models exhibit numerous opportunities for the glacial geomorphologist, such as the ability to isolate variables and study long spatio-temporal scales.

Aeolian Geomorphology

Theoretical understanding and the development of numerical models of Aeolian processes often contain empirical coefficients that need to be determined using wind tunnel tests, where variables such as grain size and wind speed/direction can be systematically controlled to investigate interactions (Dong *et al.*, 2003). Authors such as Dong *et al.* (2003) and Han *et al.* (2011) have applied experimental wind tunnel tests to understand the relationships between flow velocities and sediment entrainment under differing wind velocity, grain size and moisture scenarios. These studies have confirmed the importance of using physical models to understand aeolian mechanisms.

Biogeomorphology

Biogeomorphology, the study of the interactions between flora and fauna and the

development of landforms, is an emerging topic within geomorphology (Frostick *et al.*, 2011). Using flume experiments, Statzner *et al.* (2000) conducted ecological experiments to demonstrate that crayfish activity significantly affects sand and gravel erosion by increasing bed roughness, decreasing bedform height and altering the pool-riffle sequence downstream. More recently, Johnson *et al.* (2010) highlight that the presence of signal crayfish may affect river bed stability by modifying the microtopography and grain-grain fabric of gravel substrates which can significantly affect bed stability during subsequent flood events. Additionally, Gran and Paola (2001) used a series of physical modelling experiments to study the influence of riparian vegetation upon river morphology and braided stream dynamics. Furthermore, Tal and Paola (2010) conducted laboratory experiments to demonstrate that riparian vegetation can cause a braided channel to maintain a dynamic and single-threaded channel. In these studies, physical modelling allowed input variables, such as water discharge, sediment discharge and grain size to remain constant between runs, while vegetation density of alfalfa sprouts was varied between runs, confirming that vegetation acts to increase bank stability and reduce the number of active channels (see Figure 8).

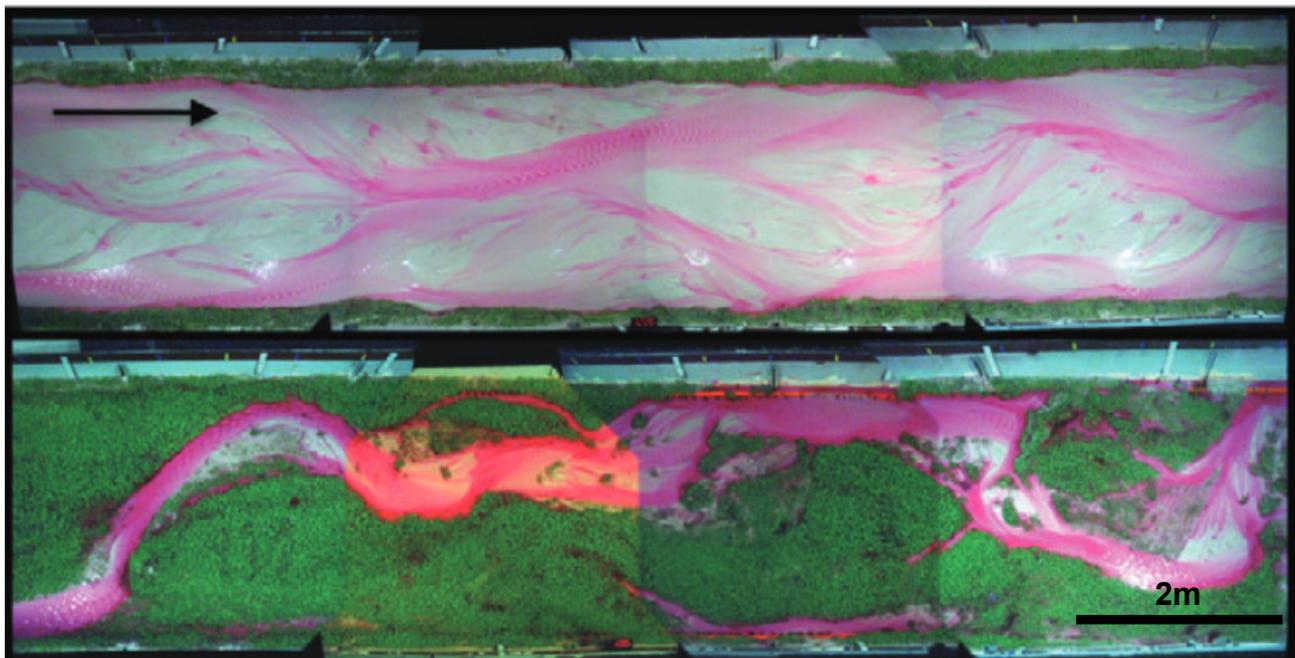


Figure 8: Transition from an un-vegetated braided channel to a dominant single-threaded channel with a vegetated floodplain in an experimental flume experiment. Source: Tal and Paola (2010).

Table 4: Summary of the advantages and disadvantages of using physical modelling and experimental laboratory techniques in geomorphological research framework. Collated using Hughes (1993), HYDRALAB (2004), Frostick et al. (2011), Heller (2011) and Sutherland and Burfuss (2011).

Physical models advantages	Difficulties associated with using physical models
Incorporation of the appropriate physical processes without simplification or assumption. Allows the reproduction of complex physical phenomena.	Potential scale effects associated with simulating model variables in incorrect ratios. Consideration must be taken during planning stages.
Experimental control within a closed system allows rapid multi-variant analysis and testing of multiple variables. Ability to exclude extrinsic parameters.	Laboratory / model effects. Factors may be misrepresented / incorrectly reproduced after simulation in a laboratory environment.
Data collected simultaneously and with relative ease over large spatio-temporal scales once model is constructed and calibrated.	Exclusion / neglect of important functions and conditions which may have been overlooked or deemed to be insignificant by the experimenter.
Large degree of experimental control allows easy simulation of infrequent or hypothetical environmental conditions which would be difficult to observe in nature.	Construction and running is potentially expensive, labour intensive and time consuming. May require appropriate and continued support and funding.
Allow instant visual feedback. Provides qualitative insight into physical processes occurring. Calibration may be assisted by visual prompts/direct contact with physical model.	Data extraction can be difficult due to measurement effects. Results obtained may not be upscaled to real-world situations / directly extended beyond the physical model.
Natural non-linear feedbacks and uncertainty in physical systems which may not be fully understood may be represented and modelled.	Construction and application may require previous experience, understanding or specific expertise.
Can be combined with other techniques to create 'hybrid/composite models', or used to calibrate or inform numerical model functioning and understanding.	May require specialist facilities and/or a large amount of space. Space constraints/lack of equipment may hinder experimentation.
Well-established technique applied to range of research applications. Numerous measurement techniques available, e.g. particle image velocimetry, Acoustic Doppler Velocimetry, pressure sensors, digital photogrammetry, laser scanning etc.	Substitution of materials may be required to ensure correct scaling. Physical model may not be fully representative of actual physical system.
May have reduced costs associated with data collection when compared to field data collection if using existing facilities/equipment.	Simulation of variables or conditions may not be possible at reduced scale within a physical modelling environment.
Control over system variables and inputs, e.g. sediment, water, vegetation. Bridges what can be simulated in the field and modelled numerically.	Equifinality may result in a misinterpretation of the fundamental processes occurring.

These physical modelling studies demonstrated the role of biota as a significant geomorphic agent. Many aspects of this field remain poorly understood but the use of physical models is of critical importance. Readers are advised to consult Thomas *et al.* (2014) and Frostick *et al.* (2014) which provide detailed overviews of the use of physical models in biogeomorphology, as well as outlines of knowledge gaps and avenues for future research.

Advantages of Physical Models

Physical models provide a number of advantages to the geomorphological user (summarised in Table 4). The main advantages of physical modelling are associated with the controlled, closed environment in which experimentation can take place. Physical models allow rapid analyses of multiple variables with a large degree of experimental control – independent variables can be altered one at a time while dependent variables can remain constant to investigate cause and effect relationships and model responses to changing variables. Additionally, physical models allow the simulation and study of infrequent, hypothetical or large spatiotemporal scale scenarios. This is significant for events which may be impossible to observe or difficult to study in the field because of the long timescales involved, e.g. the influence of autogenic mechanisms on alluvial fan evolution (Clarke *et al.*, 2010). Furthermore, physical models allow complex physical phenomena (potentially not yet described or understood) to be simulated without requiring a mathematical or theoretical simplification of governing processes (Goudie, 2003). This makes physical models an invaluable investigative tool to the geomorphological user.

Physical Model Limitations

Despite offering a number of advantages to the geomorphological user, physical modelling also has a number of shortcomings which the user must be aware of before any experimentation takes place (see Table 4).

Firstly, laboratory effects due to the limitations associated with simulating natural phenomena under a simplified and scaled

laboratory set-up may produce occurrences that are not present in natural systems. These may include cohesive and/or adhesive forces between molecules (e.g. clay or water) becoming greater than within a natural system (Schumm, 1960; Goudie, 2003). Additionally, scale effects, whereby fundamental phenomena are unable to be simulated in correct proportions to that of the physical system, may arise (Heller, 2011). These may render results misleading or incorrect. In addition, difficulties in extracting useful and transferable data from physical models may be encountered (Hooke, 1968; Isidoro *et al.*, 2012), whereby data obtained within the physical modelling environment cannot be upscaled to real world scenarios. Problems associated with equifinality, where the same end state is reached through different processes and mechanisms may also be present within physical models. Furthermore, physical models are potentially difficult to validate and determine whether the model is performing adequately because multiple model runs are required to allow adjustment of model variables until the observed effects are comparable to those observed in nature (Hooke, 1968). Validation is essential to ensure that a physical model performs relative to its real-world counterpart but is rarely considered in physical modelling. Despite this, problems with model validation are problematic in other modelling techniques.

Conclusions

Physical models permit clear visualisation, observation, demonstration and measurement of process-form interactions. This allows an understanding of complex relationships that cannot be represented mathematically, as well as allowing the verification of numerical modelling approaches (Frostick *et al.* 2011). Yalin (1971) states that physical models give the user an instant qualitative, visual insight into the processes occurring; something that is difficult in field or numerical modelling situations. Physical modelling provides an excellent tool to geomorphologists, however, users must be conscious of modelling limitations so these can be minimised (Ettema, 2000). Hughes (1993) compares a poorly scaled model to a ruler with incorrect markings – the ruler can be used to make measurements but the measurements are

guaranteed to be wrong, with incorrectly designed models always providing inaccurate predictions (Yalin, 1971).

Paola (2000) asserts that it is potentially misleading to treat even the most carefully controlled scaled model as a miniature analogue of its field system due to the limitations associated with scaling and reproducing a model under laboratory conditions. Users must be aware of the limitations of physical modelling approach, as well as procedures to address and reduce model shortcomings, before conducting such experiments.

Acknowledgements

The author would like to thank Dr. Ian Pattison and Dr. Dapeng Yu for their constructive comments, as well as the comments from two anonymous reviewers.

References

- Allen JRL. 1982. *Sedimentary Structures: Their Character and Physical Basis*. Elsevier, Amsterdam: 539.
- Ashmore PE. 1982. Laboratory modelling of gravel braided stream morphology, *Earth Surface Processes and Landforms* **7**: 201 – 225.
- Ashmore PE. 1991. Channel morphology and bed load pulses in braided, gravel-bed streams, *Geografiska Annaler* **68**: 361 – 371.
- Ashmore PE. 1993. Anabranch confluence kinetics and sedimentation processes in gravel-braided streams. In: Best JL, Bristow CS. (Eds.) *Braided Rivers*. Geological Society Special Publications 75: 129 – 146.
- Ashworth PJ, Best JL, Jones MA. 2007. The relationship between channel avulsion, flow occupancy and aggradation in braided rivers: Insights from an experimental model, *Sedimentology* **54**: 497 - 513.
- Ashworth PJ, Best JL, Leddy JO, Geehan GW. 1994. The physical modelling of braided rivers and deposition of fine-grained sediment. In: Kirkby MJ. (Eds.) *Process Models and Theoretical Geomorphology*. John Wiley & Sons, Chichester: 115 – 139.
- Braudrick CA, Dietrich WE, Leverich GT, Sklar LS. 2009. Experimental evidence for the conditions necessary to sustain meandering in course-bedded rivers, *Proceedings of the National Academy of Science USA* **106**: 16936 - 16941.
- Bryan RB, Poesen J. 1989. Laboratory experiments on the influence of slope length on runoff, percolation and rill development, *Earth Surface Processes and Landforms* **14**: 211 – 231.
- Carollo FG, Ferro V, Termini D. 2005. Flow resistance law in channels with flexible submerged vegetation, *Journal of Hydraulic Engineering* **131**: 554 – 564.
- Chanson H. 1999. *The Hydraulics of Open Channel Flow: An Introduction*, Butterworth-Heinemann, Oxford: 512.
- Chorley RJ. 1967. Models in geomorphology. In: Chorley, RJ, Haggett P. (Eds.) *Models in Geography*. Methuen, London: 59 – 96.
- Clarke LE, Quine TA, Nicholas A. 2010. An experimental investigation of autogenic behaviour during alluvial fan evolution, *Geomorphology* **115**: 278 – 285.
- Clarke LE. 2013. Experimental physical modelling. Available: http://gees-talk.blogspot.co.uk/2013_08_01_archive.html Last accessed: 04/08/2014.
- Clarke LE. 2014. The use of live vegetation in geomorphological experiments: how to create optimal growing conditions, *Earth Surface Processes and Landforms* **39**: 705 – 710.
- Cooper J, Wainwright J, Parsons AJ, Onda Y, Fukuwara T, Obana E, Kitchener B, Long EJ, Hargrave H. 2012. A new approach for simulating the redistribution of soil particles by water erosion: A marker-in-cell model. *Journal of Geophysical Research, Earth Surface*, **117**: DOI: 10.1029/2012JF002499.
- Corti G, Zeoli A, Belmaggio P, Folco L. 2008. Physical modelling of the influence of bedrock topography and ablation on ice flow and meteorite concentration in Antarctica, *Journal of Geophysical Research* **113**: 1 – 18.
- Coulthard T. 2005. Effects of vegetation on braided stream pattern and dynamics, *Water Resources Research* **41**: 1 – 9.
- Dalrymple RA. 1985. Introduction to physical models in coastal engineering. In: Dalrymple RA. (Eds.) *Physical Modelling in Coastal Engineering*. Rotterdam, The Netherlands: 3 – 9.

- De Ploey J, Mucher HJ. 1981. A consistency index and rainwash mechanisms on Belgian loamy soil, *Earth Surface Processes and Landforms*, **6**: 319 - 330.
- De Ploey J, Moeyersons J. 1975. Runoff creep of coarse debris: experimental data and some field observations, *Catena*, **2**: 275 - 288.
- De Ploey J, Savat J, Moeyersons J. 1976. The differential impact of some soil factors on flow, runoff creep and rainwash, *Earth Surface Processes*, **1**: 151 - 161.
- Dong Z, Liu X, Wang H, Wang, X. 2003. Aeolian sand transport: a wind tunnel model, *Sedimentary Geology* **161**: 71 – 83.
- Einhellig R, Svoboda C, Frizell K, Cox N. 2010. *Physical modelling of the Folsom Dam tailwater confluence area*. Proceedings of the 30th Annual USSD Conference, April 2010, Sacramento, California.
- Ettema R, Muste M. 2001. Laboratory observations of ice jams in channel confluences, *Journal of Cold Regions Engineering, ASCE* **15**: 34 – 58.
- Ettema R. 2000. *Hydraulic Modelling: concepts and practise*. ASCE, Reston, VA: 383.
- Frostick LE, Thomas RE, Johnson MF, Rice SP, McLelland SJ. 2014. *Users Guide to Ecohydraulic Modelling and Experimentation: Experience of the Ecohydraulic Research Team (PISCES) of the HYDRALAB Network*. CRC Press, Leiden, The Netherlands: 228.
- Frostick LE, McLelland SJ, Mercer TG. 2011. *User guide to physical modelling and experimentation: experience of the HYDRALAB network*. CRC Press, Leiden: 245.
- Gao G, Falconer R, Lin B. 2011. Modelling open channel flows with vegetation using a three-dimensional model, *Journal of Water Resource and Protection* **3**: 114 – 119.
- Giménez R, Govers G. 2001. Interaction between bed roughness and flow hydraulics in eroding rills, *Water Resources Research*, **37**: 791 – 799.
- Glen J. 1955. The creep of polycrystalline ice. *In Proceedings of the Royal Society of London*, **228**: 519 – 538.
- Goudie A. 2003. *Encyclopedia of Geomorphology, Volume 1 & 2*. Routledge – Taylor & Francis, London.
- Gran K, Paola C. 2001. Riparian vegetation controls on braided stream dynamics, *Water Resources Research* **37**: 3275 – 3283.
- Guy HP, Simons DB, Richardson EV. 1966. Summary of alluvial channel data experiments, 1956 – 1961, *US Geologic Survey Professional Paper* **461 – 2**: 96.
- Han Q, Qu J, Lia K, Zhu S, Zhang K, Zu R, Niu Q. 2011. A wind tunnel study of aeolian sand transport on a wetted sand surface using sands from tropical humid coastal southern China, *Environmental Earth Sciences* **64**: 1375 – 1385.
- Heller V. 2011. Scale effects in physical hydraulic engineering models, *Journal of Hydraulic Research* **49**: 293 – 306.
- Hooke RL. 1968. Model Geology: prototype and laboratory streams – discussion, *Geological Society of America Bulletin* **79**: 391 – 394.
- Hughes SA. 1993 *Physical Models and Laboratory Techniques in Coastal Engineering*. World Scientific Publishing Co., Singapore: 568.
- HYDRALAB. 2004. Strategy paper: The future role of experimental methods in European hydraulic research - towards a balanced methodology, *Journal of Hydraulic Research*, **42**: 341 - 356.
- Isidoro J, de Lima J, Leandro J. 2012. The study of rooftop connectivity on the rainfall-runoff process by means of a rainfall simulator and a physical model, *Zeitschrift für Geomorphologie* **57**: 177 – 191.
- Iverson N. 1990. Laboratory simulations of glacial abrasion: comparison with theory, *Journal of Glaciology*, **36**: 304 – 314.
- James CS, Birkhead AL, Jordanova AA, O'Sullivan JJ. 2004. Flow resistance of emergent vegetation, *Journal of Hydraulic Research* **42**: 390 – 398.
- Järvelä J. 2002. Determination of flow resistance of vegetated channel banks and floodplains. In: Bousmar D, Zech Y. (Eds.) *River Flow 2002*. International Conference on Fluvial Hydraulics, September 4 – 6, 2002, Belgium: 311 – 318.

- Johnson MF, Rice SP, Reid I. 2010. Topographic disturbance of subaqueous gravel substrates by signal crayfish (*Pacifastacus leniusculus*), *Geomorphology* **123**: 269 – 278.
- Johnson J, Whipple K. 2010. Evaluating the controls of shear stress, sediment supply, alluvial cover and channel morphology on experimental bedrock incision rate, *Journal of Geophysical Research: Earth Surface*, **115**: 1 – 21.
- Kamphuis JW. 1991. Physical Modelling in Herbich JB. (Eds.) *Handbook of Coastal and Ocean Engineering*. Gulf Publishing Company, Houston, Texas: 1152.
- Madej MA, Sutherland DG, Lisle TE, Pryor B. 2009. Channel responses to varying sediment input: a flume experiment modelled after Redwood Creek, California, *Geomorphology* **103**: 507 – 519.
- Markle DG. 1989. Physical models of coastal structures as designed and used by the US Army Corps of Engineers, *Journal of Coastal Research*, **5**: 573 - 592.
- Maynard S. 2006. Evaluation of the Micro-model: an extremely small-scale moveable bed model, *Journal of Hydraulic Engineering* **132**: 343 – 353.
- McCollum RA. 1988. *Blountstown Reach, Apalachicola River: moveable-bed model study*. Technical Report HL-88-17, US Waterways Experimental Station, Vicksburg: 39.
- Michaelides M, Wainwright J. 2002. Modelling the effects of hillslope-channel coupling on catchment hydrological response, *Earth Surface Processes and Landforms* **27**: 1441 – 1457.
- Michaelides K, Wainwright J. 2008. Internal testing of numerical model of hillslope-channel coupling using laboratory flume experiments, *Hydrological Processes*, **22**: 2274 – 2291.
- Michaelides M, Wainwright J. 2013. Modelling Fluvial Processes and Interactions. In: Michaelides K, Wainwright J. (Eds.) *Environmental Modelling: finding simplicity in complexity*. John Wiley & Sons, Chichester: 123 – 138.
- Moss AJ, Walker PH. 1978. Particle transport by continental water flows in relation to erosion, deposition, soils and human activities, *Sedimentary Geology*, **20**: 81 - 139.
- Naot D, Nezu I, Nakagawa H. 1996. Hydrodynamic behaviour of partly vegetated open channels, *Journal of Hydraulic Engineering* **112**: 625 – 633.
- Nepf HM. 1999. Drag, turbulence, and diffusion in flow through emergent vegetation, *Water Resources Research* **35**: 479 – 489.
- Paola C. 2000. Quantitative models of sedimentary basin filling, *Sedimentology* **47**: 121 – 178.
- Parsons AJ, Wainwright J. 2006. Depth distribution of interrill overland flow and the formation of rills, *Hydrological Processes*, **20**: 1511 – 1523.
- Pashe E, Rouvé G. 1985. Overbank flow with vegetatively roughened floodplains, *Journal of Hydraulic Engineering* **111**: 1262 – 1278.
- Peakall J, Ashworth PJ, Best JL. 1996. Physical Modelling in Fluvial Geomorphology: Principles, Applications and Unresolved Issues. In: Rhoads BL, Thorn CE. (Eds) *The Scientific Nature of Geomorphology*. John Wiley & Sons, Chichester: 221 – 254.
- Rameshwaran P, Shiono K. 2007. Quasi two-dimensional model for straight overbank flows through emergent vegetation on floodplains, *Journal of Hydraulic Research* **45**: 302 – 315
- Reznichenko N, Davies N, Shulmeister J. 2010. Effects of debris on ice-surface melting rates: an experimental study. *Journal of Glaciology*, **56**: 384 – 394.
- Richardson AD, Siccama TG. 2000. Are soils like sponges? *Journal of the American Water Resources Association*, **36**: 913 - 918.
- Rossetto T, Allsop W, Charvet I, Robinson D. 2011. Physical modelling of a tsunami using a new pneumatic wave generator, *Coastal Engineering* **58**: 517 – 527.
- Rowan A. 2014. Mountain glacier models. Available: <http://www.antarcticglaciers.org/glaciers-and-climate/numerical-ice-sheet-models/mountain-glacier-models/>, Last accessed: 08/10/2014/.
- Rushmer LE. 2007. Physical-scale modelling of jökulhlaups (glacial outburst floods) with contrasting hydrograph shapes, *Earth*

- Surface Processes and Landforms* **32**: 954 – 963.
- Schumm SA. 1987. *Experimental Fluvial Geomorphology*. John Wiley & Sons, New York: 413.
- Schumm SA. 1960. *The shape of alluvial channels in relation to sediment type*. United States Geological Survey, Professional Paper 352: 17 - 30.
- Shiono K, Chan T, Spooner J, Rameshwaran P, Chandler J. 2009. The effect of floodplain roughness on flow structures, bedforms and sediment transport rates in meandering channels with overbank flows: Part I, *Journal of Hydraulic Research* **47**: 5 – 19.
- Smith C. 1998. Modelling high sinuosity meanders in a small flume, *Geomorphology*, **25**: 19 – 30.
- Soares-Frazão S, Zech Y. 2008. Dam-break flow through an idealised city, *Journal of Hydraulic Research* **46**: 648 – 658.
- Slaymaker O. 1991. *Field Experiments and Measurement Programs in Geomorphology*. University of British Columbia Press, Vancouver: 225.
- Southard JB, Boguchwal LA. 1990. Bed configurations in steady unidirectional water flows, Part 2: Synthesis of flume data, *Journal of Sedimentary Petrology* **60**: 658 – 679.
- Statzner B, Fiévet E, Champagne J, Morel R, Herouin E. 2000. Crayfish as geomorphic agents and ecosystem engineers: Biological behaviour affects sand and gravel erosion in experimental streams, *Journal of Limnology and Oceanography* **45**: 1030 – 1040.
- Stephan U, Gutknecht D. 2002. Hydraulic resistance of submerged flexible vegetation, *Journal of Hydrology* **269**: 27 – 43.
- Stone BM, Shen HT. 2002. Hydraulic resistance of flow in channels with cylindrical roughness, *Journal of Hydraulic Engineering* **128**: 500 – 506.
- Sutherland J, Barfuss SL. 2011. *Composite modelling: combining physical and numerical models*. Proceedings of the 34th IAHR World Congress, June 2011, Brisbane, Australia.
- Tal M, Paola C. 2010. Effects of vegetation on channel morphodynamics: results and insights from laboratory experiments, *Earth Surface Processes and Landforms*, **35**: 1014 – 1028.
- Thomas R, Johnson M, Frostick L, Parsons D, Bouma T, Dijkstra J, Eiff O, Gobert S, Henry P, Kem P, McLelland S, Moulin F, Myrhaug D, Neyts A, Paul M, Penning E, Puijalón S, Rice S, Stanica A, Tagliapietra D, Tal M, Torum A, Voudoukas M. 2014. Physical modelling of water, fauna and flora: knowledge gaps, avenues for future research and infrastructural needs, *Journal of Hydraulic Research*, **52**: 311 – 325.
- Turnball L, Parsons AJ, Wainwright J, Anderson JP. 2013. Runoff responses to long-term rainfall variability in a shrub-dominated catchment, *Journal of Arid Environments*, **91**: 88 – 94.
- Waldron R. 2008. *Physical modelling of flow and sediment transport using distorted scale modelling*. MSc Thesis, Louisiana State University, USA.
- Warburton J, Davies TRH. 1998. The use of hydraulic models in the management of braided gravel-bed rivers. In: Klingeman PC, Beschta RL, Komar RD, Bradley JB (Eds) *Gravel-bed Rivers in the Environment*. John Wiley & Sons, New York: 832.
- Wilson C, Horritt M. 2002. Measuring the flow resistance of submerged grass, *Hydrological Processes* **16**: 2589 – 2598.
- Wilson A, Hovius N, Turowski JM. 2013. Upstream-facing convex surfaces: Bedrock bedforms produced by fluvial bedload abrasion, *Geomorphology*, **180**: 187 – 204.
- Wilson C, Hoyt J, Schnauder I. 2008. Impact of foliage on the drag force of vegetation in aquatic flows, *Journal of Hydraulic Engineering* **134**: 885 – 891.
- Yalin MS. 1971. *Theory of Hydraulic Models*. Macmillan, London: 266.

5.6.4. Modelling Geomorphic Systems: Fluvial

Michael C. Grenfell¹

¹Environmental and Water Science Division, Department of Earth Science, University of the Western Cape, Bellville, 7530, South Africa

Email: mgrenfell@uwc.ac.za



ABSTRACT: Rivers and floodplains convey and exchange water and all its constituent matter from Earth's surface to an intra-continental or oceanic sink. The associated processes of flow and flux are the foundation of all ecosystem service provision and human value derived from river environments. Numerical modelling is one of many approaches that may be used to understand these processes. It is an approach that seeks quantitative mechanistic understanding that is critical to enhancing the predictive capacity of river science, and to developing evidence-based management practices. In combination with other approaches, modelling should play a key role in constraining understanding of river responses in an uncertain future. Ongoing improvements in computing power and in the availability and accessibility of fluvial modelling codes have substantially increased the uptake of modelling as a method of investigating fluvial processes and forms. This chapter outlines the physical basis of different types of fluvial model, and illustrates the key considerations needed to select a model code with the necessary numerical complexity, to establish a physical model domain and boundary conditions, to test for sensitivity to domain, boundary and parameter variables, and to evaluate results.

KEYWORDS: Numerical modelling, fluvial processes, process-form feedback, morphodynamics

Virtual rivers

Numerical modelling addresses the fundamental mechanisms that drive fluvial processes, and the process-form feedbacks that govern river characteristics, dynamics and socio-ecological value. Numerical models and the virtual rivers they describe are useful tools because they offer the potential for full control over boundary conditions and physical laws (Kleinhans, 2010), thereby providing an abstraction of reality that is modifiable within given physical constraints. This allows one to test hypotheses derived from field data and experiments (Kleinhans, 2010), to evaluate competing explanations, to elucidate *key* controls (e.g. on chute cutoff; van Dijk *et al.*, 2014), or *necessary conditions* (e.g. for a meandering river avulsion; Slingerland and Smith, 1998). Such endeavours typically require insight that extends beyond the limits of field observation, or that is difficult to transfer from laboratory experiments.

However, numerical models provide a tool that should be embedded within a broader conceptual approach to understanding rivers, as the greatest potential for full explanation of natural river phenomena arises when results from field measurements, laboratory experiments and numerical modelling converge (Kleinhans, 2010). More broadly, the process itself of building a virtual river can be valuable to force rigour in setting hypotheses and interpreting results for field or experimental studies (Bras *et al.*, 2003).

The literature on modelling approaches and applications has burgeoned in recent years, as has the availability of commercial and non-commercial fluvial modelling codes (see Table 1 for examples). The former are subject to licence fees, while the latter comprise research or management-centred software developed by universities, government agencies, or an active free and open source (FOSS) community, and are

Table 1: Examples of fluvial channel and channel-floodplain modelling codes in common use by geomorphologists (crudely ordered according to the dimensionality of process-representation).

Fluvial modelling codes	Examples
1D channel and floodplain processes, including sediment transport, erosion and deposition	<ul style="list-style-type: none"> • Example code or numerical basis: HEC-RAS; USACE Hydrologic Engineering Center (available at http://www.hec.usace.army.mil/software/hec-ras/). 2D hydrodynamics currently under development for version 5. • Example model: Energy dissipation in step-pool systems, Chin (2003).
1D morphodynamics for meander migration and floodplain evolution; fixed-width channels	<ul style="list-style-type: none"> • Example code or numerical basis: HIPS Relation (see Parker <i>et al.</i> (2011) for a review). • Example model: bend instability and channel migration in meandering rivers (Ikeda <i>et al.</i>, 1981).
1D/quasi 2D bar dynamics at bifurcations	<ul style="list-style-type: none"> • Example code or numerical basis: nodal point relation; Bolla Pittaluga <i>et al.</i> (2003), modified by Kleinhans <i>et al.</i> (2008). • Example model: bifurcation dynamics and avulsion duration in meandering rivers, Kleinhans <i>et al.</i> (2008).
Reduced complexity 2D flow routing, sediment transport and morphodynamics from basin to reach scales; long-term, landscape controls on river form and process (e.g. trunk-tributary interaction)	<ul style="list-style-type: none"> • Example code or numerical basis: CAESAR-Lisflood; Tom Coulthard, Paul Bates (available at http://code.google.com/p/caesar-lisflood/). • Example model: Estimating sediment yield from river basins (Coulthard <i>et al.</i>, 2013).
2D depth-averaged morphodynamics for meander migration and floodplain evolution; dynamic width variation	<ul style="list-style-type: none"> • Example code or numerical basis: Nays2D; Asahi <i>et al.</i>, 2013 (available at http://i-ric.org/en/). • Example model: co-evolution of river width and sinuosity in a meandering river (Asahi <i>et al.</i>, 2013).
2D depth-averaged morphodynamics for multiple-thread channel planform dynamics	<ul style="list-style-type: none"> • Example code or numerical basis: <ul style="list-style-type: none"> - Delft3D; Deltares (available at http://oss.deltares.nl/web/delft3d/about). - HSTAR; Andrew Nicholas • Example models: dynamic planform evolution and planform transitions in large alluvial rivers (Nicholas 2013a, 2013b; Schuurman <i>et al.</i>, 2013).
2D/quasi-3D hydrodynamics and sediment transport; fixed banklines	<ul style="list-style-type: none"> • Example code or numerical basis: FaSTMECH; USGS Geomorphology and Sediment Transport Laboratory (available at http://i-ric.org/en/). • Example model: bed evolution in flow separation eddies (Nelson and McDonald, 1995).
Full 3D CFD with Large Eddy Simulation and particle tracking	<ul style="list-style-type: none"> • Example code or numerical basis: Hardy <i>et al.</i> (2005). • Example model: Transport of individual particles over a gravel bed (Hardy, 2005).

*Note: this list is not exhaustive, and some codes may overlap different application environments (e.g. Delft 3D has been used in quasi-3D mode with fixed banks to investigate the stability of bifurcations, Kleinhans *et al.*, 2008, and in depth-averaged mode with a bank erosion model to investigate dynamic planform evolution, Schuurman *et al.*, 2013). The examples indicate potential codes for use in common applications, with a focus on codes available in the public domain.*

typically provided at no cost upon request. To those wishing to learn *how to learn* about rivers using numerical models, this growth in interest is inspiring, but also daunting. Key challenges lie in finding a code that is fit for purpose, and in understanding important elements of how the code works. The aim of this chapter is to present a simplified general account of key considerations involved in selecting, running and evaluating results of numerical models for fluvial channel and channel-floodplain applications.

Overview

In broad terms, setting up a numerical fluvial model will require: i) decisions on the complexity of process representation, ii) definition of the spatial and temporal domains, iii) definition of boundary conditions and initial conditions, iv) testing for sensitivity to variation in the values of key parameters, and v) confirmation (*sensu* Oreskes *et al.*, 1994) of results. The process is iterative at all stages; e.g. developing a computational grid requires iterative refinement to meet specified quality criteria, calibration and sensitivity tests may involve iteratively adjusting parameter values to match measurements or define error limits, and confirmation of results is typically expressed in terms of a range of variability in model output derived through iteratively adjusting model variables.

Necessary numerical complexity

All models capable of representing fluvial process-form relations and morphological dynamics achieve three basic things with varying levels of complexity, realism and computational cost (Figure 1): (1) they route water ('hydrodynamics'); (2) they predict sediment transport based on flow characteristics determined by point (1); and, (3) they update the boundary morphology based on fluxes driven by point (2) ('morphodynamics').

Hydrodynamics

Fluvial model codes route water between nodes aligned on a longitudinal profile (1D), or between cells on a grid (2D, quasi-3D or 3D). Some codes combine nodes (for the channel) and a grid (for the floodplain) in lattice-network structures (Bates and De Roo, 2000; Nicholas and Quine, 2007). Each node

or cell has a location that can be defined using a physical spatial coordinate system. Node or cell properties such as elevation, bed characteristics (sediment size, density, porosity, and roughness), and water level are defined using a commensurate physical spatial coordinate system. Water routing schemes calculate fluxes of mass and momentum across the spatial domain and through time according to physical laws that may be represented with varying numerical complexity. One division of this complexity relates to the treatment of momentum, wherein numerical model codes may be grouped as either 'physics-based' or 'reduced-complexity'.

Physics-based codes solve the Navier-Stokes equations (for detailed reviews relevant to fluvial geomorphology, see Lane, 1998, Wright and Baker, 2004, and Ingham and Ma, 2005). Several orders of complexity are observed according to the treatment of turbulence in the momentum flux. For geomorphological applications the highest level of complexity is achieved using the 'Reynolds-averaged Navier Stokes equations' (RANS). These allow simulation of complex flow structures on relatively high-resolution 3D grids, but the equations must be closed using a turbulence model (Ingham and Ma, 2005). Full 3D RANS codes are required to accurately simulate complex flow structures and to simulate the trajectory of individual particles over an irregular bed (Hardy, 2005, 2008).

Within the RANS approach, there are models that represent only the time-averaged flow conditions, and more complex models that are eddy-resolving (e.g. Large Eddy Simulation, LES; Keylock *et al.*, 2005). In LES, a length scale is used to differentiate between large and small eddies, and large eddies are resolved directly on a high-resolution grid, while smaller eddies are parameterised using a sub-grid scale model (Wright and Baker, 2004). LES computations are sensitive to grid resolution and to the method of spatio-temporal discretisation employed (Wright and Baker, 2004). It is important to investigate these issues in code validation documents, especially since turbulence values computed by LES are used in sediment transport equations (discussed in the next section).

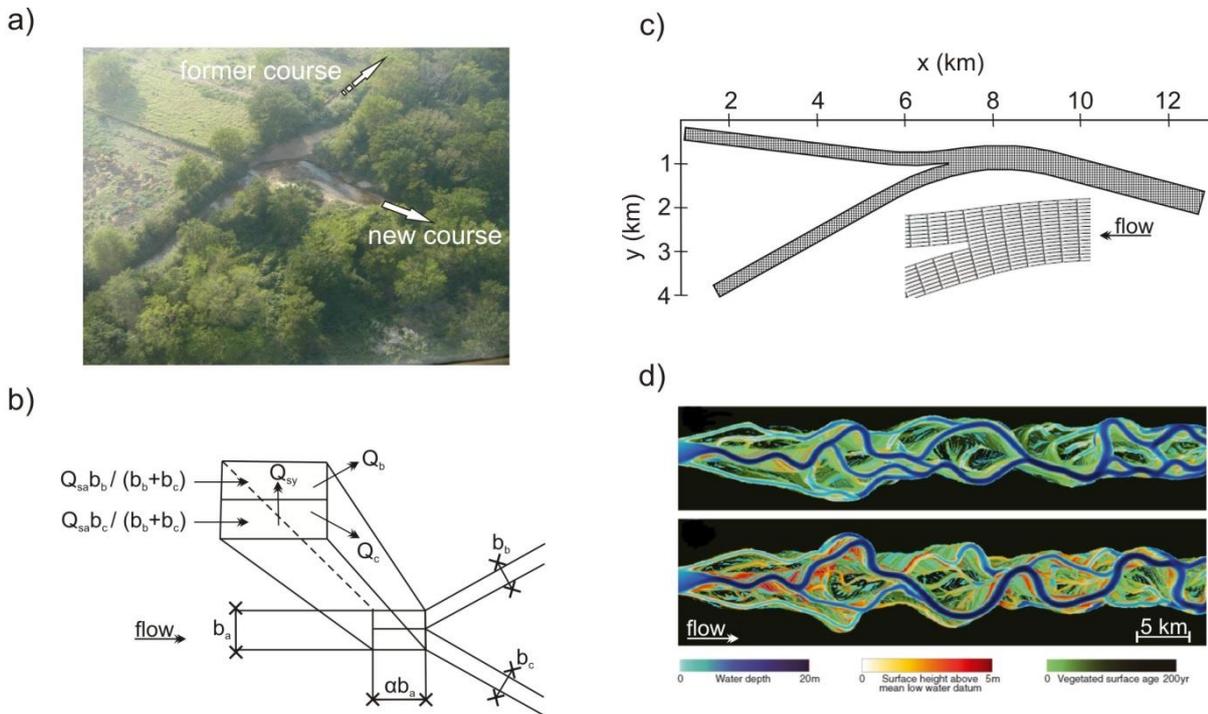


Figure 1: Versatility of numerical modelling in river environments, illustrated through the variety of possible schematisations of the physical space and dynamics of a fluvial geomorphic feature (in this case, a bifurcation) in the computational space of numerical models. a) A photograph of a natural bifurcation (avulsion) on the Mkuze River in eastern South Africa. b) Schematic 1D/quasi-2D nodal point relationship of a bifurcation on a braided river (redrawn from Bolla Pittaluga *et al.*, 2003). c) Curvilinear SWE model grid of an avulsion, with simplified domain geometry (redrawn from Kleinhans *et al.*, 2008). The curvature of the inlet reach may be varied systematically to simulate the effect of the upstream bend configuration on bifurcation dynamics. d) morphodynamic development of bifurcations and an anabranching channel network in a SWE model with a regular grid (after Nicholas, 2013a). Computational demands and process realism increase from b) to d), but the goal of these different approaches varies; c) may be used to test the realism of b), while b) may be used to examine controls on morphodynamics over longer timescales than possible with c). Parameterisation of d) requires an advanced understanding of the controls on bifurcation (such as developed by b) and c)), but seeks to understand the morphodynamics of channel network and coupled channel-floodplain development as a whole.

A further simplification involves reducing the dimensionality of the equation set by integrating over the water depth to yield depth-averaged velocities (the 2D depth-averaged Navier Stokes, or 'shallow water equations', SWE). In addition to the specification of an appropriate turbulence model, the SWE require a model for bed shear stress incorporating a friction coefficient, and parameterisation of secondary flow effects (Ingham and Ma, 2005). Several parameterisations of friction are available that use different roughness formulations (e.g. Chezy, White-Colebrook, Manning), and it is important to test for sensitivity to the parameterisation applied. Finding suitable parameter values is often best achieved through calibration procedures (Figure 2) that aim to maximise the fit

between predicted and observed data (Horritt, 2005).

Some SWE codes are 'quasi-3D' (Wright and Baker, 2004) in that they allow grid-based simulation of secondary circulation through the inclusion of a series of vertical layers (e.g. Delft3D, Lesser *et al.*, 2004). SWE codes are able to accurately reproduce a number of environmental flow processes (e.g. dynamic inundation of floodplains, Nicholas *et al.*, 2006, flow separation at bifurcations, Constantine *et al.*, 2010) at reasonable computational expense, and arguably form the basis of most river morphodynamic models at present.

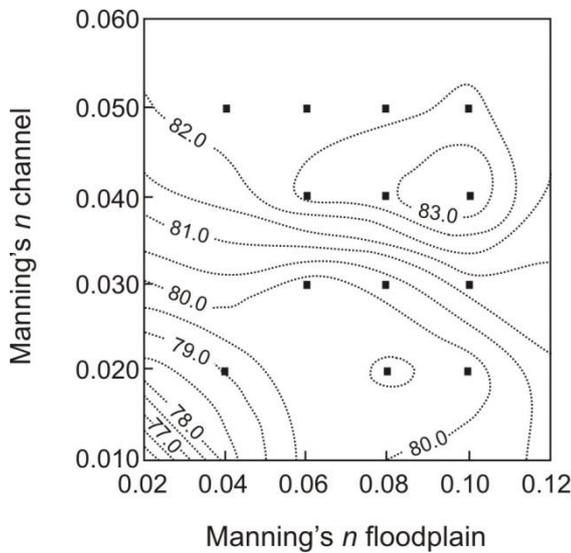


Figure 2: Parameterisation of hydraulic roughness is an inexact science (Horritt, 2005). Calibration seeks parameter values that produce the best fit between model predictions and observations, in this case of floodplain inundation extent (redrawn from Horritt, 2005). Contours represent the fit between inundation extent predicted by a SWE model and that mapped from ERS-1 SAR data.

The dimensionality of the equation set may be further reduced as in the nodal 1D 'St Venant equations', implemented in full (e.g. Brunner, 2010) or reduced to the kinematic wave approximation (e.g. Bates and De Roo, 2000). Codes based on the St Venant equations are used extensively to investigate flood inundation over complex topographies (e.g. Nicholas and Walling, 1997).

Reduced-complexity codes for channel applications use a simplified treatment of the momentum flux. Initially this involved replacing the momentum equation of the SWE by a steady, uniform flow approximation, with the goal of improving computational efficiency, often at the expense of process realism (see the review by Nicholas and Quine, 2007). Thus, one of the key challenges of reduced-complexity modelling lay in incorporating regional proxies of momentum effects to refine routing rules that otherwise operated exclusively with local (adjacent cells) information. In this regard, examples of important advancements include reducing model sensitivity to grid resolution (discussed later) by calculating slope as a weighted function of local and upstream topographic gradients (Nicholas *et*

al., 2006), and simulating river meandering by transferring regional information on bend curvature to individual cells (Coulthard and Van De Wiel, 2006).

True fluvial geomorphic models go beyond hydrodynamics to include transfer of dissolved or particulate matter (a process-study), or couple sediment fluxes with the morphology and composition of the bounding environment to simulate change (a process-form study; 'morphodynamics'). These aspects are discussed next.

Sediment transport

Fluvial geomorphic models predict a sediment transport rate at each node or cell, typically as a function of the local transport capacity, using empirical relations based on shear stress or stream power (derived from the hydrodynamics model). It is possible to model the transport of individual particles over an irregular bed using 3D codes (Hardy, 2005), and this may become increasingly important in some gravel bed river applications. However, most present applications in gravel and sand-bed rivers focus on bulk transport of material and the associated morphological response.

Several sediment transport formulae are available (e.g. Meyer-Peter and Müller, 1948, Engelund and Hansen, 1967, and van Rijn, 1993). The choice of formula is often a matter of personal preference, and sensitivity to this choice should be assessed. The Meyer-Peter and Müller (1948) formula tends to be favoured for gravel transport (e.g. Paola, 1996, and Nicholas, 2000), while the Engelund and Hansen (1967) formula is commonly applied to model fine gravel and/or sand transport (e.g. Kleinhans *et al.*, 2008, Nicholas, 2013a, and van Dijk *et al.*, 2014). Partheniades (1965) is typically used to model cohesive sediment transport (e.g. Deltares, 2014). Sediment transport over a non-erodible surface (e.g. cohesive floodplain) may be modelled using the approach of Struiksmá (1999), which preserves the integrity of local transport capacity relations by applying a correction factor to reduce the transportable layer thickness of the bed (Mosselman, 2005).

When considering the direction of sediment transport it is important to note that secondary (helical) flow and gravity-driven

transverse bed slope effects cause deviations between the direction of bed and near-bed sediment transport and that of the flow (Mosselman, 2005). Not all models correct for these effects. Work by Schuurman *et al.* (2013) and Nicholas (2013a, b) showed that the inclusion of such corrections is necessary to simulate dynamic channel planform evolution: i) secondary circulation corrections are essential to generate high-sinuosity meanders; and, ii) deflection driven by local bed slopes (determined by relative sediment mobility – the ratio of particle fall velocity to shear velocity) controls bar development. In combination with the effects of secondary circulation, these factors control bar morphodynamics, the rate of conversion of bars to floodplain through aggradation (Asahi *et al.*, 2013) and vegetation colonisation, and ultimately channel dimensions. Significant uncertainty remains in the parameterisation of bed slope effects (Mosselman, 2005; Nicholas, 2013a).

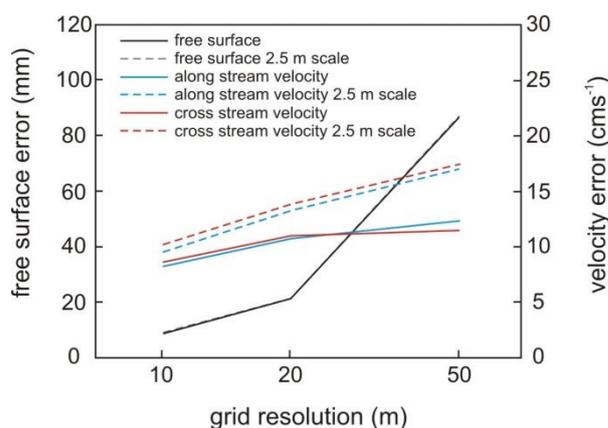
Morphodynamics

Fluvial geomorphic models update the morphology of the bounding environment at each node or cell based on the local volume change given by the balance between incoming and outgoing sediment. All forms of sediment mass balance equation are derived from work originally presented by Exner (1920, 1925), which schematises a sediment surface (e.g. hillslope, or channel bed) as a series of layers of known density, thickness, and volume (see Paola and Voller, 2005, for a full review and derivation of different forms of the equations). Such schematisation allows changes in bed level to be determined based on the flux at each control volume (e.g. cell), driven *inter alia* by the rate of gain or loss of mass within the flow, and horizontal divergence of particle flux within the flow, which are in turn driven by properties of the flow and local morphology (Paola and Voller, 2005).

Spatial domain and sensitivity

Different numerical approaches variously introduce false behaviour in hydraulics (and hence transport and/or morphological change) that is an artefact of the mathematics at play. The key note is that these artefacts may originate from or be expounded by certain aspects of the virtual

environment. One example is numerical diffusion in RANS codes that can be a problem where the grid resolution is low or the grid is poorly aligned with the primary flow direction (Patankar, 1980). For RC codes, grid-dependence of solutions may be unavoidable since the grid structure is an integral part of the process parameterisation (Nicholas, 2005). It is therefore important to understand what the potential sensitivities are for the code being used, and to work to quantify the effects of these sensitivities on model results.



*Figure 3: Testing for the effect of grid resolution on SWE model error (agreement with a high-resolution benchmark simulation, redrawn from Horritt *et al.*, 2006). The increase in error with grid cell size is due to the inability of models at lower resolution to resolve hydraulic structures such as recirculation zones, and to represent the complex domain boundary features that generate these structures (Horritt *et al.*, 2006).*

In 2D and 3D codes, spatial discretisation of the governing equations provides a framework to include the typically complex topography of the natural channel environment (Hardy, 2008). Grid resolution therefore directly influences representation of the river environment, and sensitivity of model solutions to grid resolution (Hardy *et al.*, 1999; Horritt *et al.*, 2006; Figure 3) must be tested and reported (see Lane *et al.*, 2005). Since variation in elevation determines streamwise and across-stream slope terms that are fundamental to the energy distribution within fluvial environments, the importance of an accurate and high-density elevation or bathymetric survey cannot be overstated in cases where the aim is to

investigate complex 3D flow structures (Horritt *et al.*, 2006; Hardy, 2008), to calibrate turbulence parameters in 2D hydrodynamic models (Williams *et al.*, 2013), or to predict flow paths and sedimentation processes over complex floodplain topography (Nicholas and Mitchell, 2003).

Ideally, the point density of a survey would match the grid resolution such that each cell has a measured elevation value. This is achievable in the case of LiDAR, terrestrial laser scanner, multibeam echosounder, or other continuous data. Otherwise, elevation data will comprise a series of points surveyed using ad-hoc, cross-section and/or morphologically-based approaches. These discontinuous data are interpolated onto the grid, thereby introducing an element of uncertainty. Confidence in a morphological survey can be enhanced by understanding the effect of different survey strategies and interpolation methods (see Heritage *et al.*, 2009).

Although an accurate initial morphology is typically prized in hydrodynamics applications (especially in 3D models), in the case of morphodynamic models that generate morphology, initial boundary conditions are commonly set using generalised values of key variables (slope, particle size distribution, discharge regime) that are broadly representative of a natural analogue, rather than using a specific set of highly accurate, high resolution field data (e.g. Kleinhans *et al.*, 2008). This is justified because the process representation of a model is unlikely to be consistent with processes operative at a field site (Nicholas, 2005). One approach is to set an initial morphology comprising a plane bed with a constant slope and small white noise elevation perturbations (Nicholas, 2013a).

In morphodynamic models it is important that the grid is able to accommodate the dynamics of landforms under investigation such that model output may be evaluated through comparison of properties of elevation change or planform change of modelled and measured environments. In the work of Nicholas (2013a, b) for example, the combination of a simple grid structure that can accommodate both gradual (lateral migration) and abrupt (cutoff, avulsion) channel movements, with a bank erosion

model that preserves bank height, is considered critical to the ability of models to reproduce channels that are similar in form and behaviour to natural analogues.

Temporal domain and sensitivity

Determining an optimum computational time step for a model requires a balance between preservation of numerical stability (improved by decreasing the time step) and computational efficiency (improved by increasing the time step). In codes based on the Navier-Stokes equations, the largest possible time step ensuring preservation of numerical stability may be found using an approach that ensures that hydrodynamic wave propagation does not progress beyond one cell per time step (CFL, Courant-Friedrichs-Lewy condition or number, Courant *et al.*, 1928). The CFL number is determined by model hydrodynamic properties and grid cell dimensions, and most code user manuals discuss its derivation in the context of the numerical procedure applied (e.g. Deltares, 2014). Numerical stability is ensured by scaling the time step according to the CFL condition. Instability is indicated by 'chequerboard' oscillations in hydrodynamic properties over a grid or lattice in flow simulations (Bates *et al.*, 2010; Coulthard *et al.*, 2013).

In lattice-network codes, numerical instability may be alleviated using a flow limiter (Bates and De Roo, 2000), which results in poor representation of the flow dynamics (Bates *et al.*, 2010), or adaptive time-stepping (Hunter *et al.*, 2005), which greatly increases the computational demand, especially for the high-resolution grids required in some applications. Bates *et al.* (2010) advanced the approach of Hunter *et al.* (2005) through incorporating a simple treatment of inertia based on analysis of the St. Venant equations, which allows time steps that scale linearly with the grid cell size according to the CFL condition (reviewed by Coulthard *et al.*, 2013).

Some SWE codes use different time steps for hydrodynamic and morphodynamic development by applying a morphological acceleration or scaling factor to allow simulation of long term morphological change (e.g. Lesser *et al.*, 2004; Nicholas, 2013a). Acceleration of morphological change is

achieved by multiplying sediment fluxes at each grid cell and time step by a constant factor, effectively increasing the morphodynamic time step relative to the hydrodynamic one. The selection of a suitable acceleration factor should be based on sensitivity tests (Lesser *et al.*, 2004); for example, Nicholas (2013a) showed that factors in the range 25 to 200 did not yield systematic variation in morphometric attributes of the simulated morphology (e.g. channel geometry, bar dimensions, number of branches).

Inlet/Outlet boundary conditions

Boundary conditions imposed at the inlet and outlet of the computational domain fundamentally influence simulation outcomes, and where a field analogue is considered it is the boundary conditions that define the field data needed to set up simulations (see Lane *et al.*, 1999 and Ingham and Ma, 2005 for full reviews). At the downstream (outlet) boundary a water level is specified that is either fixed through time (for steady flow simulations) or varies in relation to discharge (for unsteady flow simulations).

The main upstream (inlet) boundary requirement is a discharge rate and/or velocity profile (for 3D codes) – if the latter is required but could not be determined in the field it is acceptable to use a uniform velocity distribution (e.g. Milan, 2013) provided that the inlet is located far upstream of the region under investigation to allow development of a full flow field (Ingham and Ma, 2005). This is good practice for morphodynamic simulations as well, as morphological change at the inlet is representative of the immediate inlet flow/sediment feed rather than an outcome of channel process-form interaction.

Another method of improving the realism of inlet conditions involves the incorporation of an inlet bed perturbation comprising a rocking plane that mimics the slow migration of alternate bars (Nicholas, 2013a). The key point is that boundary perturbations may be needed in morphodynamic models to mimic disturbances that are central to the behaviour of natural systems. Morphodynamic disturbances typically propagate downstream in rivers (although upstream propagation also occurs), and in 1D meander migration models, for example, omitting inlet

perturbations may lead to channel straightening downstream of the inlet over time (Zolezzi and Seminara, 2001).

Confirmation

While hydrodynamic model attributes (e.g. depth, depth-averaged velocity) can be calibrated and confirmed according to their agreement with point or cross-sectional field measurements that are relatively easy to define at discrete points in space and time, morphodynamic model attributes are more difficult to define and thus morphodynamic model confirmation is more challenging (Bras *et al.*, 2003). Confirmation in the latter case typically involves quantifying metrics that describe the morphology and dynamics of the virtual river, and comparing these with measurements from imagery, fieldwork, laboratory experiments or results from other models that make different simplifying assumptions. Metrics used by Schuurman *et al.* (2013) to evaluate models of large braided sand-bed rivers included bed levels, braiding index, active channel width, bar length and bar shape. The focus of confirmation in morphodynamic modelling is not on reproducing an exact field analogue morphology (a perfect overlay of modelled and field forms), but on mimicking the quantifiable structure and dynamics of the natural environment.

The choice of when (in simulation time) to extract data for confirmation purposes tends to vary for hydrodynamic and morphodynamic models. Hydrodynamics simulations are typically run for relatively short timescales over which it is possible to compute an equilibrium solution for steady flow simulations ('convergence' for the specified boundary and initial conditions; Lane *et al.*, 2005), and over which unsteady flow dynamics can be compared with field measurements. Morphodynamics problems operate at longer timescales where it is very difficult to judge the 'end-point' of a simulation, even for steady flow input, due to the inherent dynamics of process-form feedbacks.

One approach is to define an 'equilibrium morphology' based on the stability of a feature (rate of change in form or process) over some multiple of the 'morphological timescale' (Miori *et al.*, 2006; Edmonds and

Slingerland, 2008). For example, Edmonds and Slingerland (2008) consider a bifurcation to be in equilibrium if there is active sediment transport in all reaches and the change in discharge ratio (partitioning of flow between the bifurcates) through time varies by less than 1% around the equilibrium value for at least 15 multiples of non-dimensional time (time elapsed over morphological time, where one unit of morphological time is determined by the time taken to transport an amount of sediment needed to fill one channel cross section).

Limitations of virtual rivers

The equations used to express the physical laws governing fluid flow are complex and cannot be solved analytically – it is only possible to estimate values of flow characteristics at discrete points in space and time (Kleinhans, 2010). Equations for sediment transport are arguably more complex, requiring several empirical closure relationships (Mosselman, 2005). Various degrees of simplification of these equations lead to questions about whether a model i) solves the correct equations, and ii) solves those equations correctly (Nicholas, 2005). Most model users are isolated from these ‘code development issues’ through the pull of a user-friendly graphical user interface, and the push of rather more obscure code validation documents. Some attention to the latter is advised (Oreskes *et al.*, 1994), as an understanding of key simplifying assumptions and sensitivities is important when interpreting model results. This has been a chief impediment to broad disciplinary engagement with modelling, but the rise of ‘open science’ initiatives (e.g. the iRIC Project, <http://i-ric.org/en/>, and CSDMS, <http://csdms.colorado.edu>), and improved access to information and support through e-resources and online forums make this an exciting time to be exploring the use of numerical models in fluvial geomorphology.

Further limitations of numerical modelling are discussed by Kleinhans (2010), and summarised hereafter. Simulating complex processes such as floodplain development with high-dimensionality codes still requires more computational power than is available to most researchers, but this is likely to change. Reduced complexity codes offer an alternative, but it is difficult to determine

whether they correctly reproduce the characteristics and dynamics of natural prototypes for the correct (physical) reasons. Advancements are being made that are especially relevant to the simulation of landscape-scale phenomena (e.g. Coulthard *et al.*, 2013).

Morphodynamic models are not well-suited to simulating the exact details and dynamics of a natural prototype, and point-by-point comparison is neither feasible nor desirable given that: i) it is not possible to specify initial and boundary conditions in sufficient detail with available measurement techniques; ii) uncertainty remains in the numerical representation of important physical processes (e.g. transverse bed slope effects); and, iii) all codes neglect important processes that may lead to minor differences in ‘real’ and ‘virtual’ form (e.g. sediment sorting). Kleinhans (2010) therefore emphasises the importance of a comparison of general characteristics such as bar morphometrics, channel planform geometry or channel network structure.

Acknowledgements

The author thanks two anonymous reviewers for constructive feedback that helped to refine this chapter.

References

- Asahi K, Shimizu Y, Nelson J, Parker G. 2013. Numerical simulation of river meandering with self-evolving banks. *Journal of Geophysical Research: Earth Surface* **118**: 1–22, DOI:10.1002/jgrf.20150.
- Bates PD, De Roo APJ. 2000. A simple raster-based model for flood inundation simulation. *Journal of Hydrology* **236**: 54–77.
- Bates PD, Horritt MS, Fewtrell TJ. 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology* **387**: 33–45, DOI: 10.1016/j.jhydrol.2010.03.027.
- Bolla Pittaluga M, Repetto R, Tubino M. 2003. Channel bifurcation in braided rivers: Equilibrium configurations and stability. *Water Resources Research* **39**, 1046, DOI:10.1029/2001WR001112.

- Bras RL, Tucker GE, Teles V. 2003. Six Myths About Mathematical Modeling in Geomorphology. In: Wilcock PR, Iverson RM. (Eds.) *Prediction in Geomorphology*, American Geophysical Union, Washington, D. C., DOI: 10.1029/135GM06.
- Brunner GW. 2010. HEC-RAS River Analysis System Hydraulic Reference Manual. USACE Hydraulic Engineering Center, Davis California, 417pp.
- Chin A. 2003. The geomorphic significance of step-pools in mountain streams. *Geomorphology* **55**: 125–137, DOI: 10.1016/S0169-555X(03)00136-3.
- Constantine JA, Dunne T, Piégay H, Kondolf GM. 2010. Controls on the alluviation of oxbow lakes by bed-material load along the Sacramento River, California. *Sedimentology* **57**: 389–407. DOI:10.1111/j.1365-3091.2009.01084.x.
- Coulthard TJ, van de Wiel MJ. 2006. A cellular model of river meandering. *Earth Surface Processes and Landforms* **31**: 123–132, DOI:10.1002/esp.1315.
- Coulthard TJ, Neal JC, Bates PD, Ramirez J, de Almeida GAM, Hancock GR. 2013. Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution. *Earth Surface Processes and Landforms* **38**: 1897–1906, DOI:10.1002/esp.3478.
- Courant R, Friedrichs K, Lewy H. 1928. On the partial difference equations of mathematical physics. *Mathematische Annalen* **100**: 32–74 (translated for IBM Journal).
- Deltares. 2014. Delft3D-Flow. Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. User Manual, Version 3.15.34158.
- Edmonds DA, Slingerland RL. 2008. Stability of delta distributary networks and their bifurcations. *Water Resources Research* **44**, W09426, DOI:10.1029/2008WR006992.
- Engelund F, Hansen E. 1967. A monograph on sediment transport in alluvial streams. Teknisk Forlag, Copenhagen.
- Exner FM. 1925. Über die wechselwirkung zwischen wasser und geschiebe in flüssen, *Akad. Wiss. Wien Math. Naturwiss. Klasse* **134(2a)**: 165–204.
- Exner FM. 1920. Zur physik der dünen, *Akad. Wiss. Wien Math. Naturwiss. Klasse* **129(2a)**: 929–952.
- Hardy RJ. 2008. Geomorphology Fluid Flow Modelling: Can Fluvial Flow Only Be Modelled Using a Three-Dimensional Approach? *Geography Compass* **2**: 215–234.
- Hardy RJ. 2005. Modelling granular sediment transport over water-worked gravels. *Earth Surface Processes and Landforms* **30**: 1069–1076, DOI:10.1002/esp.1277.
- Hardy RJ, Lane SN, Lawless MR, Best JL, Elliot L, Ingham DB. 2005. Development and testing of a numerical code for the treatment of complex river channel topography in three-dimensional CFD models with structured grids. *Journal of Hydraulic Research* **43**: 468–480, DOI:10.1080/00221680509500145.
- Hardy RJ, Bates PD, Anderson MG. 1999. The importance of spatial resolution in hydraulic models for floodplain environments. *Journal of Hydrology* **216**: 124–136.
- Heritage GL, Milan DJ, Large, AR, Fuller IC. 2009. Influence of survey strategy and interpolation model on DEM quality. *Geomorphology* **112**: 334–344.
- Horritt MS. 2005. Parameterisation, validation and uncertainty analysis of CFD models of fluvial and flood hydraulics in the natural environment. In: Bates PD, Lane SN, Ferguson RI (Eds.) *Computational Fluid Dynamics: Applications in Environmental Hydraulics*. Wiley and Sons: Chichester, UK, pp. 193–213.
- Horritt MS, Bates PD, Mattinson MJ. 2006. Effects of mesh resolution and topographic representation in 2D finite volume models of shallow water fluvial flow. *Journal of Hydrology* **329**: 306–314.
- Hunter NM, Horritt MS, Bates PD, Wilson MD, Werner MGF. 2005. An adaptive time step solution for raster-based storage cell modelling of floodplain inundation. *Advances in Water Resources* **28**: 975–991, DOI:10.1016/j.advwatres.2005.03.007.
- Ikeda S, Parker G, Sawai K. 1981. Bend theory of river meanders. Part 1. Linear development. *Journal of Fluid Mechanics* **112**: 363–377.
- Ingham DB, Ma L. 2005. Fundamental equations for CFD in river flow simulations. In: Bates PD, Lane SN, Ferguson RI. (Eds.)

- Computational Fluid Dynamics: Applications in Environmental Hydraulics*. Wiley and Sons: Chichester, UK, pp. 19–49.
- Keylock CJ, Hardy RJ, Parsons DR, Ferguson RI, Lane SN, Richards KS. 2005. The theoretical foundations and potential for large-eddy simulation (LES) in fluvial geomorphic and sedimentological research. *Earth Science Reviews* **71**: 271–304.
- Kleinhans MG. 2010. Sorting out river channel patterns. *Progress in Physical Geography* **34**: 287–326, DOI:10.1177/0309133310365300.
- Kleinhans MG, Jagers HR, Mosselman E, Sloff CJ. 2008. Bifurcation dynamics and avulsion duration in meandering rivers by one-dimensional and three-dimensional models. *Water Resources Research* **44**, W08454, DOI:10.1029/2007WR005912.
- Lane SN. 1998. Hydraulic modelling in hydrology and geomorphology: A review of high resolution approaches. *Hydrological Processes* **12**: 1131–1150.
- Lane SN, Hardy RJ, Ferguson RI, Parsons DR. 2005. A framework for model verification and validation of CFD schemes in natural open channel flows. In: Bates PD, Lane SN, Ferguson RI. (Eds.) *Computational Fluid Dynamics: Applications in Environmental Hydraulics*. Wiley and Sons: Chichester, UK, pp. 169–192.
- Lane SN, Bradbrook KF, Richards KS, Biron PA, Roy AG. 1999. The application of computational fluid dynamics to natural river channels: three-dimensional versus two-dimensional approaches. *Geomorphology* **29**: 1–20.
- Lesser G, Roelvink J, van Kester J, Stelling G. 2004. Development and validation of a three-dimensional morphological model, *Coastal Engineering* **51**: 883–915.
- Meyer-Peter E, Müller R. 1948. Formulas for bed-load transport. Paper No. 2, Proceedings of the 2nd Congress, IAHR, Stockholm, pp. 39–64.
- Milan DJ. 2013. Sediment routing hypothesis for pool-riffle maintenance. *Earth Surface Processes and Landforms* **38**: 1623–1641, DOI:10.1002/esp.3395.
- Miori S, Repetto R, Tubino M. 2006. A one-dimensional model of bifurcations in gravel bed channels with erodible banks, *Water Resources Research* **42**, W11413, DOI:10.1029/2006WR004863.
- Mosselman E. 2005. Basic equations for sediment transport in CFD for fluvial morphodynamics. In: Bates PD, Lane SN, Ferguson RI. (Eds.) *Computational Fluid Dynamics: Applications in Environmental Hydraulics*. Wiley and Sons: Chichester, UK, pp. 71–89.
- Nelson JM, McDonald RR. 1995. Mechanics and modeling of flow and bed evolution in lateral separation eddies. US Geological Survey Report, Grand Canyon Monitoring and Research Center, Flagstaff, Arizona.
- Nicholas AP. 2013a. Modelling the continuum of river channel patterns. *Earth Surface Processes and Landforms* **38**: 1187–1196, DOI:10.1002/esp.3431.
- Nicholas AP 2013b. Morphodynamic diversity of the world's largest rivers. *Geology* **41**: 475–478.
- Nicholas AP. 2005. Cellular modelling in fluvial geomorphology. *Earth Surface Processes and Landforms* **30**: 645–649, DOI:10.1002/esp.1231.
- Nicholas AP. 2000. Modelling bedload yield in braided gravel bed rivers. *Geomorphology* **36**: 89–106.
- Nicholas AP, Quine TA. 2007. Crossing the divide: Representation of channels and processes in reduced-complexity river models at reach and landscape scales. *Geomorphology* **90**: 318–339, DOI:10.1016/j.geomorph.2006.10.026.
- Nicholas AP, Mitchell CA. 2003. Numerical simulation of overbank processes in topographically complex floodplain environments. *Hydrological Processes* **17**: 727–746.
- Nicholas AP, Walling DE. 1997. Modelling flood hydraulics and overbank deposition on river floodplains. *Earth Surface Processes and Landforms* **22**: 59–77.
- Nicholas AP, Thomas R, Quine TA. 2006. Cellular modelling of braided river form and process. In: Sambrook Smith GH, Best JL, Bristow CS, Petts GE. (Eds.) *Braided Rivers: Process, Deposits, Ecology and Management*. Special Publication International Association of Sedimentologists **36**, pp. 137–151.

- Oreskes N, Shrader-Frechette K, Belitz K. 1994. Verification, validation and confirmation of numerical models in the earth sciences. *Science* **264**: 641–6.
- Paola C. 1996. Incoherent structure: Turbulence as a metaphor for stream braiding. In: Ashworth PJ, Bennett SJ, Best JL, McLelland SJ. (Eds.) *Coherent Flow Structures in Open Channels*, Wiley and Sons: Chichester, UK, pp. 705–723.
- Paola C, Voller VR. 2005. A generalized Exner equation for sediment mass balance. *Journal of Geophysical Research* **110**, F04014, DOI:10.1029/2004JF000274.
- Parker G, Shimizu Y, Wilkerson GV, Eke EC, Abad JD, Lauer JW, Paola C, Dietrich WE, Voller VR. 2011. A new framework for modelling the migration of meandering rivers. *Earth Surface Processes and Landforms* **36**: 70–86, DOI:10.1002/esp.2113.
- Partheniades E. 1965. “Erosion and Deposition of Cohesive Soils.” *Journal of the Hydraulics Division, ASCE* **91 (HY 1)**: 105–139.
- Patankar SV. 1980. *Numerical Heat Transfer and Fluid Flow*. Hemisphere Publishing Corporation.
- Schuurman F, Marra WA, Kleinhans MG. 2013. Physics-based modelling of large braided sand-bed rivers: Bar pattern formation, dynamics, and sensitivity. *Journal of Geophysical Research: Earth Surface* **118**: 2509–2527, DOI:10.1002/2013JF002896.
- Slingerland R, Smith ND. 1998. Necessary conditions for a meandering-river avulsion. *Geology* **26**: 435–438.
- Struiksmā N. 1999. *Mathematical modelling of bedload transport over non-erodible layers*. Proceedings of IAHR Symposium on River, Coastal and Estuarine Morphodynamics, Genova, 6-10 September Vol 1, pp. 89–98.
- van Dijk WM, Schuurman F, van de Lageweg WI, Kleinhans MG. 2014. Bifurcation instability and chute cutoff development in meandering gravel-bed rivers. *Geomorphology* **213**: 277–291, DOI:10.1016/j.geomorph.2014.01.018.
- van Rijn LC. 1993. *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas*. Aqua Publications, Amsterdam.
- Williams RD, Brasington J, Hicks M, Measures R, Rennie CD, Vericat D. 2013. Hydraulic validation of two-dimensional simulations of braided river flow with spatially continuous aDCP data. *Water Resources Research* **49**: 5183–5205, DOI:10.1002/wrcr.20391.
- Wright NG, Baker CJ. 2004. Environmental Applications of Computational Fluid Dynamics. In: Wainwright J, Mulligan M. (Eds.) *Environmental Modelling: Finding Simplicity in Complexity*. Wiley and Sons: Chichester, UK, pp. 335–348.
- Zolezzi G, Seminara G. 2001. Downstream and upstream influence in river meandering. Part 1. General theory and application to overdeepening. *Journal of Fluid Mechanics* **438**: 183–211.

5.6.5 Modelling Geomorphic Systems: Glacial

Ann V. Rowan¹

¹Basin Studies and Petroleum Geoscience, School of Earth, Atmospheric and Environmental Sciences, University of Manchester, Manchester, M13 9PL, UK

Email: ann.rowan@manchester.ac.uk



The cryosphere contains a rich archive of the climate record, and is an important part of the global hydrological cycle. The investigation of the glacial behaviour informs our understanding of the forces that have shaped landscapes in glaciated terrains, and is key to developing understanding of past and future climate change and global climate teleconnections. Numerical modelling of glacial systems allows us to understand quantitatively the processes that drive glaciers. Three major model types exist: those that use ice extents to understand climate changes; those that investigate the forces that control ice dynamics; and those that investigate the erosional consequences of glaciation. Glacier models can be applied with a range of spatial extents, from individual cirque glaciers (e.g. Murray and Locke, 1989; Ballantyne, 2002; Coleman et al., 2009; Hughes, 2009) to continental ice sheets (e.g. Marshall and Clarke, 1999; Hubbard et al., 2005). The basis of any glacier model is a calculation of mass balance; the relationship between ice accumulation and ablation at a given point in time, under the current climate conditions. Mass balance controls glacier dynamics, which respond to processes operating on timescales of different orders of magnitude (e.g. climate change, tectonic uplift). Increasingly complex models require more variables to be specified as inputs, and so are more difficult to apply accurately; the model builder must decide which variables to exclude. However, if the input parameters are well constrained, results from complex models should be more robust. Numerical models can be mathematically 1- or 2-D, equivalent to what is more commonly described as spatially 2-D (e.g. along a line of section) or 3-D (e.g. a map view extent), which must include ice thickness. In this section, types of glacier model that can be applied to a range of different aspects of the cryosphere are discussed, alongside methodological concerns in applying different models, and important considerations in a modelling project. However, modelling studies of smaller glaciers that are confined by topography (i.e. valley glaciers) are the focus. For a starting point for models describing ice sheets, ice shelves and marine ice margin processes, glacier hydrology and isostatic adjustment, the reader is directed towards Petrenko and Whitworth (2002), Jamieson et al. (2008), Cuffey and Patterson (2010), Benn and Evans (2010) and references therein.

KEYWORDS: Numerical modelling; glacial; mass balance; energy-balance; glacier dynamics

Glacier reconstruction

Before a glacier model can be applied, a reconstruction of the glacier of interest is necessary to define boundary conditions such as topography, temperature change and precipitation amount. Model parameters, such as albedo, the values for the exponent and the constant to be used in Glen's Flow Law (Glen, 1955), or the regional lapse rate (typically a decrease of 4–8°C per 1000 m elevation gained), also need to be defined.

Reconstruction can be based on both empirical field evidence of glacier form (e.g. Ballantyne, 2002; Kelly et al., 2004), and theoretical understanding of iceflow relationships (e.g. Benn and Hulton, 2010).

A simple tool for glacier reconstruction is the accumulation area ratio (AAR) which assumes that the accumulation area is ~65% of a glacier (a value derived from studies of modern glaciers; Meierding, 1982) to constrain the equilibrium line altitude (ELA) for a past glacier, by reconstruction from the

position of terminal moraines. However, the AAR is not suitable for some smaller glaciers, and is increasingly being replaced by the more sophisticated area-altitude balance ratio (AABR) method, which considers glacier hypsometry to calculate ELA based on the location of the zero point in the glacier energy balance (Rea, 2009).

Types of glacier model

The applications of glacier models are hugely varied, and the type, as well as the spatial and temporal scale, of model used will depend on what is being investigated. Glacier models are commonly employed to investigate local ice extent (i.e. cirque and valley glaciers) and behaviour for a given period (Carr et al., 2010), possibly in combination with a geochronological study. Glacier models are also used to investigate the relationship of large ice volumes (i.e. ice caps and ice sheets) to regional/hemispheric climate change (Pollard, 2010) and ocean circulation patterns (e.g. Schmittner et al., 2002). Possible outcomes of glacier modelling discussed here include: a regional equilibrium line altitude (ELA) for a particular climate period; a recreation of the ice that generated a particular geomorphologic record; the conditions required for glacier steady state; and definition of the important drivers of glacial erosion and sediment production.

Glacier models can be implemented using a variety of computer architectures and software, and an important consideration should be the amount of computational time that will be required to produce the results needed. A complex model running on an office-specification machine may take an undesirably long time to produce the looked-for data, whereas some problems may be solved simply by using Excel™ (e.g. Brock and Arnold, 2000; Benn and Hulton, 2010). Different model types commonly applied to glacial systems are classified and described in this section. Glacier models may be classified according to the level of complexity in their calculations, as this defines the inputs needed to create a model and the output style generated.

Mass balance models

The positive degree day model (PDD;

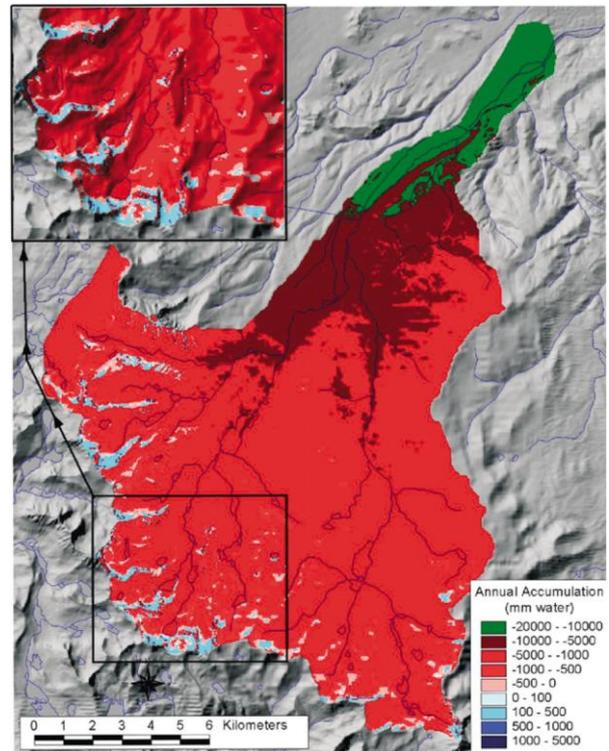


Figure 1. Result of a mass balance calculation for the Bishop Creek basin, Sierra Nevada, USA under modern climate conditions overlaid on shaded relief image of the study area. Blue indicates areas of positive net accumulation. Red indicates areas of negative net accumulation (ablation). From Plummer and Phillips (2003).

Braithwaite, 1995; Braithwaite and Raper, 2007) calculates the amount of melting that occurs during the melt season, determined by when air temperatures at the glacier surface are above a threshold temperature, usually 0–2°C.

Mass balance models investigate the change in mass of a glacier and the distribution of these changes in space and time (Cuffey and Paterson, 2010) vertically relative to the glacier surface (Oerlemans, 2008, 2010). Mass balance describes when the relationship between accumulation of snow and ablation by melting and sublimation is equal to zero. A mass balance calculation can be a result in itself (Fig. 1), or used to investigate glacier behaviour as an input to an ice dynamics model. Changing the climate conditions in the mass balance model can be used to investigate the properties of palaeoglaciers.

Energy balance models calculate explicitly the energy (heat) fluxes at the glacier surface that control snow melting and sublimation

that will affect mass balance (Brock and Arnold, 2000; Klok and Oerlemans, 2002; Braithwaite and Raper, 2007). The most important factor in ice surface energy flux is longwave radiation (Ohmura, 2001). Atmospheric radiative flux is strongly dependent on local topography and the annual position of the sun (Coleman et al., 2009), which varies over glacial timescales. Energy balance calculations can be used to analyse glacier sensitivity (Hubbard, 1999), which may have been very different for the same glacier throughout its history, and to investigate the effect of basin geometries on glacier behaviour. Energy balance calculations also feed into models describing ice melting that are used to understand how and when water is released from the glacial system (Hock, 2005).

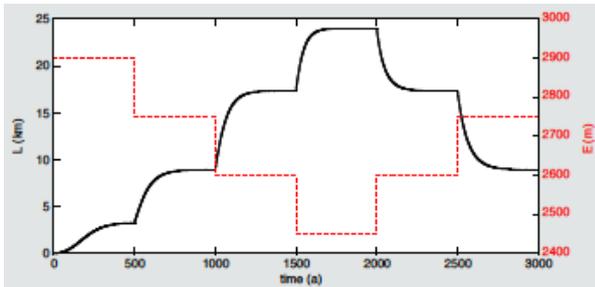


Figure 2. The response of a simple time-dependent glacier model to stepwise forcing. The steps shown by the dashed red line represent changes in ELA (E) of 150 m. Black line shows glacier length. From Oerlemans (2008).

Glacier dynamics models

The mass balance models discussed above consider glaciers in equilibrium, which is useful to understand glacier sensitivity and model consistency, but does not consider the often transient response of glaciers to external forcing (Oerlemans, 2008). Glacier dynamics models allow the investigation of time-dependent changes in glacier properties and behaviour (Oerlemans, 2008) such as change in glacier length with change in mass balance (Fig. 2). These models define the physical laws that control ice flow, and use these relationships to calculate the velocities and direction of ice flow in a glacier or ice sheet (Rutt et al., 2009). Models of ice dynamics may be very simple (e.g. Carr et al., 2010) or very complex (e.g. Huybrechts, 1990).

Glacier dynamics models are typically concerned with warm-based or polythermal glaciers or ice shelves, which flow in a much more rapid and unpredictable manner than ice bodies that are frozen to their beds. These warm-based glaciers move downstream as a result of both internal deformation and basal sliding, whereas cold-based ice bodies, such as those found inland on Antarctica, advance much more slowly, often by deformation only. The most frequently used glacier dynamics models address ice sheet and sea level response to climate change, such as the parallel ice sheet model (PISM, Fig. 3; Bueller and Brown, 2009; Winkelmann et al., 2010) and Glimmer (Rutt et al., 2009). For valley glaciers, flowline models, or 2-D iceflow models such as the Plummer and Phillips (2003) glacier model, are often used (Fig. 4).

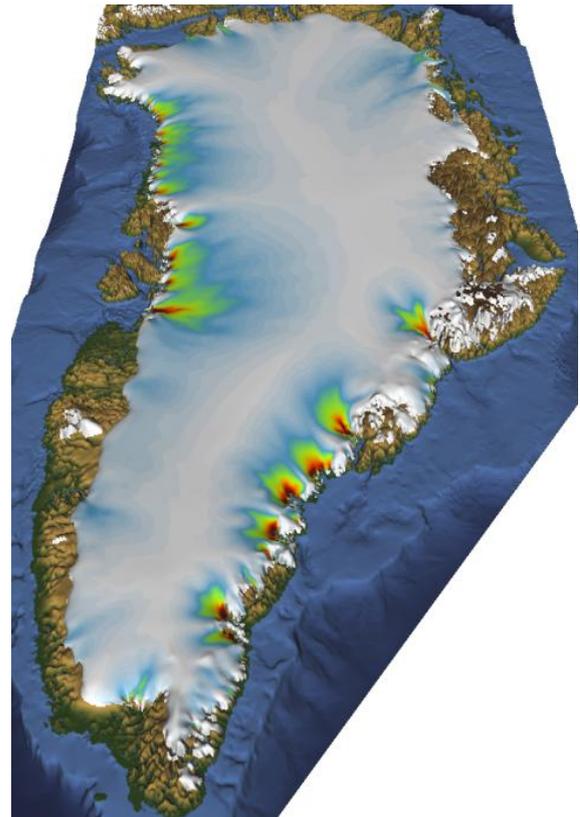


Figure 3. PISM model output for the Greenland ice sheet, showing ice surface elevation, coloured by ice thickness. From pism-docs.org.

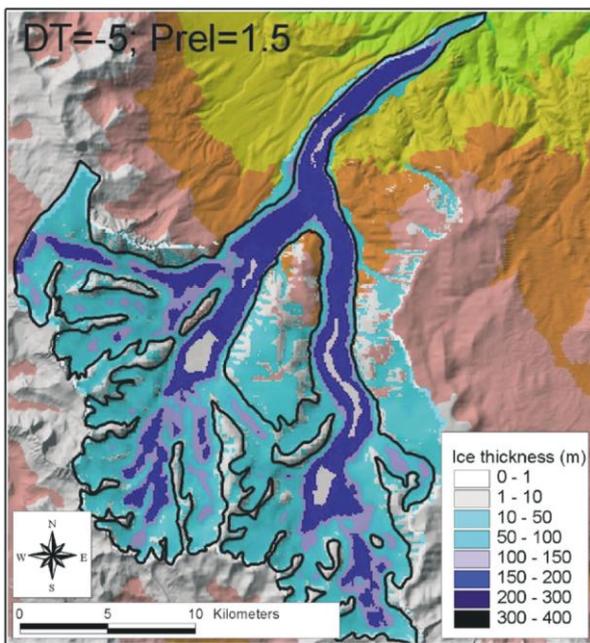


Figure 4. Iceflow model output for the Bishop Creek basin, Sierra Nevada, USA, produced by a 5°C temperature depression and a 50% increase in precipitation, equivalent to a Tioga (LGM) maximum glacier, based on field mapping of glacial trimlines and other glacial geomorphic features. From Plummer and Phillips (2003).

Flowline models (e.g. Pattyn, 1996; Oerlemans, 1997; MacGregor et al., 2000) are frequently used for valley glaciers and ice streams, to consider the transient changes that result from the forces acting along a longitudinal glacier profile (Fig. 5). Glen's Flow Law (Glen, 1955) shows that, at stresses important for normal glacier flow (50–150 kPa) the relation between shear stress change and the corresponding strain rate follows a power law (Cuffey and Patterson, 2010). Glacier dynamics models are frequently based on the shallow ice approximation (SIA; Hutter, 1983) where ice flow is driven simply by local gradients in ice surface elevation and ice thickness and all the resistance to flow is generated by the basal boundary to the ice, although development is needed to allow this model to respond well to variable bed topography (Egholm et al., 2011).

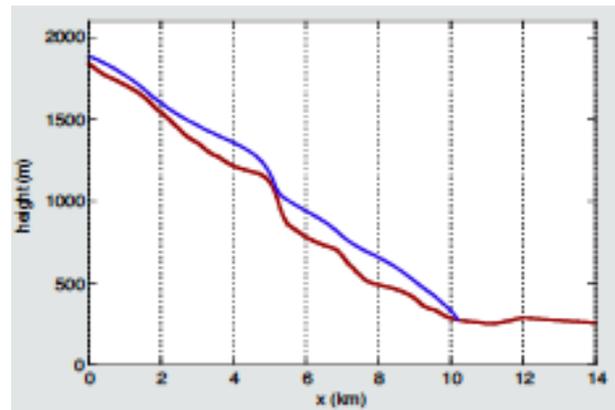


Figure 5. Flowline model of the Nigardsbreen Glacier, Norway. Blue line shows ice surface elevation relative to the glacier bed, shown by the red line, as a function of distance along a flowline (x). From Oerlemans (1997).

Glacial erosion and sediment transport models

Glacial erosion models (e.g., Braun et al., 1999; MacGregor et al., 2000; Tomkin and Roe, 2007; Tomkin, 2007, 2009; Herman and Braun, 2008; Egholm et al., 2010, 2011) investigate the complex relationships that control when and where glacial downcutting and headwall erosion occur (MacGregor et al., 2000), and the resulting landscape form (Brocklehurst and Whipple, 2007). Erosion models may represent either a valley long-profile (e.g. Brocklehurst and MacGregor, 2009), a valley cross-section (Harbor, 1992), a range cross-section (e.g. Benn and Hulton, 2010) or a swath of topography representing an entire range (Fig. 6; e.g. Egholm and Nielsen, 2010). Beneath temperate glaciers, erosion rate is usually treated as proportional to the sliding rate, but this is complicated by factors including water-pressure fluctuations and chemical dissolution (MacGregor et al., 2000). Glacier models that incorporate an erosion term usually simply define the amount of erosion as a function of the sliding rate, as this is important for quarrying and abrasion (e.g. MacGregor et al., 2000; Egholm et al., 2009) but does not take into account other erosion processes (Hooke, 1991; Alley et al., 2003).

Tectonic uplift has a crucial role to play in erosion models (Hubbard et al., 2005; Tomkin, 2007), as glaciers will respond to an increase in their elevation by becoming more erosive (Brocklehurst and Whipple, 2007;

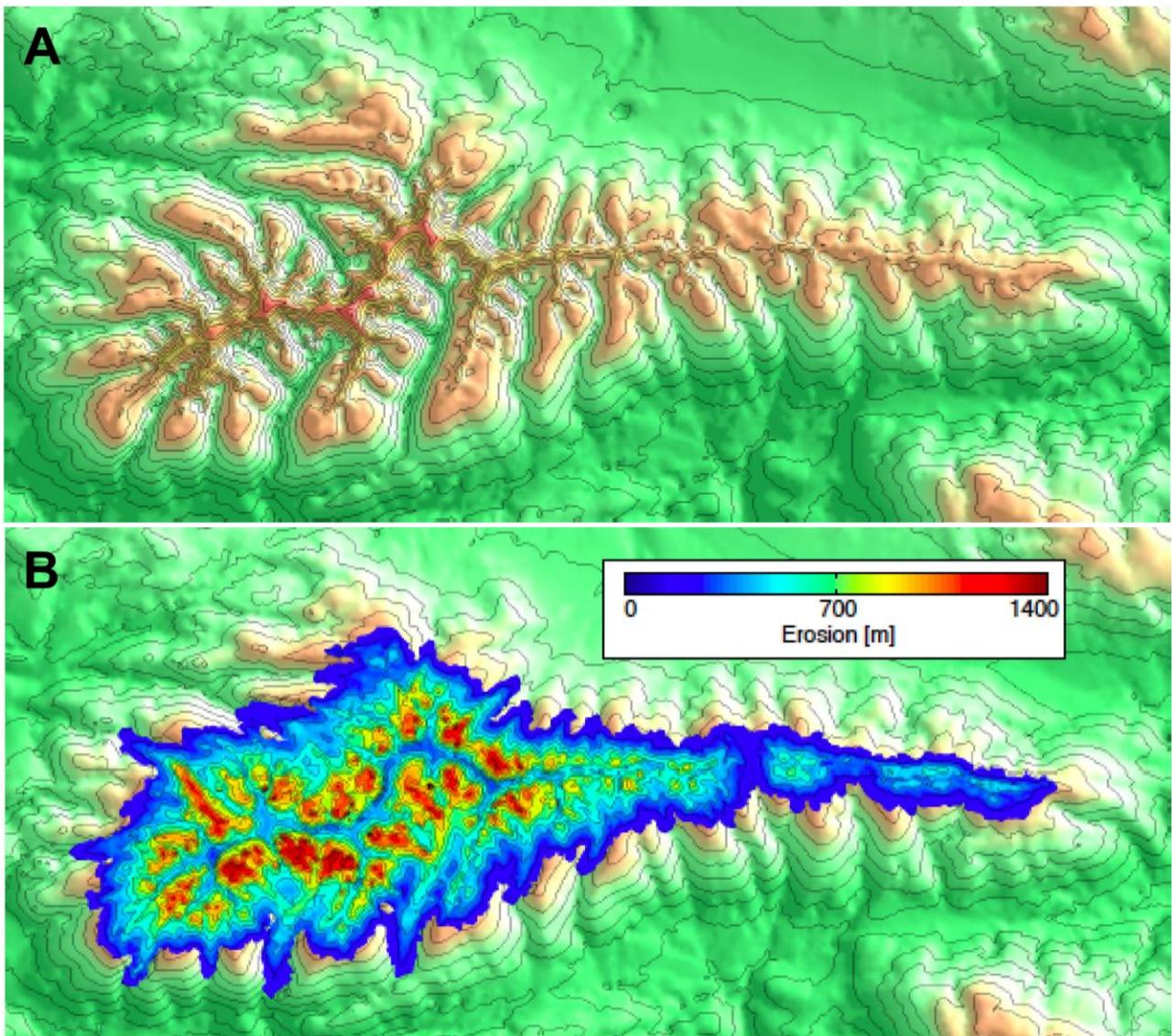


Figure 6. (A) The resulting topography after the Egholm et al. (2009) model has been applied to previously non-glaciated topography in the Sierra Nevada, Spain, for 500 ka. The postglacial DEM shows classical glacier erosion features including arêtes, cirques and hanging valleys. (B) The distribution and magnitude of erosion within the study area. Contour spacing is 200 m. From Egholm et al. (2009).

Egholm et al., 2010). Erosion is typically concentrated at the valley floor and so can lead to increased relief in glaciated landscapes, particularly when the effect of isostatic rebound is included (Molnar and England, 1990; Hubbard, 2006). However, as glacial erosion lowers the glacier bed, mass balance becomes less positive for the same climate condition and so the glacier retreats (Oerlemans, 1984; Whipple et al., 1999). These processes set up feedbacks that present challenges to glacier modelling.

Glacier sediment transport models (e.g., Dowdeswell and Siegert, 1999) consider how sediment is transported by ice and where and

when it is released from the glacial system. Sediment transport is a complex problem to parameterise (Ballantyne, 2002), but results can be usefully combined with a glacier erosion model to understand the controls of glacier dynamics on sediment flux, and so interpret the resulting sedimentological record in the context of climate change. Finally, predicting the occurrence and distribution of rock glaciers remains a considerable challenge to glacier modelling (e.g. Imhof, 1996; Brenning and Trombotto, 2006), in part due to the limited understanding of the processes that form rock glaciers (Haeberli et al., 2006).

Glacier model inputs

A glacier model will require a representation of the underlying topography. In the case of a model for a theoretical glacier this would require either a representative topography or flat surface as a starting point. If the modelling study considers the glaciation of a particular study area, then a representation of the real topography is required, usually as a digital elevation model (DEM). DEM resolution can be reduced from that of the raw data (typically 10-50 m cells) often without affecting the accuracy of model results (Marshall and Clarke, 1999; Evans et al., 2009). This has the advantage of decreasing model run-time. If modern ice or large volumes of sedimentary valley fill are present in the area covered by the DEM it may be desirable to create a DEM of the underlying bedrock surface (Jordan, 2009). However, this requires detailed knowledge of the thickness of the sedimentary fill, e.g. from boreholes or seismic surveying, to prevent introducing further error. It is therefore common in studies of glaciated regions to model glaciers using the DEM data as it stands. When modelling creates ice within the DEM this should be iteratively included in subsequent model runs.

DEMs typically occupy a regular grid of square cells, but models may also operate on irregular grid patterns (Fig. 7), which can increase model accuracy by removing the strong directional bias of orthogonal grid patterns, and allow decreased resolution where resolution is less important, such as unglaciated peaks (e.g. Alho and Aaltonen, 2008; Egholm and Nielsen, 2010). Cell density can then be increased at complex areas, such as glacier boundaries, to allow more detail to be captured (Egholm and Nielsen, 2010). Model grid resolution should be carefully chosen, as it can affect model output, such as the response time of glacier length to changing mass balance (Oerlemans, 2008).

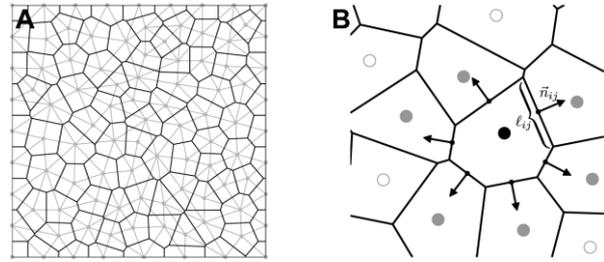


Figure 7. (A) Irregular model grid of Voronoi cells (black lines) discretised using Delaunay triangulation (grey lines) that can be used to improve model accuracy, particularly at ice body boundaries. (B) A Voronoi cell, centred on the black nodal point, Each edge of the cell has a length l_{ij} and an orientation described by the normal vector \vec{n}_{ij} . From Egholm and Nielsen (2010).

Temperature is one of the major factors driving changes in glacier mass balance (Anderson et al., 2010) and lapse rate varies with latitude (Syvitski et al., 2003). Rates of temperature change over the modelling period are also important, as variations in the duration and magnitude of climate change event are shown to be a primary control on glacier form (Brocklehurst and MacGregor, 2009). Precipitation may be more difficult to constrain than temperature if databases such as PRISM (Daly et al., 2002) are not available for the study area. Precipitation in alpine terrains is likely to be subject to pronounced orographic effects, causing spatially complex precipitation distributions strongly influenced by prevailing wind direction and local atmospheric conditions (e.g. Marshall and Clarke, 1999). Problems can arise in determining if precipitation falls as snow or rain (precipitation phase) and the effect of wind redistribution on solid-phase precipitation (e.g. Anders, 2008; Foster et al., 2010), which is also difficult to quantify (Gauer, 2001; Bernhardt et al., 2010).

Applications of glacier models

Some common and emerging aims when using different types of glacier model are:

- Palaeoclimate reconstructions to determine where ice occurred and the climate changes required to generate glaciers. This can be carried out using a mass and/or energy balance model, either by modelling glaciers for a particular

change in climate linked to a temperature proxy such as an ice core, or by modelling glaciers to fit the observed geomorphology (Fig. 4), to define the climate change needed to match this record (Boulton and Hagdorn, 2006).

- Ice extent reconstructions to define a particular climate event (e.g. Hubbard et al., 2005), often the Last Glacial Maximum (LGM) that is considered to be the last extreme climate change event.
- Calculation of glacier sensitivity; the change in mass balance with a certain climate change (e.g. Anderson and Mackintosh, 2006).
- Erosion models can be used to consider where and when glacial erosion occurs as climate changes (e.g. Egholm and Nielsen, 2010), and so understand the landscape response to climate change events such as the Middle Pleistocene Transition (e.g. Brocklehurst and MacGregor, 2009).
- To investigate glacier hydrology and generate estimates of discharge and sediment flux from the modelled glaciers, which can be used to drive downstream sediment transport models and link these results to climate change (e.g. Richards et al., 1996).
- To remotely investigate the geomorphology of other planets, such as Mars (e.g. Pelletier et al., 2010).

Models can be designed either as discrete or finite calculations depending on their purpose. Discrete element models make calculations based on the characteristics of individual points, so in the case of a mass balance model, the mass balance is calculated for each cell of the model domain, based on the properties (e.g. elevation, degree of shading, precipitation input) of that cell. Finite element and finite difference models calculate the relationships between cells of the model grid, so in the case of an iceflow model, these calculations would consider the velocity and direction of iceflow into and out of each cell from/to neighbouring cells, allowing detailed investigation of ice dynamics (Gudmundsson, 1999). Finite volume calculations offer an alternative to some of the problems commonly encountered in finite element calculations, but may be more difficult to implement

(Egholm and Nielsen, 2010).

Limitations of glacier modelling

The major limitations of many existing glacial models is that the majority of these rely on estimates of palaeoglacier properties based on assumptions about the characteristics of modern glaciers (Plummer and Phillips, 2003), and models must simplify the physical processes operating on an ice mass to describe these processes numerically (Benn and Evans, 2010). The use of assumptions about glacier properties is problematic in areas where no modern ice exists and modern ELA must also be estimated, as this introduces uncertainties into the results that are difficult to quantify. However, some models take a different approach, and reconstruct glaciers without assuming any prior knowledge of glacier form, using only local topographic and climate information as inputs (e.g. Plummer and Phillips, 2003). Uncertainty in glacier modelling exists due to the physics of ice flow being controlled by a complex set of relationships, and the wide range of processes that occur to different degrees in different glaciers. Inclusion of all these processes in a glacier model is limited by the amount of computing time that would be needed to solve these problems. For example, the SIA is not a particularly good representation of the processes that control valley glaciers (Hubbard, 2000; Egholm and Nielsen, 2010) or for ice sheets where bed topography may be pronounced (e.g. at divides and grounding zones), as it does not effectively capture longitudinal and transverse stress gradients (Pelletier et al., 2010; Egholm et al., 2011) or the contribution of basal slip to glacier movement (Gudmundsson, 2003). Models that include these stresses are called 'higher order models' (Pattyn et al., 2008). A further limitation of the SIA is the dependence of glacier response time on the grid resolution used (Oerlemans, 2008).

The non-linear variability of natural systems can be difficult to capture in a model. For example, the PDD model takes a typical approach to defining variability, calculating monthly mean temperature using the assumption that these values have a sinusoidal distribution around a mean value (Braithwaite, 1995). However, in reality, daily temperature patterns may vary in a more

unpredictable fashion. In defining the period over which a model is to run, it is important to consider the degree of transience in ice response over the history of the glacier or ice sheet; the LGM will not be representative of an ice sheet over an entire glacial cycle (Boulton and Hagdorn, 2006). Many complicated feedbacks occur within the cryosphere and these are often difficult to constrain in a model, such as the change in albedo of snow at the glacier surface as it degrades to ice, and the effect of this on the glacier energy balance, or how increased relief due to glacial erosion will affect orographic precipitation (MacGregor et al., 2009). Incorporating these feedbacks into glacial models presents a significant challenge.

The time period over which the model runs is important; glacial-scale climate changes occur over periods of 10^3 – 10^6 years, but the response time of alpine glaciers may be more rapid, with steady state reached in 10^2 years or less. Temporal resolution can introduce uncertainty: depending on the chosen model timestep, climatic conditions may be averaged over a day, a month, a year or more. This leads to problems in defining the mean values for variations in temperature and precipitation over the modelling period, and is likely to remove stochastic events and seasonal variability from the model inputs. Alternatively, models may consider only stochastic events, such as surging behaviour (e.g. Bindshadler, 1982), or jokulhlaups, and their effect on glacier dynamics and hydrography (e.g. Alho and Aaltonen, 2008). The difference in timescale over which important glacier processes occur varies by orders of magnitude. Stochastic events, such as avalanches, are important controls on glacier behaviour. Avalanches will redistribute snow mass rapidly, and expose rock surfaces to erosion, but their occurrence is difficult to predict and dependent on the characteristics of the snow and the weather conditions. The dynamics of avalanching are poorly understood as a large number of controls exist (Ballantyne, 2007), so while the angle of repose of a slope can be used as a simple threshold for avalanche initiation this is not a totally satisfactory solution.

Glacier steady state may represent the end point of a model output, and occurs where accumulation is exactly equal to ablation over

a period of many years and the glacier does not change in size. However, steady state rarely occurs in real glaciers (Cuffey and Paterson, 2010) so if a model depends on the assumption of steady state it may not be representative of the real glacier. Finally, verification of model results is frequently carried out by comparing the modelled glaciers to the geomorphologic record, often in the form of fragmentary terminal moraines, which may have more than one interpretation, i.e. a landslide- rather than climate- driven formation mechanism (e.g. Torvar et al., 2008; Vacco et al., 2010). However, it is possible to systematically assess the correlation between the model outputs and observed landforms automatically to improve the accuracy of these results and evaluate model performance (Napieralski et al., 2007).

Advantages and disadvantages of glacier modelling

A numerical model can be a powerful tool to reconstruct events for which there is no geological record and which would be difficult to understand quantitatively without modelling. However, uncertainties exist in any numerical model that attempts to describe the physical complexity of the natural environment. Glacier models require complex physical relationships to be defined, often using a single function. Increasing model complexity can increase calculation uncertainty, and it is vital that an estimate of uncertainty is made when modelling is undertaken, to allow understanding of the degree of confidence expressed by the model results. Moreover, it is often necessary to use output from other models to define the parameters needed (e.g. to quantify the degree of cooling at the LGM compared to the present day) and this can introduce a further level of uncertainty into calculations that must also be accounted for. Despite this, glacier models are useful to inform geomorphological studies of glaciated or deglaciated landscapes, frequently in tandem with field studies, where each aspect of the study (modelling and fieldwork) supports and validates each other, to provide a robust, quantitative understanding of cryosphere behaviour, that would not be possible through fieldwork alone.

Alternative methods

Numerical models of glacial systems are commonly used for a variety of geomorphic applications, as there are few suitable alternatives. Construction of a scaled physical model of a glacier in the way that is possible in fluvial flume studies would be difficult, although some aspects of glacier behaviour can be recreated in the laboratory, such as controls on ice melting (e.g. Reznichenko et al., 2010), ice flow (e.g. Glen, 1955) and erosion and sediment transport processes at the glacier bed (e.g. Iverson, 1990, 2000). However, many studies monitor and experiment with real glaciers to investigate processes in a similar way.

The best solution where modelling is not suitable is a field-based study of modern glacier dynamics and geomorphology, with all that can be inferred from this about past glacial behaviour, although ideally an integrated approach that combines field observations with modelling is used to validate results (Hubbard et al., 2005).

References

Alho, P., Aaltonen, J., 2008. Comparing a 1D hydraulic model with a 2D hydraulic model for the simulation of extreme glacial outburst floods. *Hydrological Processes*, 22(10), 1537–1547.

Alley, R., Lawson, D.E., Larson, G.J., Evenson, E.B., 2003. Stabilizing feedbacks in glacier-bed erosion. *Nature*, 424, pp 758–760.

Anders, A.M., Roe, G.H., Montgomery, D.R., Hallet, B., 2008. Influence of precipitation phase on the form of mountain ranges. *Geology*, 36(6), p.479.

Anderson, B. et al., 2010. Climate sensitivity of a high-precipitation glacier in New Zealand. *Journal of Glaciology*, 56(195), 114–128.

Ballantyne, C.K., 2007. The Loch Lomond Readvance on north Arran, Scotland: glacier reconstruction and palaeoclimatic implications. *Journal of Quaternary Science*, 22(4), 343–359.

Ballantyne, C.K., 2002. The Loch Lomond Readvance on the Isle of Mull, Scotland: glacier reconstruction and palaeoclimatic implications. *Journal of Quaternary Science*, 17(8), 759–771.

Benn, D.I., Hulton, N.R., 2010. An Excel™ spreadsheet program for reconstructing the surface profile of former mountain glaciers and ice caps. *Computers and Geosciences*, 36(5), 605–610.

Bernhardt, M., Liston, G.E., Strasser, U., Zängl, G., Schulz, K., 2010. High resolution modelling of snow transport in complex terrain using downscaled MM5 wind fields. *The Cryosphere*, 4, 99–113.

Bindschadler, R., 1983. A numerical model of temperate glacier flow applied to the quiescent phase of a surge-type glacier. *Journal of Glaciology*, 28(99), 239–265.

Boulton, G., Hagdorn, M., 2006. Glaciology of the British Isles Ice Sheet during the last glacial cycle: form, flow, streams and lobes. *Quaternary Science Reviews*, 25, 3359–3390.

Braithwaite, R., 1995. Positive degree-day factors for ablation on the Greenland ice sheet studied by energy-balance modelling. *Journal of Glaciology*, 41(137), 153–160.

Braithwaite, R.J., Raper, S., 2007. Glaciological conditions in seven contrasting regions estimated with the degree-day model. *Annals of Glaciology*, 46(1), 297–302.

Braun, J., Zwartz, D., Tomkin, J., 1999. A new surface-processes model combining glacial and fluvial erosion. *Annals of Glaciology*, 28(1), 282–290.

Brenning, A., Trombotto, D., 2006. Logistic regression modeling of rock glacier and glacier distribution: Topographic and climatic controls in the semi-arid Andes. *Geomorphology*, 81(1-2), 141–154.

Brock, B., Arnold, N., 2000. A spreadsheet-based (Microsoft Excel) point surface energy balance model for glacier and snow melt studies. *Earth Surface Processes and Landforms*, 25(6), 649–658.

- Brocklehurst, S., MacGregor, K., 2009. Numerical modelling of glacial landscape response to the Middle Pleistocene Transition. *Eos Transactions AGU*, 90(52), Fall Meeting Supplement, Abstract EP42A-02.
- Brocklehurst, S.H., Whipple, K., 2007. Response of glacial landscapes to spatial variations in rock uplift rate. *Journal of Geophysical Research*, 112(F2), F02035.
- Bueler, E., Brown, J., 2009. Shallow shelf approximation as a “sliding law” in a thermomechanically coupled ice sheet model. *Journal of Geophysical Research*, 114(F3).
- Carr, S.J., Lukas, S., Mills, S.C., 2010. Glacier reconstruction and mass-balance modelling as a geomorphic and palaeoclimatic tool. *Earth Surface Processes and Landforms*, 35(9), 1103–1115.
- Coleman, C.G., Carr, S.J., Parker, A.G., 2009. Modelling topoclimatic controls on palaeoglaciers: implications for inferring palaeoclimate from geomorphic evidence. *Quaternary Science Reviews*, 28(3-4), 249–259.
- Cuffey, K., Paterson, W., 2010. *The physics of glaciers*. Butterworth Heinemann, Oxford.
- Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., Pasteris, P., 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research*, 22(2), 99–113.
- Dowdeswell, J., Siegert, M., 1999. Ice-sheet numerical modeling and marine geophysical measurements of glacier-derived sedimentation on the Eurasian Arctic continental margins. *Geological Society of America Bulletin*, 111(7), 1080–1097.
- Egholm, D., Nielsen, S., 2010. An adaptive finite volume solver for ice sheets and glaciers. *Journal of Geophysical Research*, 115(F1), F01006.
- Egholm, D., Knudsen, M., Clark, C., 2011. Modeling the flow of glaciers in steep terrains: The integrated second-order shallow ice approximation (iSOSIA). *Journal of Geophysical Research*, 116, F02012–28.
- Egholm, D.L., Nielsen, S.B., Pedersen, V.K., Lesemann, J.-E., 2010. Glacial effects limiting mountain height. *Nature*, 460(7257), 884–887.
- Evans, D.J.A., Livingstone, S.J., Vieli, A., Cofaigh, C.O., 2009. The palaeoglaciology of the central sector of the British and Irish Ice Sheet: reconciling glacial geomorphology and preliminary ice sheet modelling. *Quaternary Science Reviews*, 28(7-8), 739–757.
- Foster, D., Brocklehurst, S.H., Gawthorpe, R.L., 2010. Glacial-topographic interactions in the Teton Range, Wyoming. *Journal of Geophysical Research*, 115(F1), p.F01007.
- Glen, J., 1955. The Creep of Polycrystalline Ice. In *Proceedings of the Royal Society of London*, 228(1175), 519-538.
- Gudmundsson, G.H., 1999. A three-dimensional numerical model of the confluence area of Unteraargletscher, Bernese Alps, Switzerland. *Journal of Glaciology*, 45(150), 219–230.
- Gudmundsson, G.H., 2003. Transmission of basal variability to a glacier surface. *Journal of Geophysical Research*, 108(B5), 2253–2262.
- Haeberli, W., Hallet, B., Arenson, L., Elconin, R., Humlum, O., Kääb, A., Kaufmann, V., Ladanyi, B., Matsuoka, N., Springman, S., Mühl, D.V., 2006. Permafrost creep and rock glacier dynamics. *Permafrost and Periglacial Processes*, 17(3), 189–214.
- Harbor, J., 1992. Numerical modeling of the development of U-shaped valleys by glacial erosion. *Bulletin of the Geological Society of America*, 104, 1364–1375.
- Herman, F., Braun, J., 2008. Evolution of the glacial landscape of the Southern Alps of New Zealand: Insights from a glacial erosion model. *Journal of Geophysical Research*, 113(F2), F02009.
- Hooke, R., 1991. Positive feedbacks associated with erosion of glacial cirques and overdeepenings. *Bulletin of the Geological Society of America*, 103, 1104–1108.
- Hubbard, A., 1999. High-resolution modeling of the advance of the Younger Dryas ice

- sheet and its climate in Scotland. *Quaternary Research*, 52(1), 27–43.
- Hubbard, A., 2006. The validation and sensitivity of a model of the Icelandic ice sheet. *Quaternary Science Reviews*, 25, 2297–2313.
- Hubbard, Alun, 2000. The Verification and Significance of Three Approaches to Longitudinal Stresses in High-resolution Models of Glacier Flow. *Geografiska Annaler: Series A, Physical Geography*, 82(4), 471–487.
- Hubbard, A et al., 2005. A modelling reconstruction of the last glacial maximum ice sheet and its deglaciation in the vicinity of the Northern Patagonian Icefield, South America. *Geografiska Annaler*, 87, 375–391.
- Hughes, P.D., 2009. Loch Lomond Stadial (Younger Dryas) glaciers and climate in Wales. *Geological Journal*, 44(4), 375–391.
- Hutter, K., 1983. *Theoretical Glaciology: Material Science of Ice and the Mechanics of Glaciers and Ice Sheets*, Reidel, Tokyo.
- Huybrechts, P., 1990. A 3-D model for the Antarctic ice sheet: a sensitivity study on the glacial-interglacial contrast. *Climate Dynamics*, 5, 79–92.
- Imhof, M., 1996. Modelling and verification of the permafrost distribution in the Bernese Alps (Western Switzerland). *Permafrost and Periglacial Processes*, 7, 267–280.
- Iverson, N., 2000. Sediment entrainment by a soft-bedded glacier: a model based on regelation into the bed. *Earth Surface Processes and Landforms*, 25, 881–893.
- Iverson, N.R., 1990. Laboratory Simulations of Glacial Abrasion: Comparison with Theory. *Journal of Glaciology*, 36(124), 304–314.
- Jamieson, S., Hulton, N., 2008. Modelling landscape evolution under ice sheets. *Geomorphology*, 97, 91–108.
- Jordan, P., 2009. Designing the DEM of the base of the Swiss Plateau Quaternary sediments. 6th ICA Mountain Cartography Workshop Mountain Mapping and Visualisation, 107–113.
- Kelly, M., Kubik, P., von Blackenburg, F., Schlüchter, C., 2004. Surface exposure dating of the Great Aletsch Glacier Egesen moraine system, western Swiss Alps, using the cosmogenic nuclide. *Journal of Quaternary Science*, 19, 431–441.
- Klok, E., 2002. Model study of the spatial distribution of the energy and mass balance of Morteratschgletscher, Switzerland. *Journal of Glaciology*, 48, 505–518.
- MacGregor, K., Anderson, R.S., Waddington, E.D., 2009. Numerical modeling of glacial erosion and headwall processes in alpine valleys. *Geomorphology*, 103(2), 189–204.
- MacGregor, K., Anderson, R.S., Anderson, S.P., Waddington, E.D., 2000. Numerical simulations of glacial-valley longitudinal profile evolution. *Geology*, 28(11), p.1031.
- Marshall, S., Clarke, G., 1999. Ice sheet inception: subgird hypsometric parameterization of mass balance in an ice sheet model. *Climate Dynamics*, 15, 533–550.
- Meierding, T., 1982. Late Pleistocene glacial equilibrium-line altitudes in the Colorado Front Range: a comparison of methods. *Quaternary Research*, 18, 289–310.
- Molnar, P., England, P., 1990. Late Cenozoic uplift of mountain ranges and global climate change: chicken or egg? *Nature*, 346, 29–34.
- Murray, D., Locke, W.W., 1989. Dynamics of the Late Pleistocene Big Timber Glacier, Crazy Mountains, Montana, USA. *Journal of Glaciology*, 35, 183–190.
- Napieralski, J., Hubbard, A., Li, Y., Harbor, J., Stroeven, A.P., Kleman, J., Alm, G., Jansson, K.N., 2007. Towards a GIS assessment of numerical ice-sheet model performance using geomorphological data. *Journal of Glaciology*, 53(180), 71–83.
- Oerlemans, J., 1997. A flowline model for Nigardsbreen, Norway: projection of future glacier length based on dynamic calibration with the historic record. *Annals of Glaciology*, 24, 382–389.

- Oerlemans, J., 2008. *Minimal Glacier Models*. Igitur Press, Utrecht.
- Oerlemans, J., 1984. Numerical experiments on glacial erosion. *Zeitschrift für Gletscherkunde und Glazialgeologie*, 20, 107–126.
- Oerlemans, J., 2010. *The microclimate of valley glaciers*. Igitur Press, Utrecht.
- Ohmura, A., 2001. Physical basis for the temperature-based melt-index method. *Journal of Applied Meteorology*, 40(4), 753–761.
- Pattyn, F., 1996. Numerical modelling of a fast-flowing outlet glacier: experiments with different basal conditions. *Annals of Glaciology*, 23, 237–246.
- Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., De Smedt, B., Gagliardini, O., Gudmundsson, G.H., Hindmarsh, R., Hubbard, A., Johnson, J.V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A.J., Pollard, D., Price, S., Rückamp, M., Saito, F., Soucek, O., Sugiyama, S., Zwinger, T., 2008. Benchmark experiments for higher-order and full Stokes ice sheet models (ISMIP-HOM). *The Cryosphere Discussions*, 2, 111–151.
- Pelletier, J., Comeau, D., Kargel, J., 2010. Controls of glacial valley spacing on earth and mars. *Geomorphology*, 116, 189–201.
- Pelletier, J.D., 2010. Numerical modeling of the late Cenozoic geomorphic evolution of Grand Canyon, Arizona. *Geological Society of America Bulletin*, 122(3-4), 595–608.
- Petrenko, V., Whitworth, R., 1983. *Physics of Ice*, Oxford University Press, Oxford.
- Plummer, M., Phillips, F., 2003. A 2-D numerical model of snow/ice energy balance and ice flow for paleoclimatic interpretation of glacial geomorphic features. *Quaternary Science Reviews*, 22(14), 1389–1406.
- Rea, B.R., 2009. Defining modern day Area-Altitude Balance Ratios (AABRs) and their use in glacier-climate reconstructions. *Quaternary Science Reviews*, 28(3-4), 237–248.
- Reznichenko, N., Davies, T., Shulmeister, J., 2010. Effects of debris on ice-surface melting rates: an experimental study. *Journal of Glaciology*, 56(197), 384–394.
- Richards, K., Sharp, M., Arnold, N., Gurnell, A., Clark, M., Tranter, M., Nienow, P., Brown, G., Willis, I., Lawson, W., 1996. An integrated approach to modelling hydrology and water quality in glacierized catchments. *Hydrological Processes*, 10, 479–508.
- Rutt, I., Hagdorn, M., Hulton, N.R.J., Payne, A.J., 2009. The Glimmer community ice sheet model. *Journal of Geophysical Research*, 114, 1–22.
- Schmittner, A., 2002. Instability of Glacial Climate in a Model of the Ocean-Atmosphere-Cryosphere System. *Science*, 295(5559), 1489–1493.
- Syvitski, J., Vorosmarty, C.J., Kettner, A.J., Green, P., 2005. Impact of Humans on the Flux of Terrestrial Sediment to the Global Coastal Ocean. *Science*, 308, 376–380.
- Tomkin, J., 2007. Coupling glacial erosion and tectonics at active orogens: A numerical modeling study. *Journal of Geophysical Research*, 112(F2), p.F02015.
- Tovar, D., Shulmeister, J. & Davies, T.R., 2008. Evidence for a landslide origin of New Zealand's Waiho Loop moraine. *Nature Geoscience*, 1(8), 524–526.
- Vacco, D.A., Alley, R.B. & Pollard, D., 2010. Glacial advance and stagnation caused by rock avalanches. *Earth and Planetary Science Letters*, 294(1-2), 123–130.
- Whipple, K., Kirby, E., Brocklehurst, S., 1999. Geomorphic limits to climate-induced increases in topographic relief. *Nature*, 401, 39–43.
- Winkelmann, R., Martin, M.A., Haseloff, M., Albecht, T., Bueler, E., Khroulev, C., Levermann, A., 2010. The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description. *The Cryosphere Discussions*, 4(3), 1277–1306.

BSG

British Society for Geomorphology