'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications

Westoby, M. J.; Brasington, J.; Glasser, N. F.; Hambrey, M. J.; Reynolds, J. M.

Published in:
Geomorphology

DOI:
10.1016/j.geomorph.2012.08.021

Publication date:
2012

Citation for published version (APA):
Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300-314. https://doi.org/10.1016/j.geomorph.2012.08.021

‘Structure-from-Motion’ photogrammetry: a low-cost, effective tool for geoscience applications

M.J. Westoby[1,*], J. Brasington[2], N.F. Glasser[1], M.J. Hambrey[1], and J.M. Reynolds[3]

*Corresponding author. Tel.: +44 1970 628 608; Fax: +44 1970 622 659
E-mail: mjwestoby@gmail.com

[1]Institute of Geography and Earth Sciences, Penglais Campus, Aberystwyth University, UK
[2]School of Geography, Queen Mary, University of London, London, UK
[3]Reynolds International Ltd., Unit 17, Mold Business Park, Mold, UK.

Abstract:

High-resolution topographic surveying is traditionally associated with high capital and logistical costs, so that data acquisition is often passed on to specialist third party organizations. The high costs of data collection are, for many applications in the earth sciences, exacerbated by the remoteness and inaccessibility of many field sites, rendering cheaper, more portable surveying platforms (i.e. terrestrial laser scanning or GPS) impractical. This paper outlines a revolutionary, low-cost, user-friendly photogrammetric technique for obtaining high-resolution datasets at a range of scales, termed ‘Structure-from-Motion’ (SfM). Traditional softcopy photogrammetric methods require the 3-D location and pose of the camera(s), or the 3-D location of ground control points to be known to facilitate scene triangulation and reconstruction. In contrast, the SfM method solves the camera pose and scene geometry simultaneously and automatically, using a highly redundant bundle adjustment based on matching features in multiple overlapping, offset images. A comprehensive introduction to the technique is presented, followed by an outline of the methods used to create high-resolution Digital Elevation Models (DEMs) from extensive photosets obtained using a consumer-grade digital camera. As an initial appraisal of the technique, an SfM-derived DEM is compared directly with a similar model obtained using Terrestrial Laser Scanning. This intercomparison reveals that decimetre-scale vertical accuracy can be achieved using SfM even for sites with complex topography and a range of land-covers. Example applications of SfM are presented for three contrasting landforms across a range of scales including; an exposed rocky coastal cliff; a breached moraine-dam complex; and a glacially-sculpted bedrock ridge. The SfM technique represents a major advancement in the field of photogrammetry for geoscience applications. Our results and experiences indicate SfM is an inexpensive, effective, and flexible approach to capturing complex topography.

Key words:

Structure-from-Motion; SfM; Close- range photogrammetry; Digital Elevation Model

1. Introduction

The last decade has witnessed a technological revolution in geomatics that is transforming digital elevation modelling and geomorphological terrain analysis. Spurred on by developments in traditional ground surveying, such as the advent of differential GPS (e.g., Brasington et al., 2000) and reflectorless, robotic total stations (e.g. Keim et al., 1999; Fuller et al., 2003), the acquisition of topographic data has been transformed most significantly by a new generation of remote sensing technologies. Airborne and more recently terrestrial laser scanning (e.g., Lohani and Mason, 2001; Jones et al., 2007; Notebaert et al., 2009; Rosser et al., 2005; Heritage and Hetherington, 2007; Hodge et al., 2009) and soft-copy photogrammetry (e.g., Lane et al., 2000; Westaway et al., 2000; Brasington et al., 2003) in particular, have revolutionized the quality of DEMs, extending their spatial extent, resolution, and accuracy.

Developments in airborne and terrestrial remote sensing have also been mirrored by advances in hydrographic surveying, in particular through single and multi-beam sonar (e.g., Parsons et al., 2005; Sacchetti et al., 2012). These acoustic soundings, capable of centimetric data spacing and 3-D point quality herald the potential to construct truly continuous, high fidelity terrain models of mixed terrestrial, freshwater and marine environments. Finally, and closing the technological loop, the acquisition of remotely sensed data from a range of cheap, lightweight platforms on which to deploy imaging sensors, such as unmanned aerial vehicles or UAVs (e.g., Lejot et al. 2007; Niethammer et al., 2012) and tethered kites and blimps (e.g., Marzolff et al., 2003; Boike and Yoshikawa, 2003; Smith et al., 2009; Vericat et al., 2009,) is gradually becoming more commonplace.

While the pace of development in geospatial technologies has been rapid, the acquisition of high quality terrain data nonetheless remains challenging in remote, high alpine environments. In these hostile landscapes, steep and unconsolidated slopes and poor satellite coverage hinders the application of ground surveys by GPS or total station. Alternative ground-based methods such as terrestrial laser scanning (TLS) are complicated by the high capital investment cost and the portability of large instruments and their power requirements in remote areas. Airborne surveys, including LiDAR and photography are also of restricted use due to the high three-dimensionality of mountainous landscapes, which results in significant line of sight losses and image foreshortening. Moreover, deploying survey platforms, including helicopters and smaller scale UAVs at altitude is highly dependent on favourable weather conditions and may often be hampered by high wind speed and cloud cover. Potential solutions may ultimately lie in the availability of high resolution satellite data, but at present the spatial resolution of the majority of existing active and passive sensors is typically too coarse to create digital elevation models (DEMs) at resolutions comparable to ground-based techniques and suitable for detailed geomorphological applications

1.1. Photogrammetric Survey Methods

In the decade or so since its emergence, automated aerial and close-range digital photogrammetry has become a powerful and widely used tool for three-dimensional topographic modelling (Remondino and El-Hakim, 2006; Matthews, 2008). The development of soft-copy triangulation and image-based terrain extraction algorithms have radically enhanced the quality of terrain data that can be derived from overlapping stereo-pairs (Chandler, 1999; Lane et al., 2000). Similarly, improvements in the cost and quality of compact and single lens reflex (SLR) cameras, and methods for the calibration of such non-metric cameras (Clarke and Fryer, 1998; Chandler et al., 2005) have democratized access to photogrammetric modelling and encouraged a wide range of uses in geomorphology. This has included monitoring river bed topography and planform (e.g. Lane, 2000; Chandler et al., 2002; Brasington and Smart, 2003; Bird et al., 2010), river bank (e.g. Barker et al., 1997; Pyle et al., 1997) and gully erosion (e.g. Betts and DeRose, 1999; Marzolff and Poesen, 2010), and in the field of glaciology, the quantification of glacier surface change (e.g. Keutterling and Thomas, 2006; Baltsavias et al., 2008). Digital photogrammetry has also been applied to a number of geological problems, including discontinuity characterization (e.g. Krosley et al., 2006; Sturzenegger and Stead, 2009) and rock slope stability analysis (e.g. Haneberg, 2008). Close-range applications have also included direct quantification of soil erosion and the morphodynamics of laboratory-scale landscape evolution models (e.g. Stojic et al., 1998; Brasington and Smart, 2003; Lane et al., 2001; Hancock and Willgoose, 2001; Rieke-Zapp and Nearing, 2005; Heng et al., 2010).

1.2. Structure-from-Motion

In this paper, we report on an emerging, low-cost photogrammetric method for high resolution topographic reconstruction, ideally suited for low-budget research and application in remote areas. ‘Structure-from-Motion’ (SfM) operates under the same basic tenets as stereoscopic photogrammetry, namely that 3-D structure can be resolved from a series of overlapping, offset images (Fig. 1). However, it differs fundamentally from conventional photogrammetry, in that the geometry of the scene, camera positions and orientation are solved automatically without the need to specify *a priori*, a network of targets which have known 3-D positions. Instead, these are solved simultaneously using a highly redundant, iterative bundle adjustment procedure, based on a database of features automatically extracted from a set of multiple overlapping images (Snavely, 2008). As described below, the approach is most suited to sets of images with a high degree of overlap that capture full three-dimensional structure of the scene viewed from a wide array of positions, or as the name suggests, images derived from a moving sensor.

Developed in the 1990s, this technique has its origins in the computer vision community (e.g. Spetsakis and Aloimonos, 1991; Boufama et al., 1993; Szeliski and Kang, 1994) and the development of automatic feature-matching algorithms in the previous decade (e.g. Förstner, 1986; Harris and Stephens, 1988). The approach has been popularized through a range of cloud-processing engines, most notably Microsoft® Photosynth™ (Microsoft, 2010), which uses SfM approaches documented in Snavely (2008) and Snavely *et al.* (2008). These tools can make direct use of user-uploaded and crowd-sourced photography to generated the necessary coverage of a target scene, and can automatically generate sparse 3-D point clouds from these photosets. The possibilities of SfM appear boundless, however, to date, the technique has rarely been used within the geosciences (e.g. Niethammer et al., 2012) and there exist few quantitative assessments of the quality of terrain products derived from this approach.

1.3. The First Principles of SfM

To determine the 3-D location of points within a scene, traditional softcopy photogrammetric methods require the 3-D location and pose of the camera(s), or the 3-D location of a series of control points to be known. Using the former, in the absence of a camera-mounted GPS and electronic compass, triangulation can be used to reconstruct scene geometry, whilst in the case of the latter, control points are manually identified in the input photographs, and a process called *resectioning*, or *camera pose estimation*, used to determine camera position. In contrast, the SfM approach requires neither of the above to be known prior to scene reconstruction. Camera pose and scene geometry are reconstructed simultaneously through the automatic identification of matching features in multiple images. These features are tracked from image to image, enabling initial estimates of camera positions and object coordinates which are then refined iteratively using non-linear least-squares minimization (as multiple solutions become available from the wide range of features in the image database (Snavely, 2008)).

Unlike traditional photogrammetry, the camera positions derived from SfM lack the scale and orientation provided by ground-control coordinates. Consequently, the 3-D point clouds are generated in a relative 'image-space' coordinate system, which must be aligned to a real-world, 'object-space' co-ordinate system. In most cases, the transformation of SfM image-space coordinates to an absolute coordinate system can be achieved using a 3-D similarity transform based on a small number of known ground-control points (GCPs) with known object-space coordinates. Such GCPs can be derived post-hoc, identifying candidate features clearly visible in both the resulting point cloud and in the field, and obtaining their coordinates by ground survey (i.e., by GPS). In practice, however, it is often easier to deploy physical targets with a high contrast and clearly defined centroid in the field before acquiring images. This approach simplifies the unambiguous co-location of image and object space targets and also ensures a reliable, well-distributed network of targets across the area of interest, enabling an assessment of any non-linear structural errors in the SfM reconstruction. It is also useful to incorporate a degree of redundancy in the GCP network to counter the possibility of sparse data in the region of the targets.

1.4. Goals of this Article

The aim of this paper is to provide an introduction to SfM and a detailed explanation of the methods employed; illustrating the steps required to generate a fully rendered 3-D model, starting from the initial acquisition of the photographic data. The focus here is to outline a practical workflow that could be applied by environmental scientists and practitioners interested in deploying SfM for geomorphological research. To achieve this, we describe a workflow that uses the freely available application bundle SFMToolkit3 (Astre, 2010) to process the photographs and produce the initial point cloud. This package contains a number of open-source applications including, in order of execution, SiftGPU (Lowe, 1999, 2004), Bundler (Snavely et al., 2008), CMVS and PMVS2 (Furukawa and Ponce, 2007; Furukawa et al., 2010), all of which may be run independently if desired.

Applications of SfM to a range of contrasting landscapes and landforms are described, including coastal cliffs, a moraine-dammed lake, and a smaller scale glacially-sculpted bedrock ridge. Importantly, we also undertake a detailed assessment of the quality of a derived topographic model, in this case a c. 300 x 300 m cliff section in Aberystwyth, Wales, through comparison with a high resolution terrain model derived from a precision terrestrial laser scan survey.

2. Method

2.1. Structure-from-Motion workflow

2.1.1. Image acquisition and keypoint extraction

The key problem that SfM addresses is the determination of the 3-D location of matching features in multiple photographs, taken from different angles. The initial processing step in the solution of this problem is the identification of features in individual images which may be used for image correspondence. A popular solution to this, and used in the methods popularized by Snavely (2008) is the Scale Invariant Feature Transform (SIFT) object recognition system. This is implemented in SFMToolkit3, through the incorporation of the SiftGPU algorithm (Lowe, 1999; 2004).

This identifies features in each image that are invariant to the image scaling and rotation and partially invariant to changes in illumination conditions and 3-D camera viewpoint (Fig. 2; Lowe, 2004). Points of interest, or ‘keypoints’, are automatically identified over all scales and locations in each image, followed by the creation of a feature descriptor, computed by transforming local image gradients into a representation that is largely insensitive to variations in illumination and orientation (Lowe, 2004). These descriptors are unique enough to allow features to be matched in large datasets.

The number of keypoints in an image is dependent primarily on image texture and resolution, such that complex images at high (often original) resolutions will return the most results. The density, sharpness, and resolution of the photoset, combined with the range of natural scene textures will, in the first instance therefore, determine the quality of the output point cloud data. Similarly, decreasing the distance between the camera and feature of interest, thereby increasing the spatial resolution of the photograph, will enhance the spatial density and resolution of the final point cloud.

Variations in the complexity, lighting, materials in individual scenes all influence the image texture so it is impossible to offer explicit guidance on the minimum number of photographs necessary for successful scene reconstruction. The minimum requirement is for corresponding features to be visible in a minimum of three photographs, however, obtaining as many images for SfM input as possible, given logistical constraints, is highly recommended as this optimizes the ultimate number of keypoint matches and system redundancy.

Particular consideration should also be given to the choice of acquisition platform. For example, small scale sites with steep slope angles are likely to be better suited to an exclusively ground-based approach, whereas low altitude aerial photography (LAAP) may provide better coverage over larger sites and those with more subdued topography. Indeed, imagery combined from multiple platforms may prove to be optimal, providing different levels of detail in different areas of the scene. When acquiring the photographs, particular attention should be taken to maximise overlap by adopting short camera baselines (i.e., the distance between successive photography positions), and obtaining as uniform coverage of the feature or landscape of interest as possible.

A wide variety of imaging sensors can be used for SfM, from video stills, through to low grade compact digital cameras. The primary requirement is well-exposed photographs of the feature(s) of interest. From our experience, ‘bigger’ is not necessarily ‘better’. Whereas image quality and resolution are improved by using increasingly expensive digital SLR models, images captured at the highest resolutions (e.g. >12 megapixel) will almost inevitably need to be re-sized (with the consequent loss of image detail) to avoid lengthy processing times. If operating in remote regions, specific consideration should be given to robustness and battery life, including methods for charging and performance in extreme temperatures.

2.1.2. 3-D scene reconstruction

Following keypoint identification and descriptor assignment, the sparse bundle adjustment system Bundler (Snavely et al., 2008) is used to estimate camera pose and extract a low-density or ‘sparse’ point cloud. Keypoints in multiple images are matched using approximate nearest neighbour (Arya et al., 1998) and Random Sample Consensus (RANSAC; Fischler and Bolles, 1987) algorithms, and ‘tracks’ linking specific keypoints in a set of pictures, are established. Tracks comprising a minimum of two keypoints and three images are used for point-cloud reconstruction, with those which fail to meet these criteria being automatically discarded (Snavely et al., 2006). Using this method, transient features such as people moving across the area of interest are automatically removed from the dataset before 3-D reconstruction begins. This also applies to non-static objects unintentionally captured in the input photoset, such as objects on the sensor, blimp or kite tethers, or helicopter landing skids. In these instances, although identical keypoints referencing such objects will be created, they are not suitable for use in scene reconstruction as their position relative to other keypoints is constantly changing, and are automatically filtered using visibility and regularization constraints (Furukawa and Ponce, 2009).

Keypoint correspondences place constraints on camera pose orientation, which is reconstructed using a similarity transformation, while minimization of errors is achieved using a non-linear least-squares solution (Szeliski and Kang, 1994; Nocedal and Wright, 1999). Finally, triangulation is used to estimate the 3-D point positions and incrementally reconstruct scene geometry, fixed into a relative coordinate system. Full automation of this process, from keypoint extraction, to accurate reconstruction of scene geometry is a clear advantage of the SfM method over traditional digital photogrammetric approaches.

The bundle adjustment package used in Bundler (Lourakis and Argyros, 2009) produces sparse point-clouds. An enhanced density point-cloud can be derived by implementing the Clustering View for Multi-view Stereo (CMVS) (Furukawa and Ponce, 2007; Furukawa et al., 2010) and Patch-based Multi-view Stereo (PMVS2) algorithms (Furukawa and Ponce, 2007). Here, camera positions derived from Bundler are used as input. CMVS then decomposes overlapping input images into subsets or clusters of manageable size, whilst PMVS2 is used to independently reconstruct 3-D data from these individual clusters (Furukawa and Ponce, 2007). The result of this additional processing is a significant increase in point density; typically approaching, or in excess of 2 orders of magnitude (Table 1).

2.1.3. Post-processing and digital elevation model generation

Transformation from a relative to absolute co-ordinate system is achieved through manual identification of GCPs in the point cloud and the computation of an appropriate transformation. Here, we use a rigid body transformation decomposed into a rotation and a translation matrix, and a scale factor. A solution to this transformation is found using a modified version of Horn's (1987) absolute orientation algorithm implemented in MATLAB®. A solution to the seven unknowns in the total transformation requires a minimum of three matching image and object space coordinates.

Targets are placed to provide maximum visibility (and thus obvious appearance in the photographs), and contrast as strongly as possible with the surroundings to aid their location in the final point-cloud before the data are transformed. Significant outliers and artefacts (e.g. erroneous peaks and troughs resulting from keypoint descriptor mismatches) and any unnecessarily reconstructed surrounding topography are manually removed at this stage.

The point clouds generated using SfM may, depending on the image set, be extremely dense, potentially incorporating upwards of 10^3 3-D points per square metre. Direct interpolation of the raw point cloud into a typically coarser resolution terrain model (e.g., metre-scale) represents a considerable computational task. Additionally, and in common with other remote sensing methods such as LiDAR or TLS, this raw point cloud may also incorporate 'off-ground' features that obscure simple geomorphological interpretation.

To improve data-handling and provide a first-order bare-earth elevation model, here we decimate the raw point cloud using a gridding procedure developed by Rychkov et al. (2012). This approach, originally designed to explore TLS point clouds, decomposes the point cloud into a regular grid, for which parameters of the local elevation distribution are extracted. These include the minimum, maximum, mean and first and higher order moments. The routine then fits a local tessellation to this reduced resolution grid, based on a local elevation estimate, and then detrends the raw point cloud, to derive a comparable set of local statistics that reflect variability above the first order grid-scale features. This simple, but computationally efficient procedure allows easy extraction of terrain models based, for example on the local minimum grid elevation, whilst retaining information on the sub-grid elevation complexity for later analysis. Visualisation of the grid-cell statistics also permits spatial analysis of the variability in point density across the entire model (see Fig. 6d).

When combined, SfM and point-cloud decimation potentially offer a powerful tool for geomorphological analysis. For example, using the approach on comparable TLS derived point clouds, Rychkov et al. (2012) were able reveal gravel-scale grain scale roughness on steeply inclined river banks. Similarly, in the absence of high-resolution imagery from space-borne platforms or aerial LiDAR, entire floodplains and valley floor reaches may be surveyed using cameras mounted on low-altitude tethered platforms or UAVs (e.g. Niethammer, 2012) and subsequently decimated to resolutions required for the extraction of geometric data required as boundary condition data for hydrodynamic modelling.

However, depending on the final application, data decimation may not be necessary and unwanted, although conventional GIS software is typically inappropriate to manage the visualization and storage of dense point data which may extend into tens or hundreds of millions of observations. For the examples presented in this paper, decimation was applied in order to facilitate direct comparison of SfM and TLS data at resolutions sufficient to represent the first order topography at the scale of interest whilst enabling simple surface generation and visualisation on a desktop PC. The final terrain products were derived by linearly resampling a TIN created by Delaunay triangulation in ArcGIS from decimated point cloud (following, Brasington et al., 2000). This model may be visualized effectively by draping the orthophoto derived from the SfM processing over this surface. The final result is a fully georeferenced, high-resolution, photo-realistic DEM. The complete workflow is shown in Fig. 3.

3. Comparison with Terrestrial Laser Scanning

3.1. Data acquisition and processing

An independent assessment of the accuracy of the SfM method was undertaken by direct comparison of raster DEMs of an exposed eroding cliff created using the SfM workflow described above and a comparable survey with TLS. The study site is Constitution Hill, a ~80 m high coastal cliff located immediately to the north of the town of Aberystwyth, Wales, UK (Fig. 4a). The lower section of the cliff is topographically complex, with land cover comprising grasses, shrubs, footpaths, near-vertical cliff faces and rockfall debris. The exposed bedrock comprises Silurian turbidites which were folded and faulted during the Caledonian orogeny. A cobble beach is located at its base. Labelled yellow targets, 1 x 1 m in size were deployed across the area of interest and used as GCPs. In total, 35 targets were distributed across the study site, in a quasi-uniform pattern allowing for topographic constraints (Fig. 4a).

In addition, three tripod-mounted Leica Geosystems HDS targets were deployed to co-register the TLS data. Scans of the hillside were acquired from three positions using a Leica Geosystems ScanStation (blue triangles in Fig. 4a). This was set to record data with a 2 cm spatial resolution at a distance of 15 m for the first scan and 1 cm for the second and third. Coincidently, a total of 889 photographs were taken of the hillside using a consumer-grade digital camera (Panasonic DMC-G10, 12 megapixel resolution, with both automatic focusing and exposure enabled), from a range of locations and perspectives, for use as input to SfM processing. Using three people, TLS and SfM data acquisition, including deployment of the SfM GCPs (and centroid location surveying) and TLS HDS targets, photograph acquisition, and TLS data acquisition took approximately 5 hours in total, of which the SfM component took approximately 2 hours. Total post-processing times were significantly longer: combined, sparse and dense point cloud generation took a total of 23.5 hours, whilst manual point cloud editing, GCP identification and transformation and DEM generation took an additional ~4 hours.

The TLS data were co-registered into a single point cloud using a 3-D similarity transformation based on manual identification of the three HDS targets in the individual point-clouds. This was implemented using Leica's Cyclone software suite. No attempt to transform the final product into an absolute coordinate system was attempted, as this would incorporate additional unnecessary errors. The data were also not projected, to allow a direct comparison of the two (TLS and SfM) matching Cartesian coordinate systems.

The extensive photoset was decomposed into three 'batches' to reduce computational demand, and input photographs re-scaled to 55% of their original resolution to reduce computational demand. The processing steps outlined in section 2 were employed, producing unreferenced sparse and dense point-clouds as output (Table 2). The SfM data were transformed to the TLS co-ordinate system through manual identification of matching GCP centroids in both datasets (Fig. 4b-d). The three SfM batches were registered individually, with no significant difference in the quality of the three transformation models, and average transformation residuals of 0.124 m, 0.058 m and 0.031 m for *xyz*.

3.2. Results

Perspective views of the sparse and dense point cloud data are presented in Fig. 5. A significant increase in the point density is immediately apparent for the dense reconstruction. After manual editing, the sparse dataset comprised 1.7×10^5 points, whilst the dense reconstruction produced 11.3×10^6 points; a 64-fold increase. This is comparable to the TLS survey density with 11.7×10^6 survey points over the same area. The resolution is sufficient to reveal bedrock structure, notably the style of folding (Fig. 6, 'C' in Fig 5a).

Interpolated DEMs derived using the workflow described in Section 2 are shown in Fig. 7. The models shown here have been extracted at 1 m resolution, and are based on the local minimum elevation in order to aid the automatic removal any off-ground observations (e.g., vegetation cover). The decimation process reduced the original number of points to a regularised grid of 8999 and 10,780 cells for the SfM and TLS data, whist retaining key summary statistics describing the sub-grid elevation variability.

There are minor differences in the spatial extent of the two models, though both resolve clearly notable topographic elements such as near-vertical faces adjacent to the main footpath (labelled ‘A’ and ‘B’ in Fig. 5a and Fig. 7a). Point density data (Fig. 7d, also Fig. 8) reveal that the highest concentrations (~70,000 - >90,000 points per m^2) are located towards the centre-north of the scene, corresponding to sizeable (up to 10 m high) exposed rock faces.

A DEM of difference (DoD) was created by subtracting SfM elevation model from that obtained by TLS (Fig. 9a). This reveals that 94 % of overlapping model differences are in the range -1.0 – 1.0 m, with 86 % between just -0.5 – 0.5 m (Fig. 9b). Spatially, positive differences (SfM higher than TLS) in elevation are concentrated towards the north-east of the dataset, with more pronounced negative disparities confined largely to the west. The negative deviations correspond to the headwall at the highest point of the scene, and steep faces which terminate at the beach, respectively (‘A’ and ‘B’ in Fig. 7a, respectively).

It is hypothesized that the largest DoD values (Fig. 9a), corresponding to areas of steeply sloping relief close to the beach and the headwall at the highest point of the datasets, may be partly attributed to weak reconstruction caused by large camera-to-feature baselines (>20 m) owing to the inaccessibility of these areas. The photogrammetric-TLS model comparison in these areas is also likely to be highly sensitive to small horizontal offsets introduced through point-cloud transformation. Across the rest of the study area, it was possible to photograph the terrain from relatively close proximity, including other steep areas which were well reconstructed, as indicated by low z_{diff} values (e.g., the centre of the study area, adjacent to the footpath).

No statistically significant relationship (r^2 = 0.19) between slope and z_{diff} exists, suggesting that DEM disparities cannot be explained by the local gradient alone. The areas with the highest differences were concentrated, almost exclusively, to the cliff section adjacent to the beach in the west of the dataset (‘B’ in Fig. 8a) and correspond to regions subjected to heavy interpolation. Dense vegetation dominates the eastern end of Constitution Hill (‘C’ in Fig. 8a) this is strongly associated with moderately high (positive) elevation differences. If comprehensive photosets with high degrees of overlap are not acquired for this type of land cover, dense vegetation cover proves problematic for effective terrain reconstruction due to the homogeneous image texture. The same also applies to surfaces including grass, snow, or sand.

The scale of elevation differences apparent between the two models appears to imply that application of repeat SfM to monitor landscape change (by DEMs of difference) would be limited to capturing only relatively large topographic dynamics (i.e., metre scale). However, on closer inspection it is apparent that the notable local deviations between the SfM and TLS datasets correspond largely to areas of relatively dense shrub and bush cover. Indeed, considerably lower deviations (i.e. ±

0.1 m or less) are found in the mostly vegetation-free (with the exception of short grass) central portion of the site closest to the beach. To illustrate this point, z_{diff} frequency distribution histograms for two sub-regions of the site (both 20 m^2 in area), representing an area of dense vegetation and largely vegetation-free ground, respectively, are shown in Fig. 10. Analysis of these data reveal that 100% of cells fall with the range -0.5 – 0.5 m for the vegetation-free region, with 89% of cells falling within this classification for the area of dense vegetation. At the decimetre scale and below, 61% and 39% of the vegetation-free and dense vegetation data, respectively, possess z_{diff} values of ± 0.1 m or less, demonstrating an appreciable decrease in elevation difference in the absence of any significant vegetation cover. Error introduced by the presence of vegetation is not an issue unique to the SfM method, but may confound topographic reconstruction using a wide range of remote surveying methods (e.g. Coveney and Fotheringham, 2011). Despite this, the accuracy assessment serves as an encouraging first appraisal of the SfM algorithms employed. A far more rigorous accuracy assessment, across a range of terrain types, is needed in future research to better elucidate the major sources of error. Similarly, a detailed investigation of systematic errors introduced by, for example, the manual identification of the GCPs in both point cloud datasets, and the subsequent impact upon the accuracy of the transformation matrix applied to the data, would also be desirable.

4. Applications of Structure-from-Motion to large- and small-scale glacial landform reconstruction

4.1 Dig Tsho moraine complex

The workflow outlined above was applied to the Dig Tsho moraine-dam complex in the Khumbu Himal, Nepal (4,400 m) (Fig. 11a). Located at the head of the Langmoche valley, the moraine dam failed on 4th August 1985 when an ice avalanche from the receding Langmoche Glacier produced a displacement-wave that overtopped the moraine dam and triggered its failure. The resulting glacial lake outburst flood (GLOF) killed five people, livestock, and destroyed valuable arable land as well as a newly-completed hydro-electric power plant (Vuichard and Zimmerman, 1986, 1987; Richardson and Reynolds, 2000). The terminal moraine is composed of a non-cohesive arrangement of locally-transported gravel, cobble, and boulder-sized debris, and measures approximately 650 m wide and 80 m high at its highest point. A sizeable breach, which measures ~40 m high and ~70 m across at its widest point, dissects the northern edge of the terminal moraine.

Following the approach detailed above, a network of 35 GCPs was established across the terminal moraine complex (Fig. 11b), and a set of 1649 photographs taken from various locations and perspectives across the site. With a team of three people, GCP deployment took approximately 10 hours due to the challenging topography and effects of working at high altitude, whilst photograph acquisition was completed in a little over 4 hours.

Feature matching and sparse bundle adjustment on three image batches (*n* = 457, 560, and 609 images apiece) produced a total of 2.2×10^4 points, which, after dense reconstruction and manual editing was increased to 13.2×10^6. The SfM processing took approximately 22 hours per batch, though access to a number of identical machines allowed batches to be processed in tandem. The data were then geo-registered and decimated gridded terrain products derived (producing a final grid of ~3.5×10^5 cells). GPS errors and transformation residuals are shown in Table 2. The final, fully georeferenced DEM is displayed in Fig. 12a.

Although the focus of photograph-acquisition was the terminal moraine and breach, background photographic information was sufficient to reconstruct the entire lake basin, including the 2 km long northern lateral moraine. As in the previous example, significant topographic detail (sub-metre scale) has been resolved. The entire breach was successfully reconstructed, and notable morphological features captured by the model include the narrow central section and expansive exit, as well as two abandoned spillways. Highest point densities are concentrated along the inner faces of the breach (where densities in excess of >8,700 points per m^2 may be found, compared to a site-wide median of 7.35 per m^2; see also Fig. 8c), the eastern limit of the distal face of the northern moraine, as well as the southern face of the relict medial moraine which dissects the terminal moraine complex (Fig. 12b). A number of interpolation artefacts are present across the scene, but are largely confined towards the south and correspond to an extensive area of snow cover.

4.2. Glacially-sculpted bedrock ridge, Cwm Cau

Cwm Cau is a west-east orientated glacial cirque, located immediately to the south of Cadair Idris (893 m) in Snowdonia National Park, Wales, UK (see Fig. 4a for location). It is carved out of folded Ordovician volcanic rocks. A plethora of glacial landforms are found inside the cirque and down-valley, including morainic hummocks and ridges and glacially sculpted bedrock ridges (Sahlin and Glasser, 2008). The latter were deemed suitable for a small-scale appraisal of the SfM technique. The bedrock ridge chosen for reconstruction is oriented west-east, is 80 m in length, 19 m across at its widest point, and approximately 6 m and 8 m high along its southern and northern flanks, respectively (Fig. 13a). Twenty-two orange targets measuring 0.1 m in diameter were used as GCPs. Given the scale of the feature, a relative decrease in target size was deemed appropriate. Using dGPS, horizontal, vertical, and combined positional accuracies of 0.002 m, 0.002 m, and 0.003 m were achieved. A total of 800 photographs were taken and used for scene reconstruction. As in the previous example, SfM processing was performed on three individual batches, taking an average of 12 hours each. Transformation residuals averaged 0.975, 0.161, and 0.422 for *x*, *y* and *z*. Photograph and GCP data were acquired at Cwm Cau in approximately 3 hours using two people, 1.5 hours of which was spent establishing the GCP network.

The final interpolated DEM is shown in Fig. 13b and Fig. 14. Due to the smaller scale of this feature, the point cloud was decimated to 10 cm spatial resolution, reducing the original number of points from 8.9 x 10^6 to 1.3 x 10^5. Analysis of decimated cell statistics revealed an average point density of 69 per 0.1 m^2 (also Fig. 8d) providing outstanding detail of the feature's surface, particularly exposed, bare rock faces found on the southern and northern flanks (Fig. 13b). Indeed, point density on the southern face exceeded 90 points per 0.1 m^3. This extreme resolution is sufficient to delineate centimetre-scale bedrock joints and other surface features such as striations. Scene reconstruction was weaker in topographically complex and partially occluded regions such as the north-eastern edge of the feature (see Fig. 14b), as a result of poorer photographic coverage and shadowing.

5. Discussion

As the above examples demonstrate, the apparent logistical advantages of SfM (limited hardware needs and portability) are, at least in part, offset by the lengthy processing times compared to 'data-ready' methods such as TLS or GPS. Keypoint descriptor extraction, matching, and sparse and dense reconstruction algorithms are computationally demanding. For example, total processing times for a typical photoset used here, numbering between 400-600 images (at 2272 x 1740 pixel resolution), range from 7-56 hours on a 64-bit system with a 2.8 GHz CPU, 512 MB GPU and 6 GB RAM. This is heavily dependent on the complexity of image texture, and may be ameliorated by reducing image size, although this will result in a consequent decrease in the number of returned keypoint descriptors, and ultimately reduced point density. Inevitably, the selection of SfM for any geoscience application must take account of this significant post-processing load and the choice of survey method will eventually be weighed against a number of factors, including cost, accessibility, experience, and fitness-for-purpose in terms of data resolution and coverage.

The sheer size of datasets generated using SfM, and for that matter those derived from allied remote sensing methods such as TLS present significant information management problems (Rychkov et al. 2012). The current generation of GIS are largely inappropriate tools for this purpose, and the effective visualization and analysis of high resolution point clouds is heavily constrained by the limited range of software available for this purpose. The use of the point cloud decimation methods described herein provide a useful strategy to down-scale these dense data whilst retaining information concerning the sub-grid variability. However, while statistical analysis of these decimated data may provide useful insights, for example to quantify surface roughness, this step inevitably results in an unwanted loss of topographic complexity. It is anticipated that future improvements in computational power will reduce run times and facilitate the use of original, high-resolution photosets for input to SfM processing, whilst the emergence of low-cost or freely-available software products (e.g. MeshLab), capable of easily manipulating point clouds far in excess of $>10^6$ of points will improve data handling and general accessibility to the approach.

The minimalistic nature of the approach also lends itself to aerial surveying. Platforms including kites, lighter-than-air blimps and unmanned aerial vehicles (UAVs) are equally portable, relatively inexpensive and capable of carrying both basic and advanced photographic equipment (e.g. Smith et al., 2009; Vericat et al., 2009; Niethammer et al., 2012), thereby vastly increasing the potential areal coverage attained. In addition, aerial datasets have the potential to reduce, or even eliminate the 'dead-ground' problem (Wolf and Dewitt, 2000), whereby objects in the foreground of an image obscure those in the background, resulting in significant data gaps. This issue applies not only to the SfM method, but to all point cloud acquisition methods highly limited by line-of-sight. For topographically simple terrain, such as outwash fans, and valley-floor floodplains, an aerial approach would be particular advantageous. However, as with stereoscopic reconstruction, steep, or near-vertical topography is likely to be problematic for the SfM technique. Further work is on-going to appraise this approach.

The example applications presented in section 4 were ideally suited to the application of the SfM technique. Minimal vegetation coverage and relatively complex, heterogeneous topography at both the meso- and micro-scales facilitate the extraction of suitable numbers of keypoint descriptors for consistent, dense point cloud coverage. Similarly, the method is ideally suited for application in (semi)arid environments. In contrast, the method's suitability for topographic reconstruction of, for example, riparian landscapes may be limited, given that, at present, only water-free surfaces would be suitable for reconstruction, and point density is likely to be limited, and of questionable accuracy, in areas of dense vegetation.

6. Conclusions

This paper has outlined a novel low-cost, ground-based, close-range terrestrial photogrammetry and computer vision approach to obtaining high-resolution spatial data suitable for modelling meso- and micro-scale landforms. The nature of the SfM method eliminates the requirement for manual identification of image control prior to processing, instead employing automatic camera pose estimation algorithms to simultaneously resolve 3-D camera location and scene geometry; this is an extremely significant advantage of the technique over traditional digital photogrammetric methods. However, as the raw SfM output is fixed into a relative co-ordinate system, particular time and attention should be taken in the establishment of a GCP network to facilitate transformation to an absolute co-ordinate system and the extraction of metric data. Taking the hypothesised effectiveness of an aerial approach into account, the terrestrial data collection method presented herein nevertheless represents an effective, financially viable alternative to traditional manual topographic surveying and photogrammetric techniques, particularly for practical application in remote or inaccessible regions.

Acknowledgements

MJW is funded by a NERC Open CASE Award, in association with Reynolds International Ltd. Sam Doyle, Heïdi Sevestre, Pippa Cowley and staff from Summit Trekking, Kathmandu, are thanked for their invaluable assistance with data collection in the Khumbu Himal. James Hickman, Colin Souness and Mark Smith are thanked for their assistance at Cwm Cau and Constitution Hill. TLS point cloud unification was performed by Mark Smith. Joseph Wheaton and an anonymous reviewer are thanked for their suggestions and comments, which significantly improved the original manuscript.

References

Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R., Wu, A.Y., 1998. An optimal algorithm for approximate nearest neighbour searching fixed dimensions. Journal of the Association for Computing Machinery 45, 891-923.

Astre, H. 2010. SFMToolkit3. Available: http://www.visual-experiments.com/demos /sfmtoolkit

Baltsavias, E.P., Favey, E., Bauder, A., Bösch, H., Pateraki, M., 2001. Digital surface modelling by airborne laser scanning and digital photogrammetry for glacier monitoring. The Photogrammetric Record 17, 243-273.

Barker, R., Hooke, J., 1997. Use of terrestrial photogrammetry for monitoring and measuring bank erosion. Earth Surface Processes and Landforms 22, 1217-1227.

Betts, H.D., DeRose, R.C., 1999. Digital elevation models as a tool for monitoring and measuring gully erosion. Journal of Applied Earth Observation and Geoinformation 1, 91-101.

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.

Boike, J., Yoshikawa, K., 2003. Mapping of periglacial geomorphology using kite/balloon aerial photography. Permafrost and Periglacial Processes 14, 81-85.

Boufama, B., Mohr, R., Veillon F., 1993. Euclidean constraints on uncalibrated reconstruction. Proceedings of the Fourth International Conference on Computer Vision, Berlin, Germany; 466-470.

Brasington, J., Smart, R.M.A., 2003. Close range digital photogrammetric analysis of experimental drainage basin evolution. Earth Surface Processes and Landforms 28, 231-247.

Brasington, J., Rumsby, B.T., McVey, R.A., 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.

Chandler, J.H., 1999. Effective application of automated digital photogrammetry for geomorphological research. Earth Surface Processes and Landforms 24, 51-63.

Chandler, J.H., Ashmore, P., Paola, C., Gooch, M., Varkaris, F., 2002. Monitoring river-channel change using terrestrial oblique digital imagery and automated digital photogrammetry. Annals of the Associated of American Geographers 92, 631-644.

Chandler, J.H., Fryer, J.G., Jack, A., 2005. Metric capabilities of low cost digital cameras for close range surface measurement. The Photogrammetric Record 20, 12-26.

Clarke, T.A. and Fryer, J.G. 1998. The development of camera calibration methods and models. Photogrammetric Record, 16, 51-66.

Coveney, S., Fotheringham, A.S. 2011. Terrestrial laser scan error in the presence of dense ground vegetation. The Photogrammetric Record 26, 307-324.

Fischler, M.A., Bolles, R.C., 1987. Random Sample Consensus: a paradigm for model fitting with applications to image analysis and automated cartography. In: Martin, A.F., Oscar, F. (Eds.), Readings in Computer Vision: Issues, Problems, Principles, and Paradigms. Morgan Kaufmann Publishers Inc., London, pp. 726-740.

Förstner, W., 1986. A feature-based correspondence algorithm for image matching. International Archives of Photogrammetry and Remote Sensing, 26: 150-166.

Fraser, C.S., Cronk, S. 2009. A hybrid measurement approach for close-range photogrammetry. ISPRS Journal of Photogrammetry and Remote Sensing 64, 328-333.

Fuller I.C., Large, A.R.G., Milan, D., 2003. Quantifying channel development and sediment transfer following chute cutoff in a wandering gravel-bed river. Geomorphology 54, 307-323.

Furukawa, Y., Ponce, J., 2007. Accurate, dense, and robust multi-view stereopsis. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 17-22 June, Minneapolis, USA, 1-8.

Furukawa, Y., Curless, B., Seitz, M., Szeliski, R., 2010. Clustering view for multi-view stereo. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 13-18 June, San Francisco, USA, 1434-1441

Hancock, G.R., Willgoose, G.R., 2001. The production of digital elevation models for experimental model landscapes. Earth Surface Processes and Landforms 26, 475-490.

Haneberg, W.C., 2008. Using close range terrestrial digital photogrammetry for 3-D rock slope modeling and discontinuity mapping in the United States. Bulletin of Engineering Geology and the Environment 67, 457-469.

Harris, C., Stephens, M.. 1988. A combined corner and edge detector. Proceedings of the Fourth Alvey Vision Conference, Manchester, 147-151.

Heng, B.C.P., Chandler, J.H., Armstrong, A., 2010. Applying close-range digital photogrammetry and soil erosion studies. The Photogrammetric Record 25, 240-265.

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, K., 2009. In-situ characterisation of grain-scale fluvial morphology using Terrestrial Laser Scanning. Earth Surface Processes and Landforms 34, 954-968.

Horn, B.K.P. 1987. Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America 4, 629-642.

Jones, A.F., Brewer, P.A., Johnstone, E., Macklin, M.G., 2007. High-resolution interpretative geomorphological mapping of river valley environments using airborne LiDAR data. Earth Surface Processes and Landforms 32, 1574-1592.

Keim, R.F., Skaugset, A.E., Bateman, D.S., 1999. Digital terrain modelling of small stream channels with a total-station theodolite. Advances in Water Resources 23, 41-48.

Keutterling, A., Thomas, A., 2006. Monitoring glacier elevation and volume changes with digital photogrammetry and GIS at Gepatschferner glacier, Austria. International Journal of Remote Sensing 27, 4371-4380.

Krosley, L.K. Shaffner, P.T., Oerter, E., Ortiz, T., 2006. Digital ground-based photogrammetry for measuring discontinuity orientations in steep rock exposures. Proceedings of the 41st U.S. Symposium of Rock Mechanics, Golden, Colorado, pp. 1-13.

Lane, S.N., 2000. The measurement of river channel morphology using digital photogrammetry. The Photogrammetric Record 16, 937-961.

Lane, S.N., James, T.D., Crowell, M.D., 2000. Application of digital photogrammetry to complex topography for geomorphological research. The Photogrammetric Record 16, 793-821.

Lane, S.N., Chandler, J.H., Porfiri, K., 2001. Monitoring river channel and flume surfaces with digital photogrammetry. Journal of Hydraulic Engineering 127, 871-877.

Lejot, J., Delacourt, C., Piégay, H., Fournier, T., Trémélo, M-L., 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.

Lohani, B., Mason, D.C., 2001. Application of airborne scanning laser altimetry to the study of tidal channel geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing 56, 100-120.

Lourakis, M.I.A., Argyros, A.A., 2009. SBA: A software package for generic sparse bundle adjustment. ACM Transactions on Mathematical Software 36: 1-30.

Lowe, D.G., 1999. Object recognition from local scale-invariant features, International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60, 91-110.

Marzolff, I., Poesen, J., 2009. The potential of 3-D gully monitoring with GIS using high-resolution aerial photography and a digital photogrammetry system. Geomorphology 111, 48-60.

Marzolff, I., Ries, J.B., Albert, K.D., 2002. Kite aerial photography for gully monitoring in Sahelian landscapes . In Proceedings of the Second Workshop of the EARSeL Special Interest Group on Remote Sensing for Developing Countries, 18-20 September, Bonn, Germany, 2-13.

Matthews, N.A., 2008. Aerial and close-range photogrammetric technology: providing resource documentation, interpretation, and preservation. Technical Note 428, Bureau of Land Management, Denver, Colorado.

Microsoft®. 2010. Photosynth™. Available: http://www.photosynth.net.

Niethammer, U., James, M.R., Rothmund, S., Travelletti, J., Joswig, W., 2012. UAV-based remote sensing of the Super Sauze landslide: Evaluation and results. Engineering Geology 128, 2-11.

Nocedal, J., Wright, S.J., 1999. Numerical optimization. Springer: New York.

Notebaert, B., Verstraeten, G., Govers, G., Poesen, J., 2009. Qualitative and quantitative applications of LiDAR imagery in fluvial geomorphology. Earth Surface Processes and Landforms 34, 217-231.

Parsons, D. R., J. L. Best, O. Orfeo, R. J. Hardy, R. Kostaschuk, S. N. Lane., 2005. Morphology and flow fields of three-dimensional dunes, Rio Paraná, Argentina: Results from simultaneous multibeam echo sounding and acoustic Doppler current profiling, Journal of Geophysical Research, 110, F04S03.

Pyle, C.J., Richards, K.S., Chandler, J.H., 1997. Digital photogrammetric monitoring of river bank erosion. The Photogrammetric Record 15, 753-764.

Remondino, F., El-Hakim, S., 2006. Image-based 3-D modelling: a review. The Photogrammetric Record 21, 269-291.

Remondino, F., Fraser, C. 2006. The International Archives of the Photogrammetry and Remote Sensing Commission V Symposium 'Image Engineering and Vision Metrology', 266-272.

Richardson, S.D., Reynolds, J.M., 2000. An overview of glacial hazards in the Himalayas. Quaternary International 65/66: 31-47.

Rieke-Zapp, D.H., Nearing, M.A., 2005. Digital close range photogrammetry for measurement of soil erosion. The Photogrammetric Record 20, 69-87.

Rosser, N.J., Petley, D.N., Lim, M., Dunning, S.A., Allison, R.J., 2005. Terrestrial laser scanning for monitoring the process of hard rock coastal cliff erosion. Quarterly Journal of Engineering Geology and Hydrogeology 38, 363-375.

Rychkov, I., Brasington, J., Vericat, D., 2012. Computational and methodological aspects of terrestrial surface analysis based on point clouds. Computers & Geosciences 42, 64-70.

Sacchetti, F., Benetti, S., Georgiopoulou, A., Shannon, P.M., O'Reilly, B.M. Dunlop, P., Quinn, R., Cofaigh, C.O., 2012. Deep-water geomorphology of the glaciated Irish margin from high-resolution marine geophysical data. Marine Geology 291, 113-131.

Sahlin, E., Glasser, N.F., 2008. Geomorphological map of Cadair Idris, Wales. Journal of Maps 6: 299-314.

Smith, M., Chandler, J., Rose, J., 2009. High spatial resolution data acquisition for the geosciences: kite aerial photography. Earth Surface Processes and Landforms 34: 155-161.

Snavely, N., 2008. Scene Reconstruction and Visualization from Internet Photo Collections, unpublished PhD thesis, University of Washington, USA.

Snavely, N., Seitz, S.N., Szeliski, R., 2006. Photo Tourism: Exploring image collections in 3D, ACM Transactions on Graphics, New York, pp. 1-12.

Snavely, N., Seitz, S.N., Szeliski, R., 2008. Modeling the world from internet photo collections. International Journal of Computer Vision 80: 189-210. Spetsakis, M.E., Aloimonos, Y., 1991. A multi-frame approach to visual motion perception. International Journal of Computer Vision 6: 245-255.

Stojic, M., Chandler, J.H., Ashmore, P., Luce, J., 1998. The assessment of sediment transport rates by automated digital photogrammetry. Photogrammetric Engineering and Remote Sensing 64, 387-395.

Sturznegger, M., Stead, D., 2009. Close-range terrestrial digital photogrammetry and terrestrial laser scanning for discontinuity characterization on rock cuts. Engineering Geology 106, 163-182.

Szeliski, R., Kang, S.B., 1994. Recovering 3-D shape and motion from image streams using nonlinear least squares. Journal of Visual Communication and Image Representation 5: 10-28.

Vericat, D., Brasington, J., Wheaton, J., Cowie, M., 2009. Accuracy assessment of aerial photographs acquiring using lighter-than-air blimps: low cost tools for mapping river corridors. River Research and Applications 25: 985-1000.

Vuichard, D., Zimmerman, M., 1986. The Langmoche flash-flood, Khumbu Himal, Nepal. Mountain Research and Development 6: 90-94.

Vuichard, D., Zimmerman, M., 1987. The 1985 catastrophic drainage of a moraine-dammed lake, Khumbu Himal, Nepal: cause and consequences. Mountain Research and Development 7: 91-110.

Westaway, R.M., Lane, S.N., Hicks, D.M., 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.

Wolf, P.R., Dewitt, B.A., 2000. Elements of Photogrammetry with Applications in GIS (3rd edition), McGraw-Hill, Boston.

Fig. captions:

Table 1. Processing batch description and SfM output for Constitution Hill, Dig Tsho, and Cwm Cau.

Table 2. GCP positional accuracies and transformation residuals (Dig Tsho and Cwm Cau only). TLS system position taken as true for transformation of Constitution Hill SfM data.

Fig. 1. Instead of a single stereo pair, the SfM technique requires multiple, overlapping photographs as input to feature extraction and 3-D reconstruction algorithms.

Fig. 2. Lowe's (2004) Scale Invariant Feature Transform (SIFT) algorithm decomposes a given image (left) into a database of 'keypoint descriptors' (right). Lines represent individual keypoints, proportionally scaled according to the radius of the image region (pixels) containing the keypoint (SIFT code available: http://www.cs.ubc.ca/~lowe/keypoints/).

Fig. 3. From photograph to point-cloud: the Structure-from-Motion workflow.

Fig. 4. Constitution Hill, Aberystwyth. (a) Aerial perspective of the site. Area reconstructed using SfM and TLS is highlighted in red, GCP positions are shown in yellow, and TLS position are shown in blue. Inset map shows relative location of Constitution Hill (CH), and Cwm Cau (CC). Imagery georeferenced to UTM Zone 30N coordinate system. (b) 1 m^2 tarpaulin squares were used as GCPs on Constitution Hill. Centroid positions were recorded using dGPS. (c) A GCP target, as it appears in a photograph, and; (d) as represented in the dense point cloud.

Fig. 5. Perspective views of Constitution Hill. (a) Panorama of the survey area (with GCPs clearly visible), and; reconstructed (b) sparse and (c) dense point clouds. See text for description of A, B, and C in (a).

Fig. 6. Exposed oblique section of a fold in Silurian turbidites at the base of Constitution Hill as captured in (a) an input photograph, and as resolved (b) in the dense point cloud. Prominent (lighter) layers are sandstone, softer (darker) layers are mudstone. See annotation 'C' in Fig. 5 for location.

Fig. 7. Final interpolated DEMs of Constitution Hill using (a) TLS and (b) SfM data; (c) aerial photograph of the site; (d) point density map. **A** and **B** refer to associated labels in Fig. 5a. **VF** and **DV** refer to, respectively, vegetation-free and densely vegetated sub-regions analysed in Fig. 10. Data georeferenced to UTM Zone 30N coordinate system.

Fig. 8. Per-cell point density frequency distribution plots for: (a) Constitution Hill SfM DEM; (b) Constitution Hill TLS DEM; (c) Dig Tsho SfM DEM (note: upper limit of x axis manually truncated at 150 points per m^2 to preserve histogram form; refer to Fig. 12b for complete range of density values); (d) Cwm Cau SfM DEM..

Fig. 9. DEM of difference results (TLS – SfM). (a) Spatial representation of the z_{diff} frequency distribution. Data georeferenced to UTM Zone 30N coordinate system. (b) z_{diff} frequency distribution. **A**: the headwall at the highest point of the surveyed area; **B**: near-vertical cliffs; **C**: dense vegetation cover; **D**: interpolation error.

Fig. 10. Local scale (20 m^2) z_{diff} frequency distribution data for an area with (a) dense vegetation cover, and (b) vegetation-free. As anticipated, elevation differences in the absence of vegetation are considerably lower.

Fig. 11. Dig Tsho glacial lake. **(a)** View towards the terminal moraine complex at Dig Tsho, Khumbu Himal, Nepal. A large breach, produced by a Glacial Lake Outburst Flood is clearly identifiable to the right of the photograph; **(b)** the spatial extent of SfM reconstruction. GCP locations also shown. (photo: M Westoby).

Fig. 12. Dig Tsho SfM data products. (a) Oblique view of the final hill-shaded model of Dig Tsho. Clearly identifiable morphological features include the 1985 breach, as well as two abandoned spillways (AS) which cross the terminal moraine. Current lake extent shown for reference. The up-valley end of the lake has not been fully reconstructed, resulting in interpolation artefacts; (b) oblique view showing per-cell (1 m^2) point densities. Data transformed to UTM Zone 45N geographic coordinate system.

Fig. 13. Glacially-sculpted bedrock ridge, Cwm Cau, Snowdonia. (a) Photograph of the southern face of the feature; (b) hill-shaded DEM, demonstrating the capability of the SfM data to resolve small scale (cm) features.

Fig. 14. Oblique, hill-shaded views of the final SfM DEM Cwm Cau bedrock ridge, showing: (a) the south and (b) north faces. Areas of low point density resulted in extensive interpolation across elevated sections. Data transformed to UK Ordnance Survey National Grid coordinate system (OSGB36 datum).

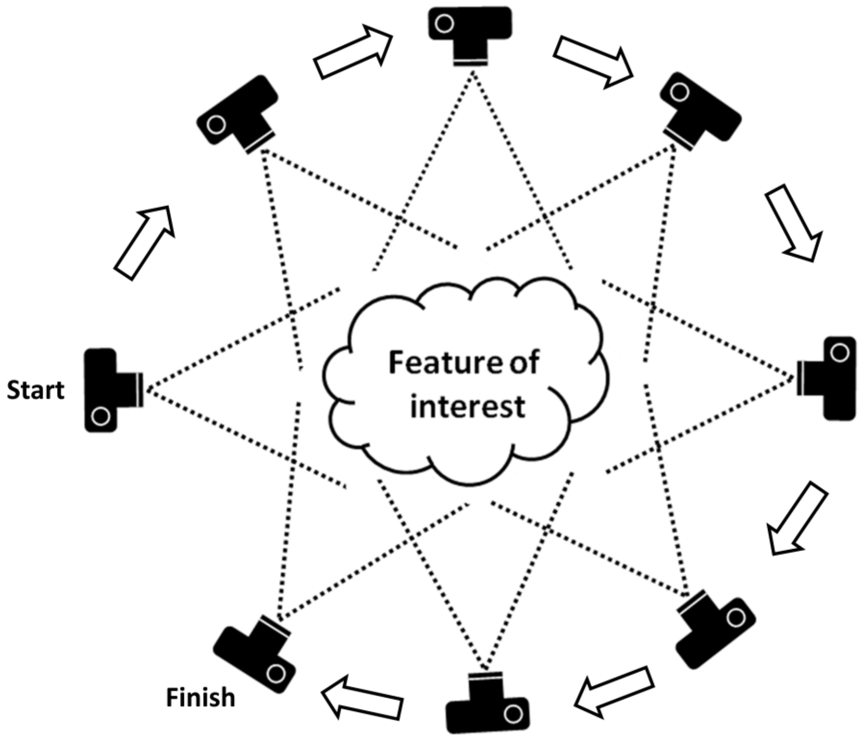

Fig. 1. Instead of a single stereo pair, the SfM technique requires multiple, overlapping photographs as input to feature extraction and 3D reconstruction algorithms.

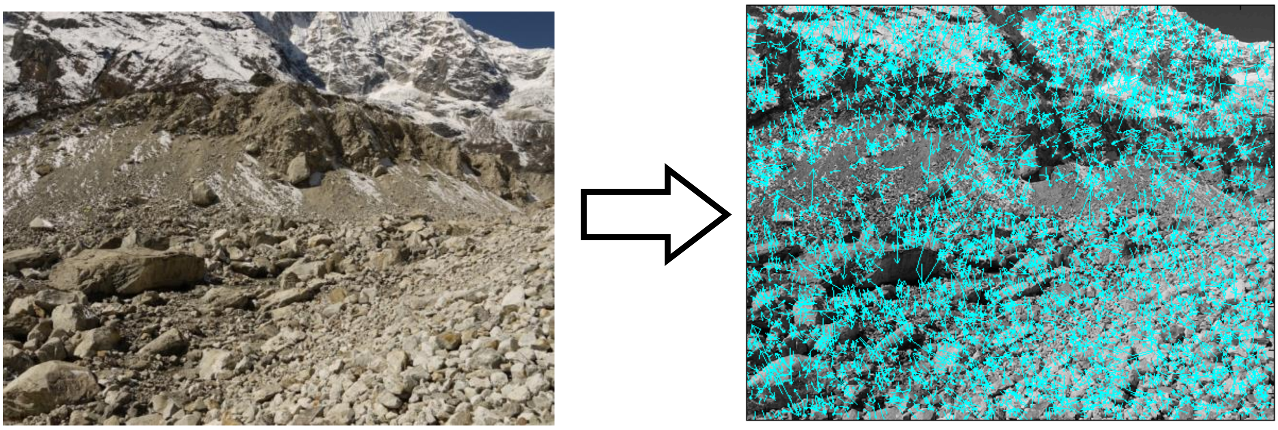

Fig. 2. Lowe's (2004) Scale Invariant Feature Transform (SIFT) algorithm decomposes a given image (left) into a database of 'keypoint descriptors' (right). Lines represent individual keypoints, proportionally scaled according to the radius of the image region (pixels) containing the keypoint (SIFT code available: http://www.cs.ubc.ca/~lowe/keypoints/)

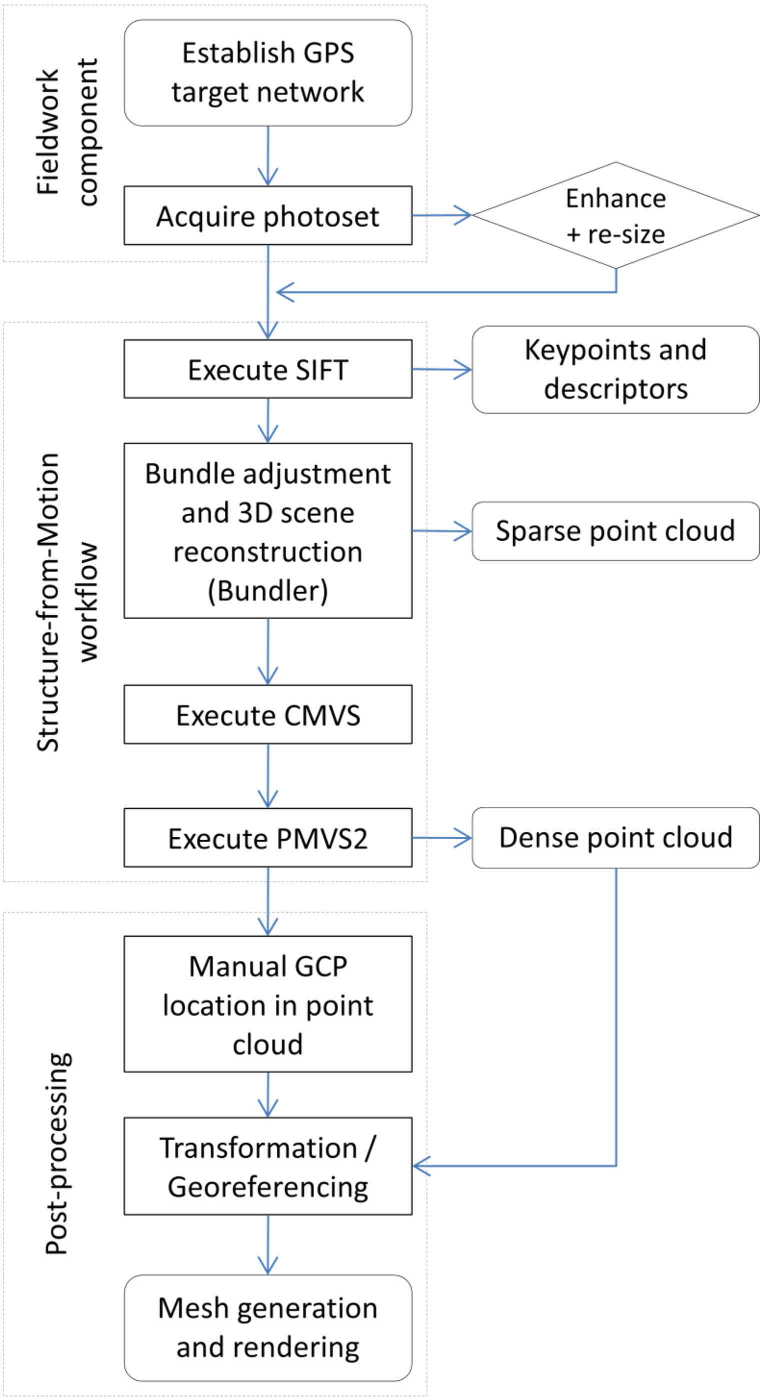

Fig. 3. From photograph to point-cloud: the Structure-from-Motion workflow.

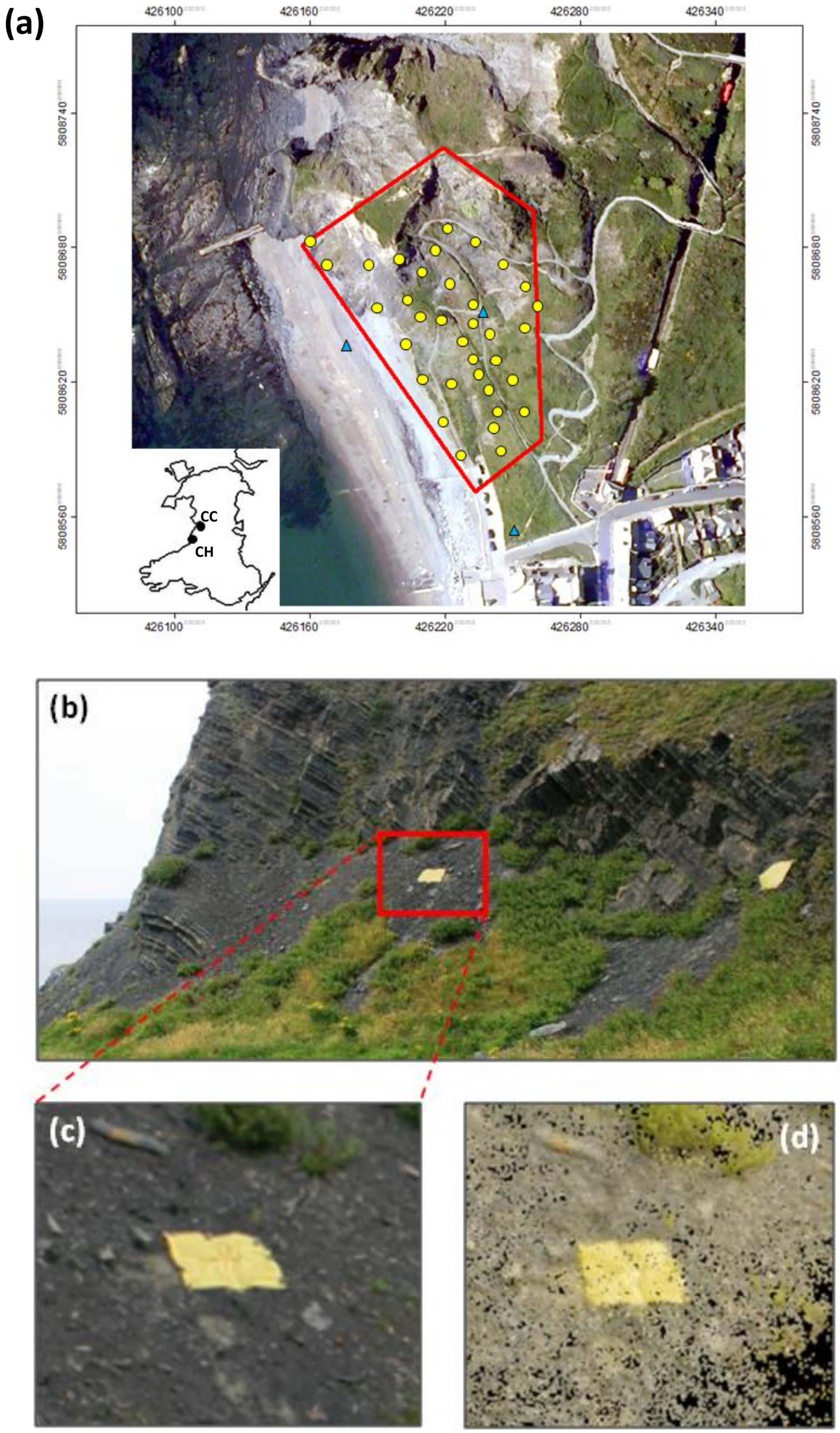

Fig. 4. Constitution Hill, Aberystwyth. (a) Aerial perspective of the site. Area reconstructed using SfM and TLS is highlighted in red, GCP positions are shown in yellow, and TLS position are shown in blue. Inset map shows relative location of Constitution Hill (CH), and Cwm Cau (CC). Imagery georeferenced to UTM Zone 30N coordinate system. (b) 1 m^2 tarpaulin squares were used as GCPs on Constitution Hill. Centroid positions were recorded using dGPS. (c) A GCP target, as it appears in a photograph, and; (d) as represented in the dense point cloud.

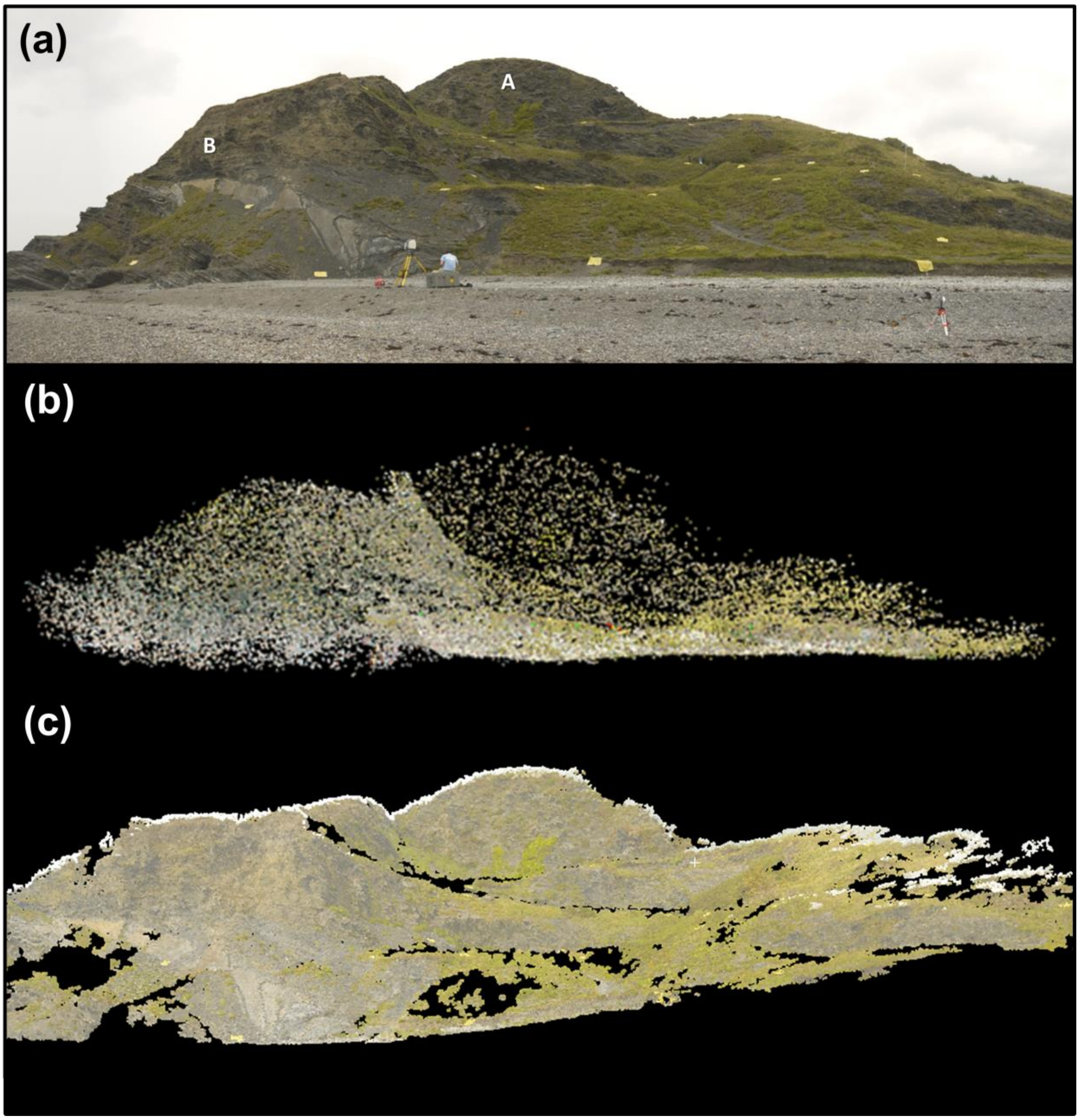

Fig. 5. Perspective views of Constitution Hill. (a) Panorama of the survey area (with GCPs clearly visible), and; reconstructed (b) sparse and (c) dense point clouds. See text for description of A and B in (a).

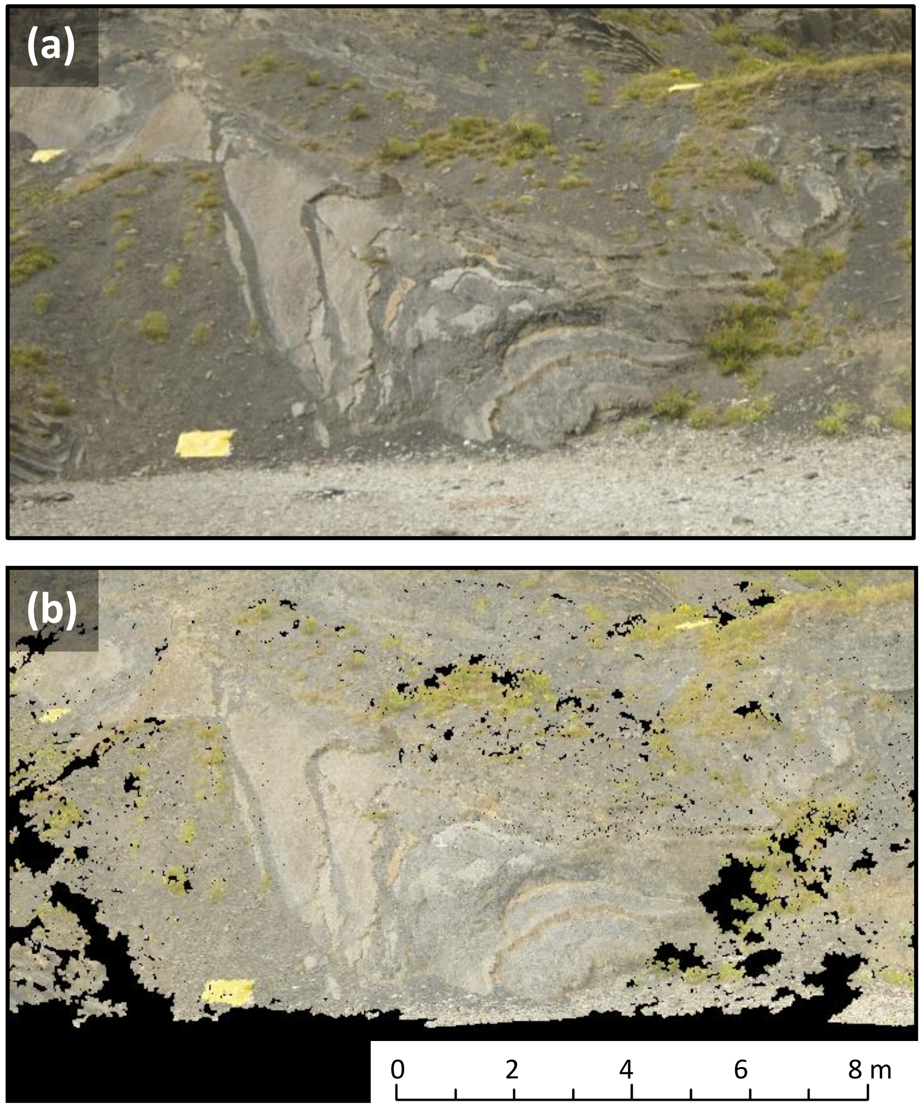

Fig. 6. Exposed oblique section of a fold in Silurian turbidites at the base of Constitution Hill as captured in (a) an input photograph, and as resolved (b) in the dense point cloud. Prominent (lighter) layers are sandstone, softer (darker) layers are mudstone. See annotation 'C' in Fig. 5 for location.

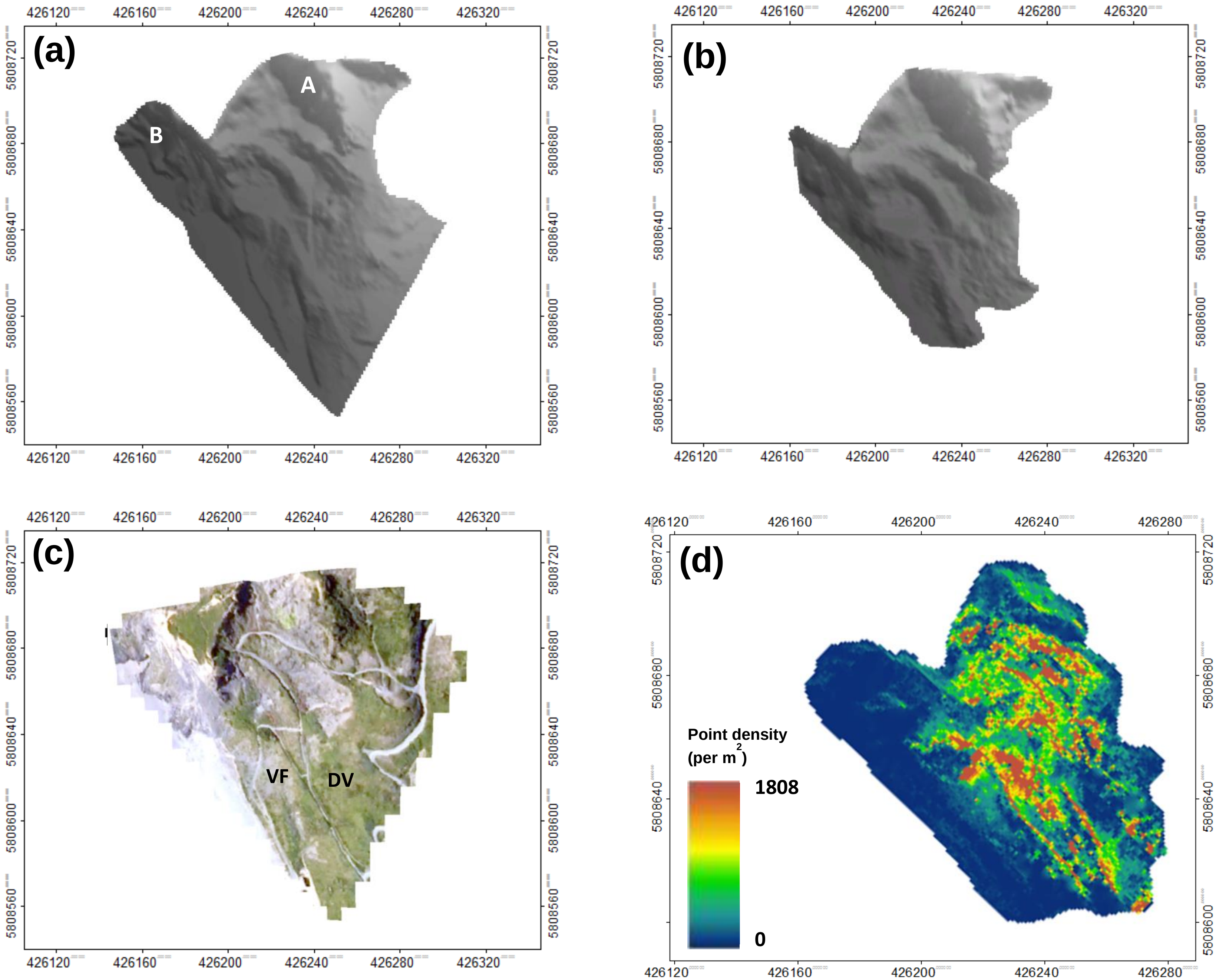

Fig. 7. Final interpolated DEMs of Constitution Hill using (a) TLS and (b) SfM data; (c) aerial photograph of the site; (d) point density map. **A** and **B** refer to associated labels in Fig. 5a. **VF** and **DV** refer to, respectively, vegetation-free and densely vegetated sub-regions analysed in Fig. 10. Data georeferenced to UTM Zone 30N coordinate system.

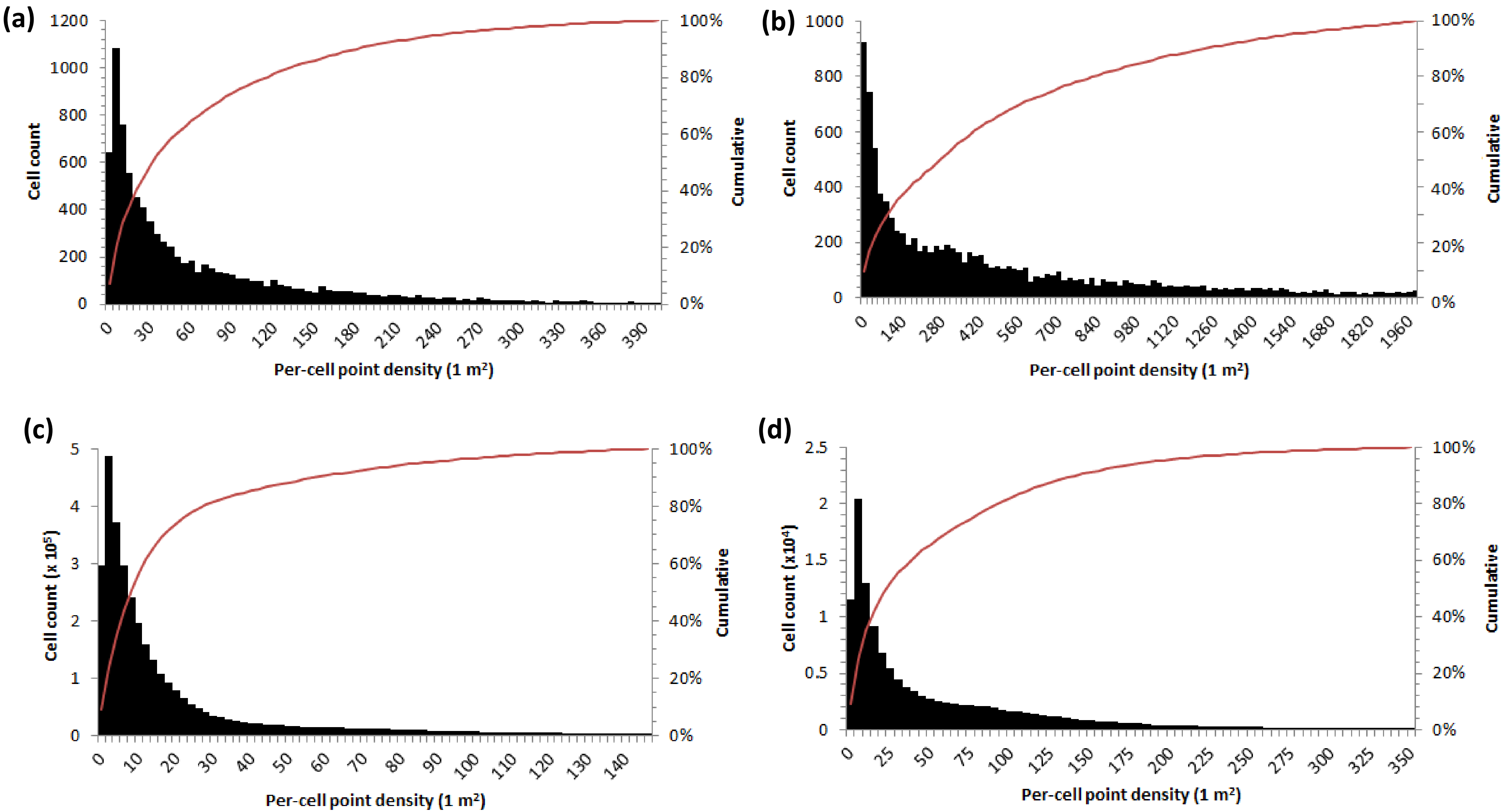

Fig. 8. Per-cell point density frequency distribution plots for: (a) Constitution Hill SfM DEM; (b) Constitution Hill TLS DEM; (c) Dig Tsho SfM DEM (note: upper limit of x axis manually truncated at 150 points per m^2 to preserve histogram form; refer to Fig. 12b for complete range of density values); (d) Cwm Cau SfM DEM.

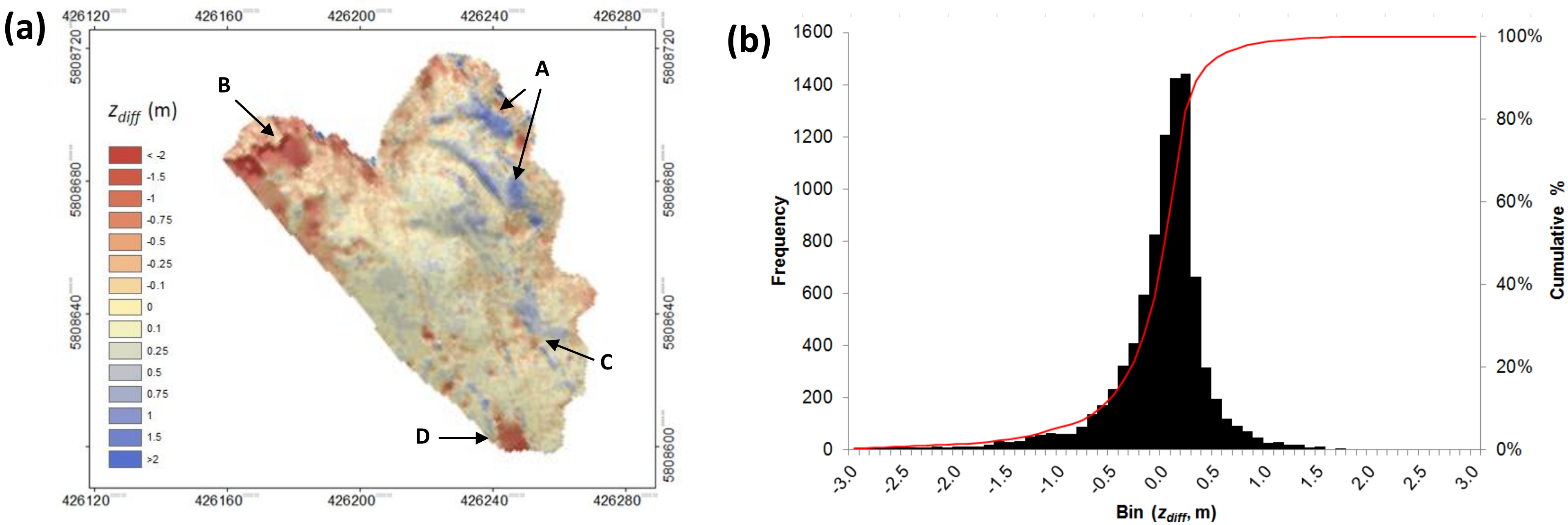

Fig. 9. DEM of difference results (TLS – SfM). (a) Spatial representation of the z_{diff} frequency distribution. Data georeferenced to UTM Zone 30N coordinate system. (b) z_{diff} frequency distribution. **A**: the headwall at the highest point of the surveyed area; **B**: near-vertical cliffs; **C**: dense vegetation cover; **D**: interpolation error.

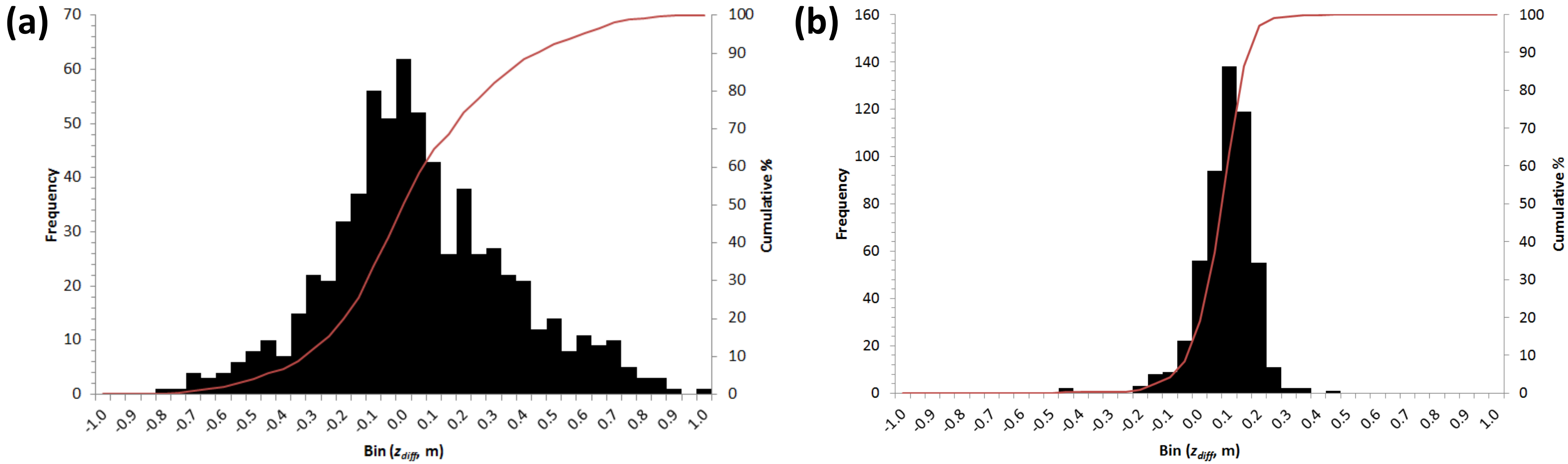

Fig. 10. Local scale (20 m^2) z_{diff} frequency distribution data for an area with (a) dense vegetation cover, and (b) vegetation-free. As anticipated, elevation differences in the absence of vegetation are considerably lower. See **DV** and **VF** in Fig. 7 for locations.

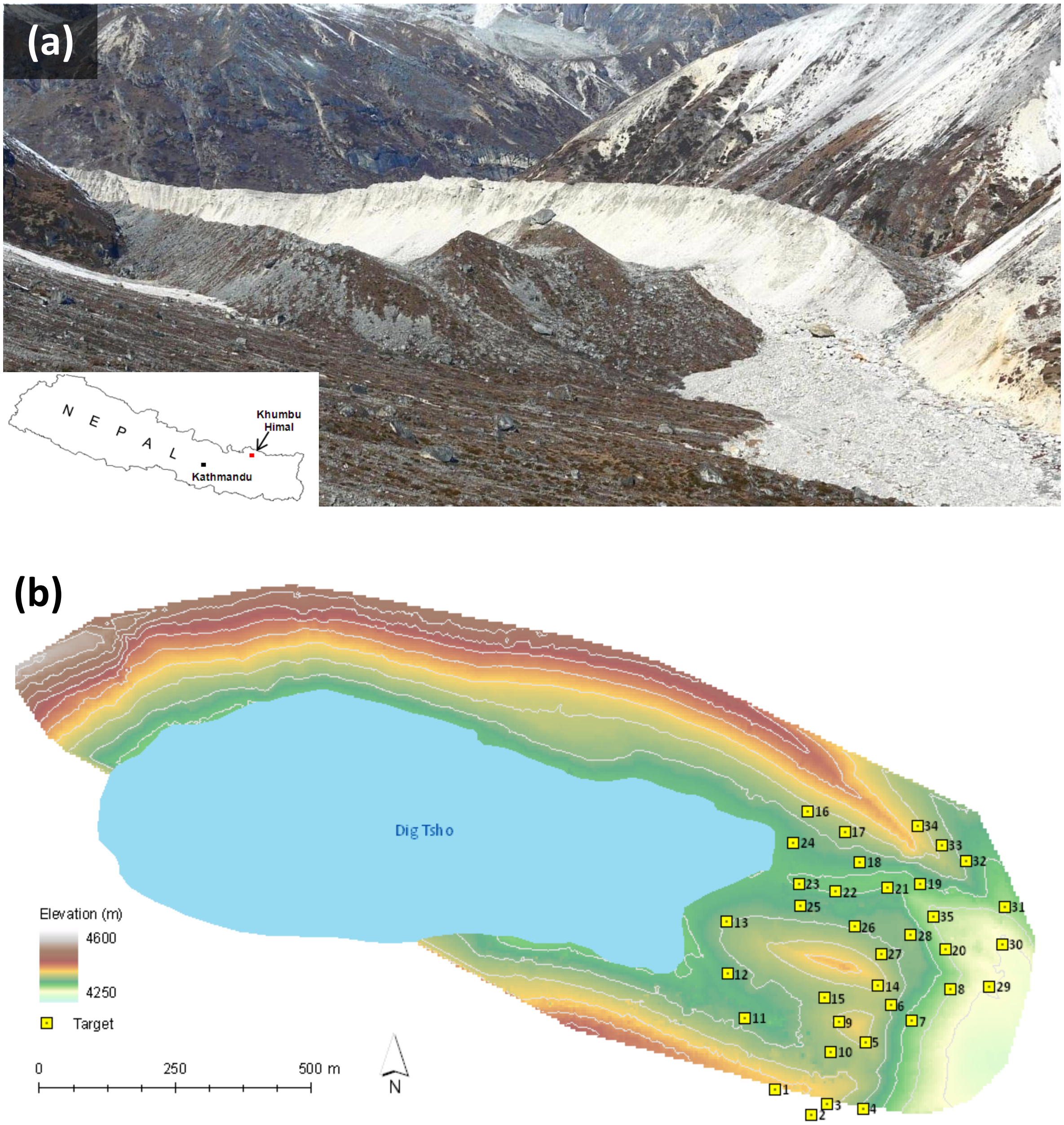

Fig. 11. Dig Tsho glacial lake. **(a)** View towards the terminal moraine complex at Dig Tsho, Khumbu Himal, Nepal. A large breach, produced by a Glacial Lake Outburst Flood is clearly identifiable to the right of the photograph; **(b)** the spatial extent of SfM reconstruction. GCP locations also shown. (photo: M Westoby).

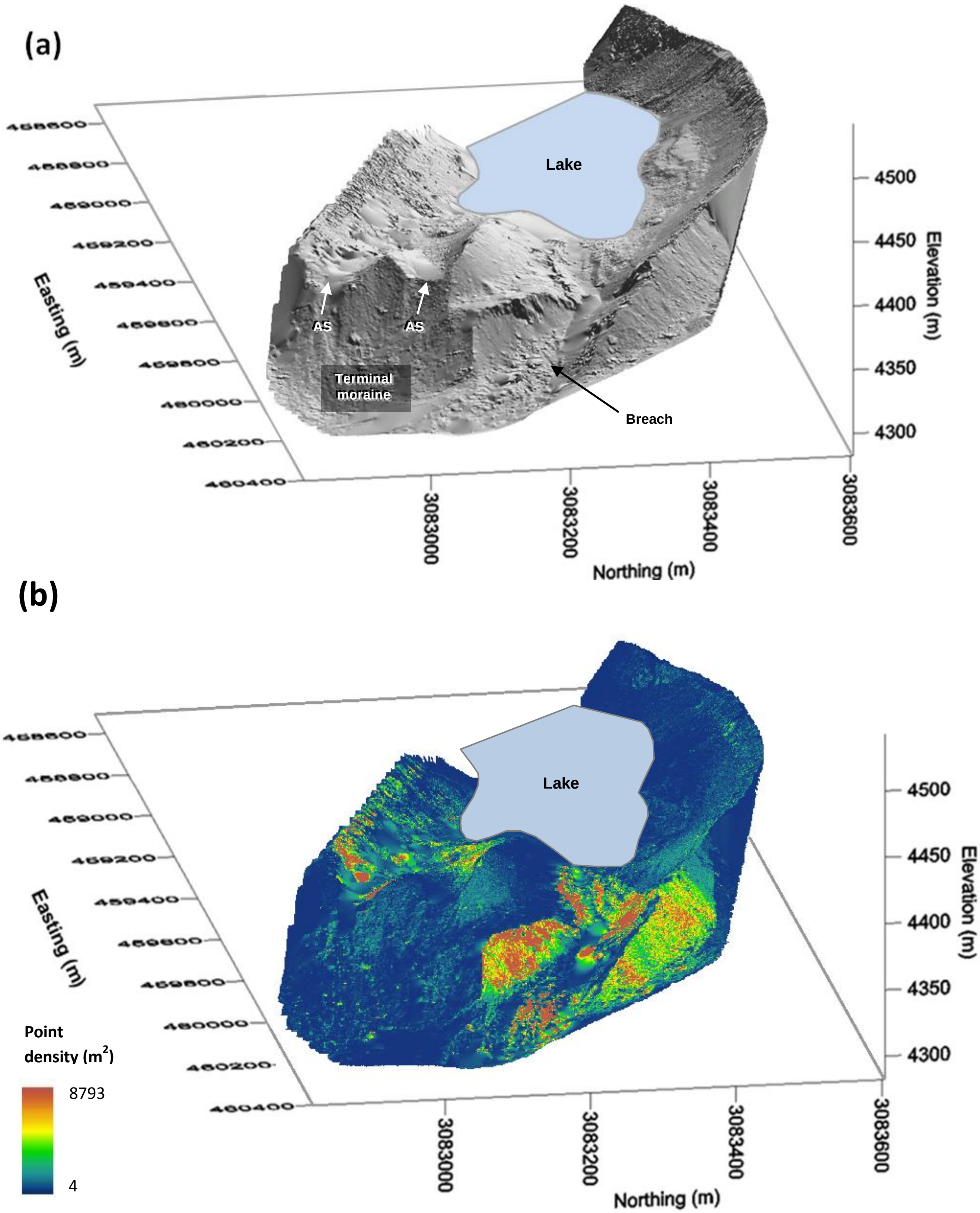

Fig. 12. Dig Tsho SfM data products. (a) Oblique view of the final hill-shaded model of Dig Tsho. Clearly identifiable morphological features include the 1985 breach, as well as two abandoned spillways (AS) which cross the terminal moraine. Current lake extent shown for reference. The up-valley end of the lake has not been fully reconstructed, resulting in interpolation artefacts; (b) oblique view showing per-cell (1 m^2) point densities. Data transformed to UTM Zone 45N geographic coordinate system.

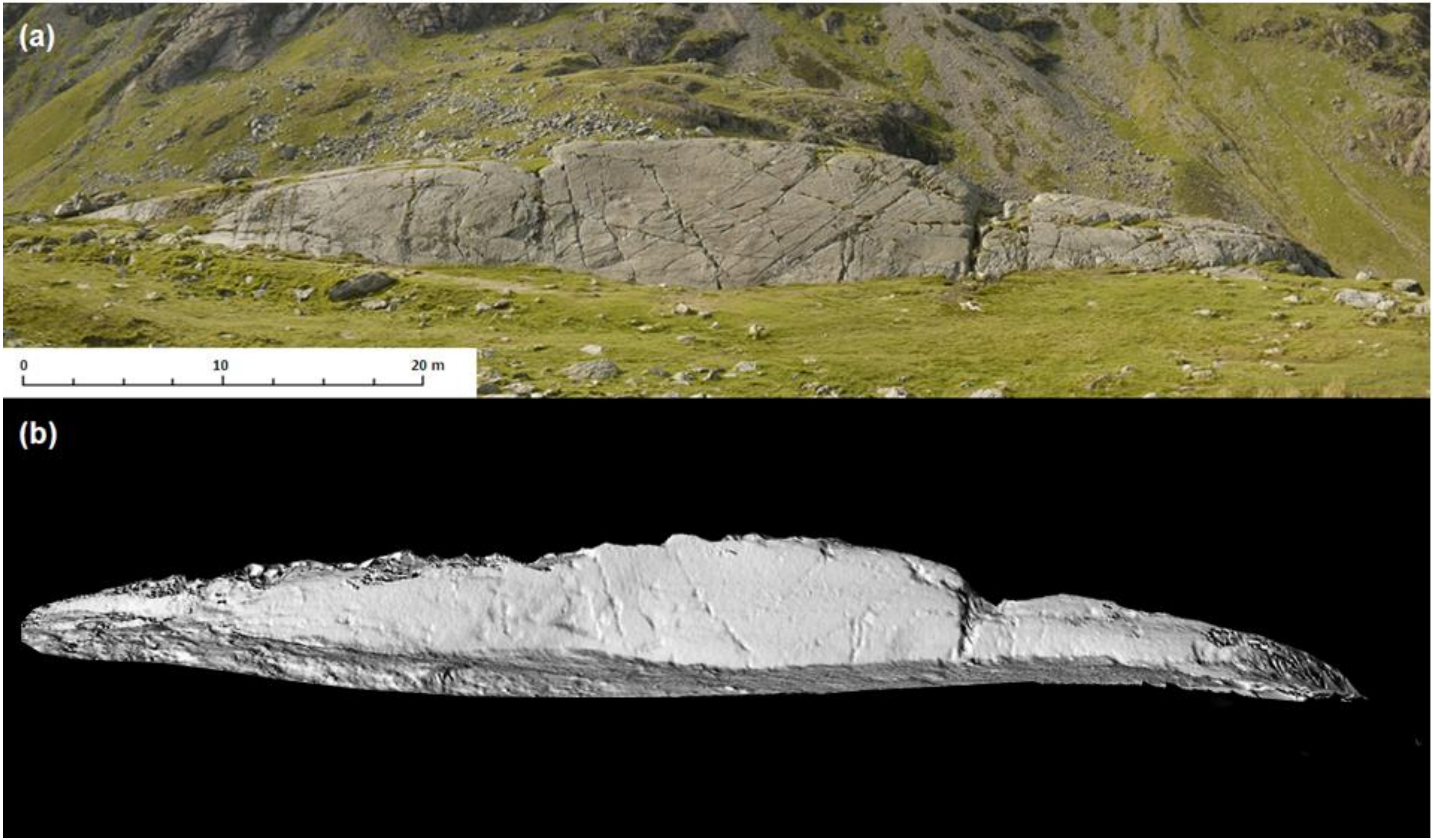

Fig. 13. Glacially-sculpted bedrock ridge, Cwm Cau, Snowdonia. (a) Photograph of the southern face of the feature; (b) hill-shaded DEM, demonstrating the capability of the SfM data to resolve small scale (cm) features.

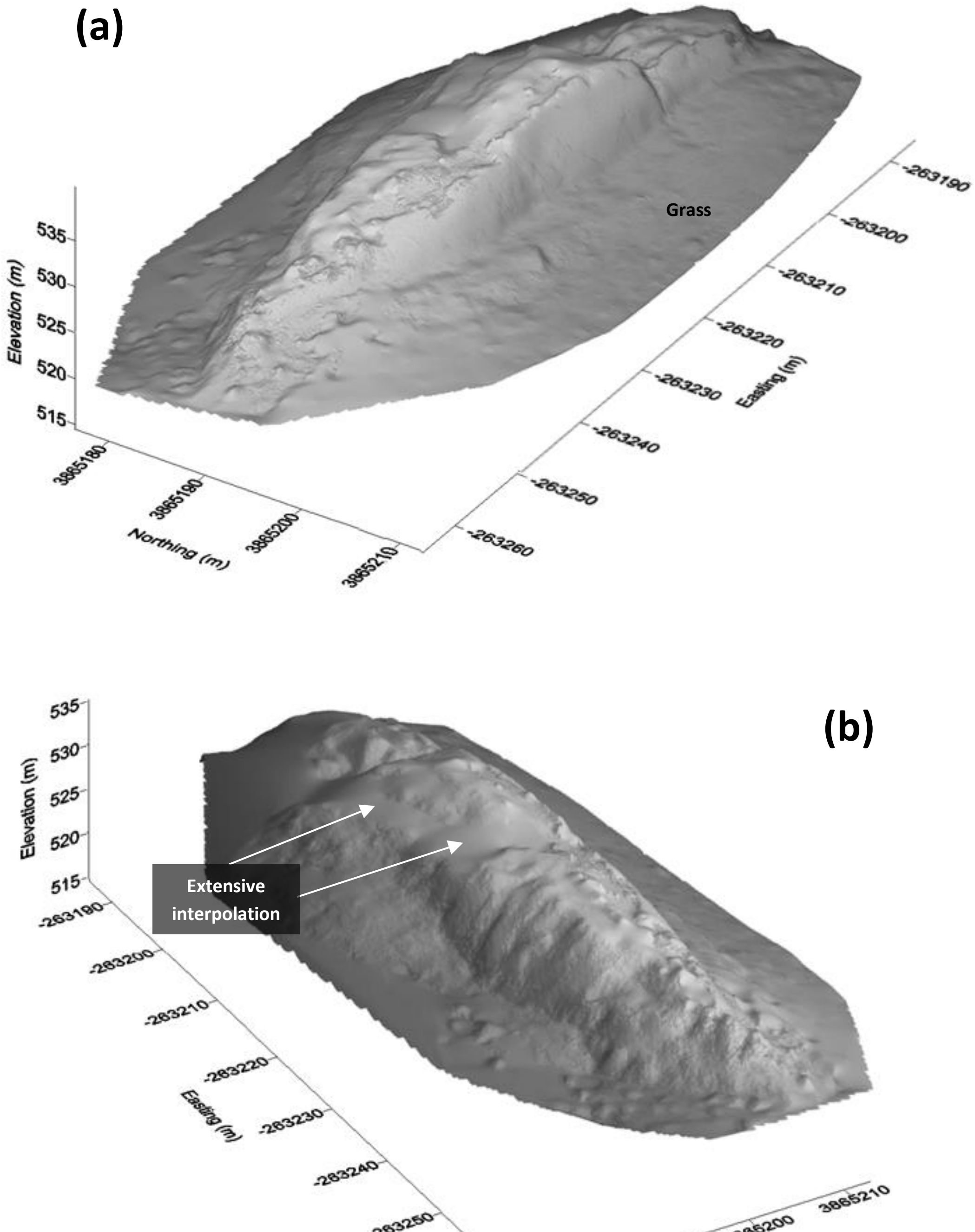

Fig. 14. Oblique, hill-shaded views of the final SfM DEM Cwm Cau bedrock ridge, showing: (a) the south and (b) north faces. Areas of low point density resulted in extensive interpolation across elevated sections. Data transformed to UK Ordnance Survey National Grid coordinate system (OSGB36 datum).

Table 2. Processing batch description and SfM output for Constitution Hill, Dig Tsho, and Cwm Cau.

Site / batch	No. of photos	Sparse points	Dense points
Constitution Hill			
CH_Ba1	159	5.8×10^4	2.9×10^6
CH_Ba2	411	8.2×10^4	5.8×10^6
CH_Ba3	286	3.8×10^4	2.7×10^6
Dig Tsho			
DT_Ba1	457	6.1×10^4	4.3×10^6
DT_Ba2	560	7.2×10^4	1.0×10^7
DT_Ba3	609	9.1×10^4	8.3×10^6
Cwm Cau			
CC_Ba1	264	7.0×10^4	5.1×10^6
CC_Ba2	536	1.2×10^5	7.3×10^6

Table 1. GCP positional accuracies and transformation residuals (Dig Tsho and Cwm Cau only). TLS system position taken as true for transformation of Constitution Hill SfM data.

Site	Ground-truth or GCP uncertainty (m; mean)			Batch	Transformation residuals (m)		
	x + y	z	xyz		x	y	z
Constitution Hill (Wales)	-	-	-	CC_Ba1	0.196	0.156	0.020
				CC_Ba2	0.076	0.064	0.035
				CC_Ba3	0.016	0.072	0.005
Dig Tsho (Nepal)	0.110	0.195	0.226	DT_Ba1	1.022	1.158	2.917
				DT_Ba2	0.630	0.694	3.241
				DT_Ba3	1.015	0.482	4.653
Cwm Cau (Wales)	0.002	0.002	0.003	CC_Ba1	0.538	0.117	0.164
				CC_Ba2	1.411	0.205	0.679