Progress in
Physical Geography

Crowd-sourcing structure-from-motion data for terrain modelling in a real-world disaster scenario: a proof of concept

Journal:	*Progress in Physical Geography*
Manuscript ID	PPG-17-085.R2
Manuscript Type:	Main Article
Keywords:	structure from motion, disaster risk reduction, crowd-sourcing, point cloud, camera phone, sfm, drr, natural hazards, terrain models, dtm

Abstract:	Structure-from-motion (SfM) photogrammetry techniques are now widely available to generate digital terrain models (DTMs) from optical imagery, providing an alternative to costlier options such as LiDAR or satellite surveys. SfM could be a useful tool in hazard studies because its minimal cost makes it accessible even in developing regions, and its speed of use can provide updated data rapidly in hazard-prone regions. Our study is designed to assess whether crowd-sourced SfM data is comparable to an industry standard LiDAR dataset, demonstrating potential real-world use of SfM if employed for disaster risk reduction purposes. Three groups with variable SfM knowledge utilized 16 different camera models, including four camera phones, to collect 1001 total photos in one hour of data collection. Datasets collected by each group were processed using VisualSFM, and the point densities, accuracies, and distributions of points in the resultant point clouds (DTM skeletons) were compared. Our results show that the point clouds are resilient to inconsistency in users' SfM knowledge: crowd-sourced data collected by a moderately informed general public yields topography results comparable in data density and accuracy to those produced with data collected by highly-informed SfM users or experts using LiDAR. This means that in a real-world scenario involving participants with a diverse range of expertise, topography models could be produced from crowd-sourced data quite rapidly and to a very high standard. This could be beneficial to disaster risk reduction as a relatively quick, simple, and low-cost method to attain rapidly updated knowledge of terrain attributes, useful for the prediction and mitigation of many natural hazards.

I Introduction

Many natural hazards have a synergistic, cascading or repetitive character, for example: heavy rains or hurricanes can, within hours or days, cause landslides that will direct subsequent flooding or mass wasting (Wieczorek et al., 2001); lake-dam breakouts are a documented hazard of volcanic eruptions when volcanic flows block a waterway but later fail, resulting in flash flooding (Künzler et al., 2012); perhaps most significantly, hazards commonly reoccur in the same regions, often with little or no respite between events (Dykes and Welford, 2007). In disaster risk reduction (DRR), structure-from-motion photogrammetry (SfM) has great potential to be beneficial as a technology that hastens terrain modelling for these purposes. An up-to-date understanding of terrain can be critical to the timely forecasting of potential natural hazard scenarios: van Westen et al. (2008) argued for "the importance of obtaining imagery as soon as possible after the occurrence of a major triggering event, so that accurate event-based landslide maps can be made, which in turn will make it possible to derive landslide probability maps" (van Westen et al., 2008), and this reasoning can be applied not only to landslides but also floods, volcanic hazards, avalanches, and more. With a step-by-step workflow utilizing mainly free and open source software, an SfM terrain model can be produced from raw data in as little as four to five hours. Additionally, the cost of SfM when compared to LiDAR, professional surveying, or terrestrial laser scanning is minimal: it requires only a camera and a computer.

'Crowd-sourcing' is the act of outsourcing a task to a crowd; in this case, the task being outsourced is SfM image collection. In many situations – for example during an emerging crisis, or in the aftermath of a major event – it may be highly desirable to generate or refresh topographic models rapidly, and without having to wait for experts and equipment to arrive on the scene. Crowd-sourcing has already been trialled in DRR for data analysis. The Humanitarian OpenStreetMap Team used crowd-sourced assessments of satellite imagery to assess damage and guide first responders to areas of need in the wake of Typhoon Haiyan in 2013 (Zastrow, 2014), but this approach is limited by the availability of recent satellite surveys. The USGS has been using social media to elicit participation from civilians in earthquake-prone regions in the "Did You Feel It?" campaign to document the geographic and temporal extent of tremors (USGS, 2013). Building on that, researchers at Stanford have used Twitter data to improve the accuracy of real-time earthquake propagation in ShakeMaps (USGS, 2015).

SfM is an *accessible* alternative to traditional terrain modelling methods due to its (1) affordable cost, (2) low barriers of required expertise, (3) rapid turnaround time, and (4) relative ease of use. For these reasons, it is a good option to consider in disaster-prone regions with limited resources and in regions that would benefit from frequently revised terrain models. While SfM itself is more accessible than expensive and labour-intensive alternatives that cannot be deployed as quickly (e.g. LiDAR, satellite imaging, geodetic surveys, etc.), it can be even more efficient if the input data are collected via crowd-sourcing.

The application of SfM to the geosciences and geohazards is still advancing (Fonstad et al., 2012; Gomez-Gutierrez et al., 2015; James and Robson, 2012; James and Varley, 2012; Micheletti et al., 2015). Crowd-sourcing has potential in these applications primarily because it allows for a greater area of coverage in a lesser amount of time than would be possible by scientifically controlled image collection alone, particularly if unmanned aerial vehicles (UAVs) or other airborne platforms are not available. Though prior studies have shown the SfM utility of smartphone-based photo collection (Micheletti et al., 2015), *crowd-sourcing* imagery remains a new avenue for SfM in the geosciences. This study tests the minimum level of SfM familiarity necessary for crowd-sourcing to optimize input image quality (defined in Section IV) to produce sufficient output terrain models.

The overall objective of this proof of concept study is to test whether crowd-sourced SfM data can produce digital terrain models (DTMs) that are sufficiently complete for the purposes of natural hazard scenario modelling. For this application, we consider the "best" DTMs would be low-cost and quick to produce, with a minimum spatial resolution of 1 data point per 10 m^2 (any finer resolution can always be down-sampled to suit the needs of specific sites or numerical models). We emphasize this point about "best" DTMs because for DRR purposes, topographic data has a minimum requirement for accuracy and resolution, but equally as important is that topography data must be *accessible* to users in the hazard-affected area.

1 Structure-from-motion (SfM)

Structure-from-motion (SfM) is a computer vision technology and a type of digital photogrammetry. It comprises a series of algorithms that cross-correlate points in collections of digital images to create 3D digital recreations of the scene, and it can therefore be used to model topographic surfaces from aerial or ground-level photographs (see Fig. 1). The first widely available SfM software, Bundler (http://www.cs.cornell.edu/~snavely/bundler), was published in 2006 and used for the "PhotoTourism Project," a digital reconstruction of popular landmarks from crowd-sourced photos found on the Internet (Snavely et al., 2006; 2008). Bundler is also freely available in a ready-to-use package online (Harle, 2010). Subsequently, many more SfM algorithms and software have been produced, including Photosynth (now discontinued) (Microsoft, 2008; Microsoft et al., 2010) and VisualSFM (Wu, 2007; 2011; Wu et al., 2011).

[insert Figure 1]

SfM-based software is easier to use than earlier photogrammetric software due to improved automation in processing photographic data. The robust SfM algorithms, including the "scale invariant feature transform" or SIFT algorithm (Lowe, 2004), facilitate processing of photographs from different angles, positions and distances from an object without user intervention.

SfM uses digital photos as input data and can benefit from, but does not require, camera calibration or information about the precise positioning of cameras. The 3D

surface model outputs are initially arbitrarily scaled and oriented, but they can be georeferenced with the use of ground control points (GCPs) or a reference image. If artificial GCP targets have not been deployed, natural features identified in orthophotos or satellite imagery can be used, with their coordinates extracted using GIS software (James and Varley, 2012; Verhoeven et al., 2012; Westoby et al., 2012). In some cases where a reference topographic dataset exists, error in the georeferencing process can be minimized through techniques such as iterative-closest-point (ICP) refinement (Besl and McKay, 1992). Consequently, the spatial and elevational accuracy of SfM digital terrain models (DTMs) will correlate with the accuracy of the reference data. However, the overall completeness of SfM output also depends on the photos used, including the quantity, resolution, focal distance, and most significantly, how comprehensively the photos were matched to one another.

Microsoft's Photosynth, Wu's VisualSFM, and similar proprietary software based on SfM (e.g. Agisoft Photoscan, and dozens more), have been used in diverse applications: in architecture to model buildings without the need for travel (Pomaska, 2009; Remondino et al., 2012; Snavely et al., 2008); in archaeology to create detailed digital copies of relics (Kersten and Lindstaedt, 2012) and to geo-orient and map dig sites (Verhoeven et al., 2012). In geohazards, SfM has been used to map lava dome growth (James and Varley, 2012), monitor landslide dynamics (Lucieer et al., 2013), and assess active lava flow emplacement (Tuffen et al., 2013).

Recent geomorphological applications of SfM have demonstrated that accuracies may be comparable to those more expensive technologies often used as 'best when available' (e.g. LiDAR, terrestrial laser scanning, etc.). These technologies can yield centimetric or even millimetric margins of error. Studies of SfM have shown favourable comparison against terrestrial laser scanning in a variety of geomorphic localities (Westoby et al., 2012), centimetre-scale accuracy and point density similar to LiDAR for a fluvial plain (Fonstad et al., 2012) and meter-resolution digital elevation models (DEMs) for dome growth observation (James and Varley, 2012).

Crowd-sourced SfM has been tested in applications to architecture (Snavely et al., 2006; 2008; 2010), where it yielded digital models visually consistent with the architectural landmarks. Yet, because the aim of creating these digital models was visual completeness, the architectural models were never quantitatively analysed to assess the effects of crowd-sourced images, and regardless, terrain and topography present different challenges. Crowd-sourced SfM in terrain modelling has yet to be either qualitatively or quantitatively assessed in the literature, so this case study examines a mostly gratis SfM workflow as a proof of concept for applications in disaster risk reduction.

II Study area

Our study area was the Agios Georgios crater on Nea Kameni Island, Santorini (Greece) (see Fig. 2). The intra-caldera island of Nea Kameni is the site of the most recent volcanic activity at Santorini. It comprises mainly dacitic lava flows and domes that have gradually emerged above sea level during a series of eruptions since 1570 (Nomikou et al., 2014; Pyle and Elliott, 2006). Its subaerial/submarine morphology and

structures have been recently mapped in detail using a combination of LiDAR and bathymetry (Nomikou et al., 2014). Agios Georgios is a small volcanic explosion crater (36° 24' 16.87" N, 25° 23' 44.32" E) on Nea Kameni (see aerial orthophoto in Fig. 2) (Druitt et al., 1999; Hellenic Cadastre, 2014), formed near the summit of the Georgios Dome, which was extruded during a major eruption from 1866 – 1870 (Fouqué, 1879). It is small, measuring approximately 8250 m^2 (75 meters E-W, and 110 meters N-S), and accessible via a tourist trail that also provides panoramic vistas of the surrounding Santorini caldera (Nomikou et al., 2014). Although the last eruption of Nea Kameni was in 1950, degassing continues through fumaroles and diffuse emissions (Parks et al., 2013; Tassi et al., 2013) near the island's summit, and in 2011-2012 a period of seismic and geodetic unrest (Newman et al., 2012; Parks et al., 2012) highlighted the continuing magmatic activity beneath Santorini.

[insert Figure 2]

The Agios Georgios crater was chosen as a case study site to test crowd-sourced SfM for several reasons. The crater was ideal because the circumnavigable path and unvegetated volcanic terrain allow for 360° of visual continuity when collecting photographs. This minimizes the likelihood of error due to visual discontinuity, allowing us to focus on error due to collected image quality. While Agios Georgios is morphologically straightforward for SfM, it is also well placed for this crowd-sourcing study due to a ready group of participants (University of Oxford undergraduate fieldtrip), and pre-existing high-resolution topography data (LiDAR: Nomikou et al. (2014); Pyle and Elliott (2006)) against which the SfM results from this study could be compared.

III Methods of data collection

Our proof of concept case study on Agios Georgios crater demonstrates the generation of SfM terrain models from crowd-sourced images as an analogue to potential real-world image collection scenarios using SfM technology for purposes in disaster risk reduction. The LiDAR against which the SfM results are compared was sourced from the Airborne Research and Survey Facility (ARSF) data collection mission EU12_12, carried out on 16 May 2012. This mission lasted four hours and the overall cost of data collection was £20,000. Average point density of the mission was 2.1 per m^2 and 2.4 per m^2 for the data subset used in this study. Additional details of LiDAR methodology are presented by Nomikou, et al. (2014).

On a fieldtrip to Nea Kameni Island in September 2013, 17 undergraduates were separated into three groups and asked to participate in image collection. Group A represented the laypeople with negligible knowledge of SfM, Group B was moderately informed about SfM, and Group C represented highly informed people. The experimental directives are summarized in Table 1, and the complete experimental directives can be accessed in supplemental material for this manuscript online. All participants were instructed to use only the information presented in their briefing, to exclude any external information, and not to share information across groups. None of the participants reported a prior knowledge of SfM techniques.

The Oxford undergraduate group involved in this study does not represent the "general public" as a whole, but their knowledge of SfM was probably not significantly different to what may be expected from the rest of the population. Experimental design for this case study attempted to negate cognitive bias by limiting interaction across groups before and during image collection. This case study may not be a strict analogue for crowd-sourcing disaster risk reduction data from the general public, but it presents a representative illustration as a proof of concept study. As the first crowd-sourcing study of its kind, we aim to show the technique's potential, and our study could be beneficially expanded upon by using larger and more diverse participant groups that could incorporate a thorough assessment of users' skill levels in image collection.

Group	# Participants	SfM Knowledge	Example Roles	Methodology
A	8	No familiarity	Tourist, amateur photographer, travel blogger, tour provider	Collect a minimum of 50 photos at random (no further directive)
B	5	Some familiarity	Citizen scientist, science student, observatory intern	Collect a minimum of 50 photos using the 'Rule of 3'(see Section III for details)
C	4	Significant familiarity	Trained SfM user, scientific collaborator	Collect a minimum of 50 photos of terrain in specified field area as described in the field guidebook provided to participants

Table 1. Summary of experimental set up. The complete experimental directives can be accessed in digital supplemental material for this manuscript.

Group A comprised eight students acting as the general public, or laypeople. The directive for Group A consisted of nothing more than instructions to collect a minimum of 50 photos per person. Participants in this group were asked to choose one of four roles representing persons in the scientifically uninformed public. The roles were: a tourist, a local tour provider, an amateur photographer, and a travel blogger. These roles were chosen as representative of the "layperson" demographic because they are likely persons who, in a real-world scenario, would incidentally possess photos of topography (similar set-up to Crandall and Snavely (2012), and Snavely, et al. (2006, 2008, 2010)). Participants were asked which role they selected to portray and how it impacted their approach to photo collection: A 'travel blogger' said she "used occasional filters to make social media posts… tried to include people, ships, tours, etc." A 'local tour provider'

said that her photos included "landscapes that tourists would want to see," while a 'tourist' noted that most of her photos were "silly people photos… having fun." An 'amateur photographer' used many camera setting filters and tried to achieve "arty views."

Group B represented the "citizen scientists" or the moderately informed public. The directive for this five-person group included the 'Rule of 3' - very basic, generally accepted guidelines to achieve satisfactory SfM results:

"When taking your photos, use 'the rule of 3': each point of interest in the photo must appear in a minimum of three photos, from three different perspectives that overlap by at least 60%. Set your camera resolution to 5M or 8M and turn off the image stabilizer setting."

This information was meant to represent the maximum amount of information that could be rapidly absorbed by the public without a high degree of background knowledge. It is a basic directive and reasonably straightforward. There were two example roles for this group, the 'concerned citizen' – a local community member with a vague understanding of volcanic risks, who received the directive from online resources such as Photosynth.net; and the 'intern' – a student or technician-level scientist at the local observatory who was asked to collect images for a group research project. Feedback from the 'interns' included that they "attempted a more scientific approach," and "took photos of the same feature from several angles." The 'concerned citizens' described their approaches as having "photos of everything indiscriminately," and "lots of photos of the crater but with little understanding of geological significance."

The four participants of Group C were given much more information about the project goals and SfM method. The directive for this group was presented in an eight page 'handbook' style format, with an overview of SfM technology, examples of use, specific instructions for how to most effectively employ the method, and a detailed description of the study area and how to access it via circumnavigation. This group described their photo collections as being "methodical," "informed," and "focused on [the volcanic] crater."

Students were instructed to use any available camera to take the photos. This variability was an intentional part of the experimental design, meant to replicate the real-world collection of photos from the general public. The equipment used and images collected are presented in Table 2.

Group	Camera Model	Resolution (MP)	# Photos	
A	Nikon D3100	14	61	388 total
	Fujifilm Finepix JV170	14	54	
	Nikon Coolpix S3300	16	58	

	iPhone 4S (phone)	8	14	
		6	10	
	iPhone 5S (phone)	8	39	
		18	1	
	Panasonic Lumix DMC FZ48	6	62	
	Olympus VG150	12	53	
	Olympus E-PL5	16	36	
	LG-E400 (phone)	3	57	
	iPhone 5 (phone)	8	52	
B	iPhone 5 (phone)	8	109	313 total
	Panasonic Lumix DMC FS35	16	53	
	Olympus X875	8	42	
	Fujifilm Finepix A170 A180	5	73	
	Canon Powershot A460	5	44	
C	Panasonic Lumix DMC S5	16	94	300 total
	Panasonic Lumix TZ20	5	89	
ALL	16 models	8 resolutions	1001 photos	

Table 2. Summary of images collected. Note that while the directive for Groups B and C asked for photos to be 5 or 8 megapixels, not all participants followed these instructions.

Participants were allotted one hour in which to collect a suggested minimum of 50 photos per person. The objective of the experiment was to obtain three distinct data sets, differentiated by users' SfM knowledge, for the purpose of determining whether crowd-sourced SfM for terrain studies is feasible in a real-world image collection scenario. Figure 3 illustrates the types of photos collected by each group.

Cognitive bias is a genuine concern in studies involving human participation; this study attempted to minimize bias by restricting interaction between participant groups. Other sources of potential error in image collection, e.g. poorly focussed or artistically filtered images, are welcomed in this study as they present an opportunity to assess the SfM output as a function of input image quality.

[insert Figure 3]

IV Methods of data analysis

The computing hardware for the analysis was a 64-bit PC running Windows 7 Enterprise (2009) with Intel Core i7-2600 processor, 3.40 GHz CPU, 8.00 GB RAM (with 7.83 GB available).

The collected photos were first randomly culled to equal sets of 300 using the RAND() function in Excel. For Group A, this removed 88 of the collected photos. For Group B this removed 13 collected photos. Group C collected exactly 300 photos so all were used. Group ALL was produced by culling data sets A, B, and C to 100 photos each using RAND() and then aggregating the 300 photos, producing a randomized but evenly distributed selection from the three user groups.

After collection, the quality of the photos was appraised in accordance with the general recommendations for SfM processing: minimal background (including sky), filling the frame with the subject matter (in this case, the crater), suggested resolution of 5M to 8M, no cropping or colorized filters, and no or minimal foreground distractions. Figure 3 provides illustrative examples of the types of photos seen in each Group's image set. As expected, Group A photos barely met the general requirements for SfM, Group B photos mostly met the requirements, and Group C photos nearly fully met the requirements.

Each of the four image sets were then processed using the SfM software VisualSFM version 0.5.22 (Wu et al., 2011). VisualSFM was chosen as the software for SfM analysis because it is free (as compared to paid-for software e.g. Agisoft Photoscan), making it accessible to even low-resource regions. Compared to "black-box" programs such as Photosynth, VisualSFM allows the user a flexible degree of control over image processing, although not as controlled as the proprietary PhotoScan. Programs that run solely in command line (e.g. Bundler) can be intimidating to new users and VisualSFM's graphical user interface (GUI) increases accessibility by reducing barriers to non-expert use. The GUI shows the locations of each photo when it was taken, a benefit that can be useful or simply interesting for a user to know. For these reasons, VisualSFM provides a good application for controlled SfM processing in a potential real-world scenario.

The VisualSFM workflow consists of several phases of the structure-from-motion process after photos are uploaded: feature identification, feature matching across photos, and 3-D reconstruction of points (refer to Fig. 1). The final phase yields a 'sparse point cloud' of data points. This is not a gridded DTM, but rather a distribution of points in 3-D space. A DTM is produced later by interpolating between points in the point cloud to yield a regular grid. Although point density and spatial resolution in the resultant DTM can be improved with continuation of the workflow to the 'multiview stereo' (MVS) or 'dense reconstruction' stage, our workflow was considered complete after the sparse reconstruction. MVS uses the SfM findings to seed much more thorough pixel-by-pixel matching (Furukawa, 2010; Furukawa et al., 2010). For this study, SfM sparse point clouds were considered adequate because the parameters of interest were comparative

point density and accuracy – measures that can be compared across either sparse or dense clouds for all datasets – and the processing time for SfM without MVS is much faster, therefore more useful in DRR for reasons outlined in Section I. The first section of Table 3 summarizes the SfM process for each Group.

[insert Figure 4]

The following analysis methods are presented in flow chart form in Figure 4, which illustrates two parallel workflows for SfM georeferencing: one workflow in the free program CloudCompare, the other in the paid program ArcGIS.

Point clouds from groups B, C, and ALL were edited in the open source software program Meshlab (Cignoni et al., 2008). In Meshlab editing, outlying points were removed using the point-picker tool, and the arbitrary coordinate systems of the SfM point clouds were re-oriented to real world X/Y/Z axes using the axis rotation tool.

CloudCompare analysis began with coarse point cloud matching of SfM datasets to the LiDAR data, cropped to the region of interest (see Figure 5). Since the LiDAR is georeferenced, using it as the reference layer in the registration process will result in a georeferenced point cloud for each study group. Georeference refinement in CloudCompare involved three steps: first, the "match bounding box center" tool was used, followed by the "match scales" tool, which utilized the LiDAR as a reference and principal component analysis as the matching criterion. The third step was fine registration using iterative closest point (ICP) analysis in the "fine registration" tool based on the algorithm pioneered by Besl and Mckay (1992). Aligned and registered point clouds are shown in Figure 7, with accompanying error values in Table 4.

The datasets in this study did not include ground control points (GCPs). Considering the end objective of the study was to assess SfM's utility to disaster risk reduction, this presented an opportunity to explore GCP-independent methods of georeferencing that would theoretically be applicable to any other site. While the CloudCompare alignment and registration functions yielded acceptable errors for DRR (see Table 4), we also wanted to demonstrate a GCP-free georeferencing workflow for real-world applications which might be more familiar to users less acquainted with SfM. As ArcGIS is a program commonly used for georeferencing and can be applied to reference data beyond point clouds (e.g. orthophotos), it is important to explore workflows that apply to this and other GISs.

For conversion into DTMs, the edited SfM point clouds were exported from Meshlab as .las files, imported into ArcGIS as a .las dataset (ESRI, 2017), and converted to a triangulated irregular network (TIN) in the ArcGIS 3D Analyst toolbox. The LiDAR data were similarly cropped to the study area's extent, and converted to a TIN. (For comparison purposes, the orthophoto from the LiDAR mission was used to draft the extent parameters for the study area based on the crater outline, which is illustrated in Figure 5). TIN files were rasterized using natural neighbors sampling in ArcGIS for georeferencing against LiDAR in ArcGIS (Sibson, 1981).

[insert Figure 5a]

[insert Figure 5b]

Using the georeferencing toolbar in ArcGIS, the LiDAR data were used as a reference surface on which each SfM group layer was individually fit to display, and the auto-registration function in ArcGIS was used to generate control points based on spectral signatures (ESRI, 2017a). Depending on the layer, between 4 and 6 control points were automatically generated (Figure 6). The DTMs were georeferenced using the "adjust" transformation for continuous data. The adjust transformation combines a polynomial transformation based on a global least-squares fitting (LSF) algorithm along with a local TIN interpolation technique (ESRI, 2017b).

The transformed and referenced point clouds generated in CloudCompare were exported to new .las files, re-imported to ArcGIS, and new TINs were generated to represent the new and adjusted values as a proof of concept. The resulting TINs can be compared to the LiDAR reference in Figure 9.

Model error following the CloudCompare and ArcGIS techniques summarized above is discussed at length in earlier work, and sources cited therein (Aguilar et al., 2006; Erdogan, 2009; Micheletti et al., 2015; Raaflaub and Collins, 2006). There exist many methods for DTM interpolation and georeferencing, etc., each with its own uncertainty considerations. This study is less concerned with absolute error from data manipulation, and more concerned with relative error between different image collections. Relative errors in Z were assessed through the maps of elevational difference (with respect to the LiDAR data) for each data set (Figure 8). RMSEs from CloudCompare alignment and registration are presented in Table 4. Table 5 shows values generated from the RMSE function in ArcGIS, which averages a sample of 5000 data points (only 3 to 8% of total points for these datasets). It is possible that the selection of these points could affect error analysis based on the distribution of the points within the cloud (see Figure 8 for maps of point accuracy distribution). A full uncertainty analysis is beyond the scope of this work.

[insert Figure 6]

V Results

1 VisualSFM results

Individual image sets resulted in multiple point cloud models in VisualSFM. Multiple models are produced in SfM when some images cannot be matched to others within the set. Only the most complete model of each image set was subsequently analysed as 'usable output'. The input images, and direct and usable output results were related to each other in order to produce 'derived output' values (see Table 3).

To quantify the differences between the SfM datasets, several known values are compared in Table 3. *SfM runtime* is the amount of time that each Group's image set took

to process in VisualSFM (prior to point cloud editing and analysis in other programs). *Usable output* acknowledges that not all the results from VisualSFM are pertinent to the study: only a subset of photos from the input images are ultimately used, and only a subset of the points in the point cloud are representative of the study area. *Derived output* relates input values to output. *Utilization* is calculated as: the number of photos contributing to the model / the number of input photos, (e.g. For Group B, 224 / 300 = 0.747); *percent outliers* is a percentage calculated as: (the total initial number of points in a model – the number of usable points after editing) / (the number of points in model), (e.g. For Group B, (60,948 – 59,709) / 60,948 = 0.02); and *density* is a measure of the number of usable points after editing per m^2 in the study area (e.g. For Group B, 59,709 / 8250 m^2 = 7.2 m^{-2}).

Category	Criterion	A	B	C	ALL
SfM runtime (min)	ID time	4	3	3	5
	Image match time	156	144	237	167
	3D reconstruction time	17	6	9	15
	Total time	177	153	249	187
Usable output	# Photos in model	N/A	224	299	214
	# Points in model	N/A	60,948	162,839	97,029
	# Usable edited points	N/A	59,709	159,220	93,631
Derived output metrics	Utilization ratio[a]	N/A	0.747	0.997	0.713
	Percent outliers[b]	N/A	2.0%	2.2%	3.5%
	Density[c]	N/A	7.2 m^{-2}	19.3 m^{-2}	11.3 m^{-2}

Table 3. Summary of image analysis. Note that N/A in the column for Group A is due to the 300 random Group A photos not having produced a usable SfM output model.
[a] How many photos from the input data sets were used to construct the model used in analysis.
[b] How many points were removed from the point cloud in the editing process.
[c] Number of usable edited points per m^2 of study area (8250 m^2).

It is interesting to note that Group A, when processed as a whole (388 photos) produced a sparse and incomplete point cloud of the crater, but it was definitely recognizable. Upon removing 88 photos (any 88 photos, as the authors ran several iterations of RAND() for Group A), the VisualSFM output for this group became

unrecognizable. This reflects that for purely "incidental" photographs of a given area, a substantial number of photos were required to produce a point cloud of the area. An interesting topic of further study would be to assess the relationship between incidental photos and the fewest number of photos necessary to yield recognizable results in various circumstances.

Group ALL, containing 100 photos from each A, B, and C, produced a usable model but integrated the lowest number of photos in its image set. This model was nearly on par with the photo utilization seen in Group B, but slightly lower, likely due to the inclusion of photos from Group A. Interestingly, with a utilization ratio of 0.713, we know that Group ALL did incorporate some of the 100 photos from Group A (else we would have expected utilization to remain below 0.667). This indicates that while Group A alone may have produced unserviceable point clouds, there was still valuable data captured in the photos, which may have just needed stronger cohesion across the image set to produce a recognizable point cloud.

Also interesting to note is that while utilization ratios varied across the SfM datasets, the relative proportion of outliers (points that were edited out from the cloud) remained relatively consistent. Ranging only from 2.0 to 3.5% of the total points per point cloud, it indicates that VisualSFM does quite a good job of eradicating false matches even with broad variability in datasets. Surprisingly, Group C actually contained proportionally more outliers than Group B. The authors attribute this to areas on the perimeter of the point cloud: the rim of the crater was captured in more detail by Group C, and a higher proportion of these points were often confused with background sky points. Most were therefore removed in MeshLab, although Figures 8 and 9 still show some noticeable effects of noise, especially at the south end of the crater.

In VisualSFM, the point clouds for Groups B, C, and ALL had point densities 3 to 9 times greater than provided by LiDAR (see Table 6). These point densities can be visualized in Figure 7 where the SfM point clouds are aligned to the LiDAR. It is likely that some of this success is due to the shape of the study area: a circumnavigable crater of small-ish dimensions (8250 m^2) with no obfuscating vegetation facilitates image collection. Further work could explore the limits of the crowd-sourced technique over variable baselines from the subject matter, and apply crowd-sourcing to a variable range of "real-world" challenges in study areas, for example: issues with line of sight, heavy vegetation, or photos collected obliquely (or otherwise not the convergent photo collection made possible by the crater in this study).

[insert Figure 7]

2 Alignment and georeferencing

The use of iterative closest point analysis (ICP) is well-explored in the SfM community, and we followed the procedure of Micheletti, et al. (2015), using CloudCompare align and registration functions to finely match the SfM datasets to the LiDAR data (refer to Figure 7 above). Our procedure differed in that the initial stage of course alignment was selected to match SfM datasets to the LiDAR based on the centre

of gravity of the point cloud. In the absence of ground control points (GCPs), this was a sensible approximation to use as a starting point.

	A	B	C	ALL
RMSE (meters)	N/A	1.05	0.88	0.88

Table 4. Point clouds' RMSE – root mean squared error (absolute magnitude) between SfM point clouds and LiDAR as a result of alignment and registration in CloudCompare.

For the resultant RMSE after coarse alignment and refined ICP registration, we found that SfM datasets C and ALL were just barely sub-metric in error. We believe this to be a function of the point density distribution in the point cloud itself (as compared to Group B), which would have affected the centre of gravity alignment approximation that preceded ICP.

A challenge in a real-world scenario for SfM image collection is the absence of known GCPs. Theoretically, a location may contain landmarks with known coordinates, but it is equally as feasible that a hazard would alter these natural GCPs. In this study, we circumvented the need for manually identified GCPs, even those naturally occurring. LiDAR can be used as reference material in both CloudCompare and ArcGIS (or other GIS), and this study explores georeferencing in both programs.

	A	B	C	ALL
RMSE (meters)	N/A	0.22	0.34	0.21

Table 5. DTMs' RMSE – root mean squared error between DTMs (derived from SfM) and LiDAR data, as measured in ArcGIS.

In ArcGIS, we found that RMSE was resoundingly sub-metric for all datasets. As compared to CloudCompare alignment and registration, including ICP, the GIS tools yielded lower RMSE across all datasets. Importantly, the RMSEs in Tables 4 and 5 are not exactly comparable – Table 4 represents error for point cloud to point cloud (SfM-LiDAR) analysis, whereas Table 5 represents DTM to DTM (SfM-LiDAR) analysis. Still, the overall trends are interesting: Group ALL came out ahead in both error analyses, although Group C had lower error in its point cloud, whereas Group B had lower error in the DTM. It's possible that the rasterization of data infilled portions of the Group B point cloud to lower overall error for the DTM, and it is also possible that the outlying points in Group C (discussed in Section 3) introduced higher error to the DTM.

We are not suggesting that either method would be appropriate for all geoscience applications, as applications requiring centimetric accuracy will still require expensive equipment such as a differential GPS to measure GCP coordinates. Still, our results suggest that either CloudCompare or GIS-based georeferencing could be used with reasonable confidence for future DRR work where control points are not available.

3 Z-accuracy, systematic error, and completeness

Systematic error in the image sets was minimized through collection of convergent photos. As a function of the circumnavigable crater morphology, photos were largely directed inwards towards the crater, particularly in image sets for Groups B and C. This minimized the systematic 'doming' error found in parallel photo collection studies (James and Robson, 2014). Tables 4 and 5 show RMSE following point cloud registration and DTM georeferencing, respectively.

Overall Z-accuracy can be assessed through a cloud-to-cloud or DTM-to-DTM comparison and a standard method has yet to be determined (Smith et al., 2016). In this study, we used CloudCompare to draw z-accuracy maps of each irregularly distributed SfM point cloud compared to the evenly distributed 2.4 points per m^2 LiDAR data. While RMSE as discussed in the previous section is an averaged measure across the full set of points, z-accuracy maps show how elevational accuracy in various segments of the field area responded to data collection and analysis. This can help to illuminate the best uses of SfM for real-world terrain mapping.

[insert Figure 8]

As we expect based on point cloud density, the cloud-to-cloud z-accuracy map for Group B has a few areas of missing data throughout the crater. Note that all groups have an area of no data to the northwest of the crater, which correlates to the location of the footpath. Group C shows a map of greatest completion, and fewest areas without data. However, Group C also shows the greatest z-error with respect to the LiDAR reference surface (mean error 0.175 meters, standard deviation 1.2 meters). In Figure 9, we can see that this can be attributed to outlying points above the plane of the surface. One limitation of the method used for this study is the manual outlier removal in MeshLab, which can be difficult when the undesirable points lie within the interior of the crater. A next step for this dataset would be to employ automated outlier detection and a smoothing function prior to DEM interpolation (as the TIN visualizations in Fig. 9 show, outlier points are indeed visible).

As with RMSE analysis, Group ALL once again performed best in z-accuracy mapping. The Gaussian curves accompanying the maps show the mean for each dataset (Group ALL is closest to 0 meters, at a mean z-error of 0.039 meters), and the kurtosis of each curve quantifies the most apparent visual property of the maps: Group ALL demonstrates more precise z-accuracy about the mean than either Group B or Group C (s.d. 0.94, 1.06, and 1.2 meters, respectively).

VI Discussion

Allowing for site-specific geomorphologies, sub-metric DTM data for the purposes of hazard modelling is generally unnecessary due to the fact that the overall accuracy of the numerical simulations themselves is not yet at a sub-metric standard. Thus, the SfM results (without dense MVS reconstruction) produced by mixed or moderately informed users (all image sets except Group A) are presumed to be more than

adequate for the purposes of numerical hazard simulations in DRR for the terrain investigated in this work (see Point Densities listed in Table 6).

Based on other studies, DTMs of 10-m to 1-m spatial resolution are presumed sufficient for disaster risk reduction purposes: in simulations using the numerical model LAHARZ to simulate the natural hazard of lahars, DTMs of 10-m and 1-m resolution produced insignificant changes to the model outputs, and variability in results was more directly related to site morphology or input parameters of the LAHARZ simulation (e.g. flow volume) (Huggel et al., 2008; Munoz-Salinas et al., 2009; Stevens et al., 2003); for granular flow modelling using numerical simulation TITAN-2D, 5-m and 10-m resolution DTMs were found to yield similar output, with acceptable results from 30-m data, and unacceptable results with anything coarser (Capra et al., 2011); in a study of slope failures and debris flows using LAHARZ, the source of greatest uncertainty was not DEM resolution, but rather volumetric estimates of the events (Magirl et al., 2010); similarly, assessment of a major flooding event concluded that uncertainty in output had more to do with the hydraulic model than the 1-m DEM (Roca and Davison, 2010); and finally, floodplain modelling scenarios were shown to be similar at 3 to 5-m spatial resolution, with beneficial applications at up to 10-m (Charrier and Li, 2012).

TIN visualizations of the LiDAR and SfM datasets are presented in Figure 9, without smoothing applied, to provide a general sense of how crowd-sourced images could rapidly produce DTMs for hazard modelling.

[insert Figure 9]

Group	SfM Familiarity	Point Density (m^{-2})	Density Distribution and Completeness[1]	Z-Accuracy Ranking[2]	Total Time[3] (hours)	Total Cost[4] (estimated)	DRR Utility Ranking
A	*(1st)* None	*(5th)* N/A	*(5th)* N/A	*(5th)* N/A	*(5th)* N/A	£2000	*5th*
B	*(2nd)* Some	*(3rd)* 7.2	*(4th)* Slightly uneven	*3rd*	*(1st)* 4.3	£2000	*2nd*
C	*(4th)* High	*(1st)* 19.3	*(2nd)* Relatively uniform, fully complete	*4th*	*(3rd)* 5.9	£2000	*3rd*
ALL	*(3rd)* Mixed	*(2nd)* 11.3	*(3rd)* Relatively uniform, minor gaps	*2nd*	*(2nd)* 4.9	£2000	*1st*

LiDAR	*(5th)* Expert	*(4th)* 2.4	*(1st)* Very uniform	*1st*	*(4th)* 6.5	£25000	*4th*

Table 6. Comparison of LiDAR and SfM datasets' utility to the field of disaster risk reduction (DRR). A scoring system of 1 through 5 is used to rank each dataset, averaged for a final DRR utility ranking.
[1]*Density distribution and completeness* is based on Figures 7 and 8.
[2]*Z-accuracy ranking* refers to the z-accuracy maps in Figure 8.
[3]*Total time* is the approximate time to collect and process data, then convert to a terrain model.
[4]*Total cost* is an order-of-magnitude rough estimation covering the general cost of equipment, labour, and computing resources for that particular method of data collection. In this table we have accounted for the use of licensed software ArcGIS, but a workflow exclusively built on free platforms would further lower the cost of SfM.

Table 6 shows the SfM datasets compared to LiDAR, in terms of utility for real world use in disaster risk reduction (DRR). We have considered in the assessment of *utility* all of the categories across the top row. By our estimation, SfM images from a mixture of sources (Group ALL) yield the most useful topographic data that can be produced for DRR purposes. Due to the increased z-accuracy error in Group C, which necessitates smoothing prior to interpolation, the slightly faster processing time and roughly comparable data density make Group ALL an attractive alternative. For deployment in a fast-paced hazard-prone region, the cost and labour-intensive nature of LiDAR were considered to be significant in comparison to SfM, leading to its lower ranking in Table 6.

Further, Group ALL exhibits better value, in that a subset of its 300 photos are incidental and the utilization ratio of all photos (Table 3) was lower. This indicates that in a real-world scenario, Group ALL uses less input data to achieve better output data, and for that reason is considered better value and therefore better utility than Groups B or C. As with any cost-benefit analysis, the returns outpacing the input is the ultimate tie-breaker.

The most obvious benefit of the GCP-free georeferencing methods in this study is the suitability to crowd-sourced data where expertly deployed GCPs are unavailable. However, something to consider is that the static scene in this study is at an advantage compared to a pre/post disaster terrain comparison. Without known control points, a greater degree of uncertainty is unavoidable when comparing post-disaster topography to its pre-disaster reference data, as we cannot be certain what has or has not changed. There may be strategies to mitigate uncertainty – for example, using a larger study area that encompasses not only the disaster-affected region, but also surrounding regions that remain unchanged. This could potentially anchor a post-disaster SfM image set within the proper X/Y/Z positions, then allowing for direct comparison of the pre/post image sets.

In terms of scale, crowd-sourced SfM is presented here for a small study area as a proof-of concept. Collecting adequate coverage of larger field areas would present a new challenge for continued applications of crowd-sourced topography data, but evidence

from other fields of study suggests that it should ultimately prove possible for geohazards. At the city scale, 100 years' of historical photographs have been used for '4D' digital reconstructions of Atlanta (Schindler et al., 2007), and crowd-sourced photos and textual annotation are also used in an augmented reality application where users share information about points of interest (Ioannidi et al., 2017). Most interestingly, kilometre-scale urban and architectural scenes have been successfully reconstructed with SfM by using photos crowd-sourced from the internet (Crandall et al., 2013). With the preponderance of photos shared on social media platforms (times of disaster being no exception), it would be fascinating to create SfM reconstructions of larger-scale disaster sites using internet photos, and compare resultant DTMs against more traditional topography data.

In terms of potential for real-world applications, we are not suggesting SfM as a new best-practice when options such as LiDAR are available, but rather as an approach to supplement non-existent or lower resolution topographic data. In low-resource but hazard-prone regions, topographic maps are in some cases not kept current due to the prohibitive cost of data collection and processing. Frequently, the best data may be a medium resolution contour map, or the most recent global satellite survey (coarse resolution). For these regions, SfM presents a suitable alternative for ad hoc data supplementation.

SfM is suitable to community involvement in regions that may not have a scientific team omnipresent. While it may be unadvisable that people venture to hazardous regions in order to collect images, discussions at the United Nations' Global Platform for Disaster Risk Reduction in 2013 showed that local communities feel deeply invested in hazard scenarios and are eager to participate (UN-ISDR, 2013). One example of trained citizens cooperating with scientists is the *vigias* in Ecuador, who serve as local volcano monitors reporting to the regional monitoring body (Sword-Daniels et al., 2011); similar cooperative efforts may prove beneficial in other areas. Further, the UN-ISDR Hyogo Framework for Action, and its successor HFA2 (Sendai Framework for Disaster Risk Reduction 2015-2030), identified 12 key components for increasing resilience to natural hazards, key among which are fluidized communication between involved entities, and active participation from stakeholders (UN-ISDR, 2013). In a world where communication and participation are paramount, involving stakeholders in a process as important as image collection could be immensely beneficial to not only scientific analyses, but overall greater resilience in DRR.

VII Conclusions

Due to the rapidly evolving, concomitant, or recurrent nature of natural hazards, the field of disaster risk reduction must utilize tools and techniques that assist in rapid assessment of hazard scenarios. SfM is a technique that can be used to model topography, and it is cheaper and quicker than topography-modelling alternatives such as LiDAR. With crowd-sourcing, SfM can be even more rapidly deployed and over larger areas than through expert administration alone.

While earlier work developed SfM-based methodologies using crowd-sourced photos of architectural landmarks, this study shows that crowd-sourced images can also be used to model terrain. Crowd-sourcing in the geosciences is an exciting next step for SfM practitioners, and we hope that this proof of concept study will encourage further exploration of this technique. This study has demonstrated that although the quality of collected images may vary widely, the results of SfM still produce point clouds cohesive enough for further topographic analysis. Importantly, these results indicate that an image set of mixed user knowledge (Group ALL) will tend towards being more complete and more indicative of cohesive input image subsets, rather than not.

Additionally, this study has applied a georeferencing technique that bypasses the need for manually selected ground control points (GCPs) – a necessity in a real-world analogue. We have demonstrated that this technique is not as accurate as dedicated GCP-based georeferencing, but that it is sufficiently accurate for the purposes of DRR. One addition we hope to make in the future is to replace the use of ArcGIS with a free and open source GIS platform, ensuring that the entire workflow is accessible in regions with high need but limited resources.

This proof of concept study shows that crowd-sourced SfM is sufficiently robust to produce topographical models with a data density on par with or exceeding that of similar LiDAR surveys. The findings show that average user ability in SfM image collection yields better-than-average results, indicating that if just a subset of the input images is high quality, the output results for a small study area will have sub-metric resolution and accuracy. Using SfM for DRR has the advantages of being ten times cheaper than LiDAR and at least 25% faster from image collection to finished DTM. Additionally, it supports the process of DRR as defined by the UNISDR (UN-ISDR, 2013), particularly in that it increases stakeholder participation and fluidizes communications between stakeholders. With additional studies on utility and the mechanics of deployment, crowd-sourced SfM could become a useful tool for minimizing disaster in an increasingly digital world.

References

Aguilar FJ, Aguilar MA, Agüera F, et al. (2006) The accuracy of grid digital elevation models linearly constructed from scattered sample data. *International Journal of Geographical Information Science* 20: 169–192.

Besl PJ and McKay ND (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2): 239–256.

Capra L, Manea VC, Manea M, et al. (2011) The importance of digital elevation model resolution on granular flow simulations: A test case for Colima volcano using TITAN2D computational routine. *Natural Hazards* 59: 665–680.

Charrier R and Li Y (2012) Assessing resolution and source effects of digital elevation models on automated floodplain delineation: A case study from the Camp Creek Watershed, Missouri. *Applied Geography* 34: 38–46.

Cignoni P, Corsini M and Ranzuglia G (2008) Meshlab: an open-source 3d mesh processing system. *Ercim News* 73: 47-48.

Crandall D and Snavely N (2012) Modeling people and places with internet photo collections. *Communications of the ACM* 55: 52–60.

Druitt TH, Edwards L, Mellors RM, et al. (1999) *Santorini Volcano*. Geological Society of London Memoirs, vol. 19.

Dykes AP and Welford MR (2007) Landslides in the Tandayapa Valley, northern Andes, Ecuador: implications for landform development in humid and tectonically active mountain ranges. *Landslides* 4: 177–187.

Erdogan S (2009) A comparision of interpolation methods for producing digital elevation models at the field scale. *Earth Surface Processes and Landforms* 34: 366–376.

ESRI (2017) ArcGIS Desktop 10.5.1.

ESRI (2017a) Georeferencing a raster automatically. *desktop.arcgis.com* Available from: http://desktop.arcgis.com/en/arcmap/10.5/manage-data/raster-and-images/georeferencing-a-raster-automatically.htm (accessed 11 May 2018).

ESRI (2017b) Fundamentals for georeferencing a raster dataset. *desktop.arcgis.com* Available from: http://desktop.arcgis.com/en/arcmap/10.5/manage-data/raster-and-images/fundamentals-for-georeferencing-a-raster-dataset.htm (accessed 11 May 2018).

Fabris M, Menin A and Achilli V (2011) Landslide displacement estimation by archival digital photogrammetry. *Italian Journal of Remote Sensing / Rivista Italiana di Telerilevamento* 43: 23–30.

Fonstad MA, Dietrich JT and Courville BC (2013) Topographic structure from motion: a new development in photogrammetric measurement. *Earth Surface Processes and Landforms* 38: 421-430.

Fouqué F (1879) *Santorin et ses éruptions*. Paris: Masson et compagnie.

Furukawa Y (2010) Clustering Views for Multi-view Stereo (CMVS). *www.di.ens.fr*. Available from: http://www.di.ens.fr/cmvs/ (accessed 16 May 2013).

Furukawa Y, Curless B, Seitz SM, et al. (2010) Towards Internet-scale Multi-view Stereo. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 1–8. doi: 10.1109/CVPR.2010.5539802

Giradeau-Montaut D, Bougacha S, Bey A, et al. (2007) CloudCompare.

Gill JC and Malamud BD (2014) Reviewing and visualizing the interactions of natural

hazards, *Reviews of Geophysics* 52: 680-722 doi: 10.1002/2013RG000445

Gomez-Gutierrez A, Juan de Sanjose-Blasco J, Lozano-Parra J, et al. (2015) Does HDR Pre-Processing Improve the Accuracy of 3D Models Obtained by Means of two Conventional SfM-MVS Software Packages? The Case of the Corral del Veleta Rock Glacier. *Remote Sensing* 7: 10269–10294.

Harle J (2010) Tactical Space Lab | Bundler Photogrammetry Package. *tacticalspace.org*. Available from: http://tacticalspace.org/archives/bundler-photogrammetry-package/ (accessed 16 May 2013).

Hellenic Cadastre (2014) National Cadastre & Mapping Agency S.A. *ktimatologio.gr*. Available from: http://www.ktimatologio.gr/sites/en/Pages/Default.aspx (accessed 18 December 2014).

Huggel C, Schneider D, Miranda PJ, et al. (2008) Evaluation of ASTER and SRTM DEM data for lahar modeling: A case study on lahars from Popocatepetl Volcano, Mexico. *Journal of Volcanology and Geothermal Research* 170: 99–110.

Ioannidi A, Gavalas D and Kasapakis V (2017) Flaneur: Augmented exploration of the architectural urbanscape. In: IEEE, pp. 529–533.

James, MR and 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

James MR and 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.

James MR and Varley N (2012) Identification of structural controls in an active lava dome with high resolution DEMs: Volcan de Colima, Mexico. *Geophysical Research Letters* 39(22): 1–5.

Kersten TP and Lindstaedt M (2012) Image-based low-cost systems for automatic 3D recording and modelling of archaeological finds and objects. In: Ioannides M, Fritsch D et al., (eds), Progress in Cultural Heritage Preservation. EuroMed 2012. Lecture Notes in Computer Science 7616, 1-10.

Künzler M, Huggel C and Ramírez JM (2012) A risk analysis for floods and lahars: Case study in the Cordillera Central of Colombia. *Natural Hazards* 64: 767–796.

Lowe DG (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60: 91–110.

Lucieer A, de Jong SM and Turner D (2013) Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Progress in Physical Geography* 38: 97–116.

Magirl CS, Griffiths PG and Webb RH (2010) Analyzing debris flows with the statistically calibrated empirical model LAHARZ in southeastern Arizona, USA. *Geomorphology* 119: 111–124.

Micheletti N, Chandler JH and Lane SN (2015) Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms* 40: 473–486.

Microsoft (2008) Photosynth. *photosynth.net*, Microsoft Live Labs. Available from: http://photosynth.net/Background.aspx.

Microsoft, Argyros AA and Lourakis MIA (2010) Horizon matching for localizing unordered panoramic images. *Computer Vision and Image Understanding* 114: 274–285.

Munoz-Salinas E, Castillo-Rodriguez M, Manea V, et al. (2009) Lahar flow simulations using LAHARZ program: Application for the Popocatepetl volcano, Mexico. *Journal of Volcanology and Geothermal Research* 182: 13–22.

Newman AV, Stiros S, Feng L, et al. (2012) Recent geodetic unrest at Santorini Caldera, Greece. *Geophysical Research Letters* 39: L06309 doi: 10.1029/2012GL051286

Nomikou P, Parks MM, Papanikolaou D, et al. (2014) The emergence and growth of a submarine volcano: The Kameni islands, Santorini (Greece). *GeoResJ* 1-2: 8–18.

Parks MM, Biggs J, England P, et al. (2012) Evolution of Santorini Volcano dominated by episodic and rapid fluxes of melt from depth. *Nature Geoscience* 5: 749–754.

Parks MM, Caliro S, Chiodini G, et al. (2013) Distinguishing contributions to diffuse CO2 emissions in volcanic areas from magmatic degassing and thermal decarbonation using soil gas Rn-222-delta C-13 systematics: Application to Santorini volcano, Greece. *Earth and Planetary Science Letters* 377: 180–190.

Pomaska G (2009) Utilization of Photosynth Point Clouds for 3D Object Reconstruction. *22nd CIPA Symposium*, Kyoto, Japan.

Pyle DM and Elliott JR (2006) Quantitative morphology, recent evolution, and future activity of the Kameni Islands volcano, Santorini, Greece. *Geosphere* 2: 253-268.

Raaflaub LD and Collins MJ (2006) The effect of error in gridded digital elevation models on the estimation of topographic parameters. *Environmental Modelling and Software* 21: 710–732.

Remondino F, Del Pizzo S, Kersten TP, et al. (2012) Low cost and open source solutions for automated image orientation - A critical overview. In: Ioannides M, Fritsch D, Leissner J, et al. (eds), *Progress in Cultural Heritage Preservation*, Springer-Verlag Berlin Heidelberg, pp. 40–54.

Roca M and Davison M (2010) Two dimensional model analysis of flash-flood processes: application to the Boscastle event. *Journal of Flood Risk Management* 3: 63-71.

Schindler, G., Dellaert, F., & Kang, S. B. (2007). Inferring temporal order of images from 3D structure. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Sibson, R (1981) A brief description of natural neighbor interpolation. *Interpreting Multivariate Data*, pp. 21-36.

Smith MW, Carrivick JL and Quincey DJ (2016) Structure from motion photogrammetry in physical geography. *Progress in Physical Geography* 40: 247–275.

Snavely N, Seitz SM and Szeliski R (2006) Photo tourism: Exploring photo collections in 3D. *ACM Transactions on Graphics* 25: 835–846.

Snavely N, Seitz SM and Szeliski R (2008) Modeling the world from Internet photo collections. *International Journal of Computer Vision* 80: 189–210.

Snavely N, Simon I, Goesele M, et al. (2010) Scene reconstruction and visualization from community photo collections. *Proceedings of the IEEE* 98: 1370–1390.

Stevens N, Manville V and Heron D (2003) The sensitivity of a volcanic flow model to digital elevation model accuracy: experiments with digitised map contours and interferometric SAR at Ruapehu and Taranaki volcanoes, New Zealand. *Journal of Volcanology and Geothermal Research* 119: 89–105.

Sword-Daniels V, Wardman J, Stewart C, et al. (2011) Infrastructure impacts, management and adaptations to eruptions at Volcán Tungurahua, Ecuador, 1999-2010. *GNS Science Report*: 1–76.

Tassi F, Vaselli O, Papazachos CB, et al. (2013) Geochemical and isotopic changes in the fumarolic and submerged gas discharges during the 2011-2012 unrest at Santorini caldera (Greece). *Bulletin of Volcanology* 75: –15.

Tuffen H, James MR, Castro JM, et al. (2013) Exceptional mobility of an advancing rhyolitic obsidian flow at Cordón Caulle volcano in Chile. *Nature Communications*, Nature Publishing Group 4: 2709.

UN-ISDR (2013) *Post-2015 Framework for Disaster Risk Reduction (HFA2)*. Geneva: United Nations International Strategy for Disaster Reduction.

USGS (2013) Did You Feel It? *earthquake.usgs.gov*. Available from: http://earthquake.usgs.gov/earthquakes/dyfi/ (accessed 19 May 2013).

USGS (2015) ShakeMaps., https://earthquake.usgs.gov/data/shakemap/

van Westen CJ, Castellanos E and Kuriahose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, Elsevier 102(3-4): 112–131.

Verhoeven G, Doneus M, Briese C, et al. (2012) Mapping by matching: A computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science* 39: 2060–2070.

Wackrow R and Chandler JH (2008) A convergent image configuration for DEM extraction that minimises the systematic effects caused by an inaccurate lens model. *Photogrammetric Record* 23: 6–18.

Wackrow R and Chandler JH (2011) Minimising systematic error surfaces in digital elevation models using oblique convergent imagery. *Photogrammetric Record*, 26: 16–31.

Westoby MJ, Brasington J, Glasser NF, et al. (2012) 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179: 300–314.

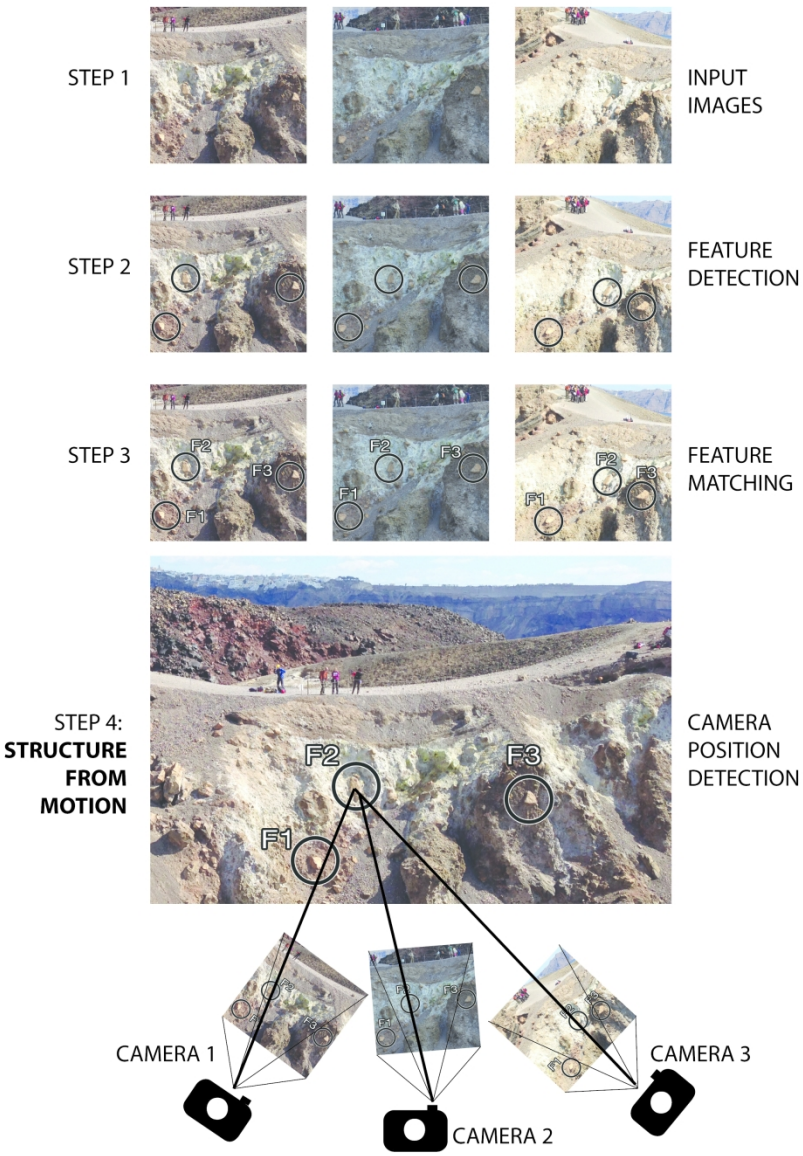
Wieczorek GF, Larsen MC, Eaton LS, et al. (2001) *Debris-flow and flooding hazards associated with the December 1999 storm in coastal Venezuela and strategies for mitigation*. USGS. Available from: http://pubs.usgs.gov/of/2001/ofr-01-0144/.

Wu C (2007) SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT). *cs.unc.edu*. Available from: http://cs.unc.edu/~ccwu/siftgpu/ (accessed 16 May 2013).

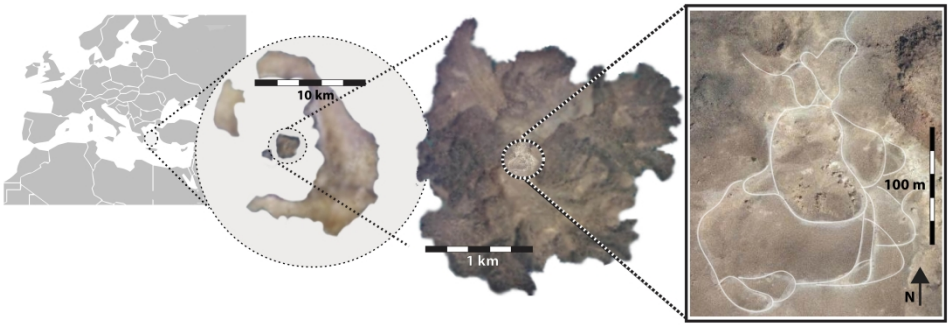
Wu C (2011) VisualSFM : A Visual Structure from Motion System. *homes.cs.washington.edu*. Available from: http://homes.cs.washington.edu/~ccwu/vsfm/ (accessed 16 May 2013).

Wu C, Agarwal S, Curless B, et al. (2011) Multicore bundle adjustment. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE: 3057–3064.

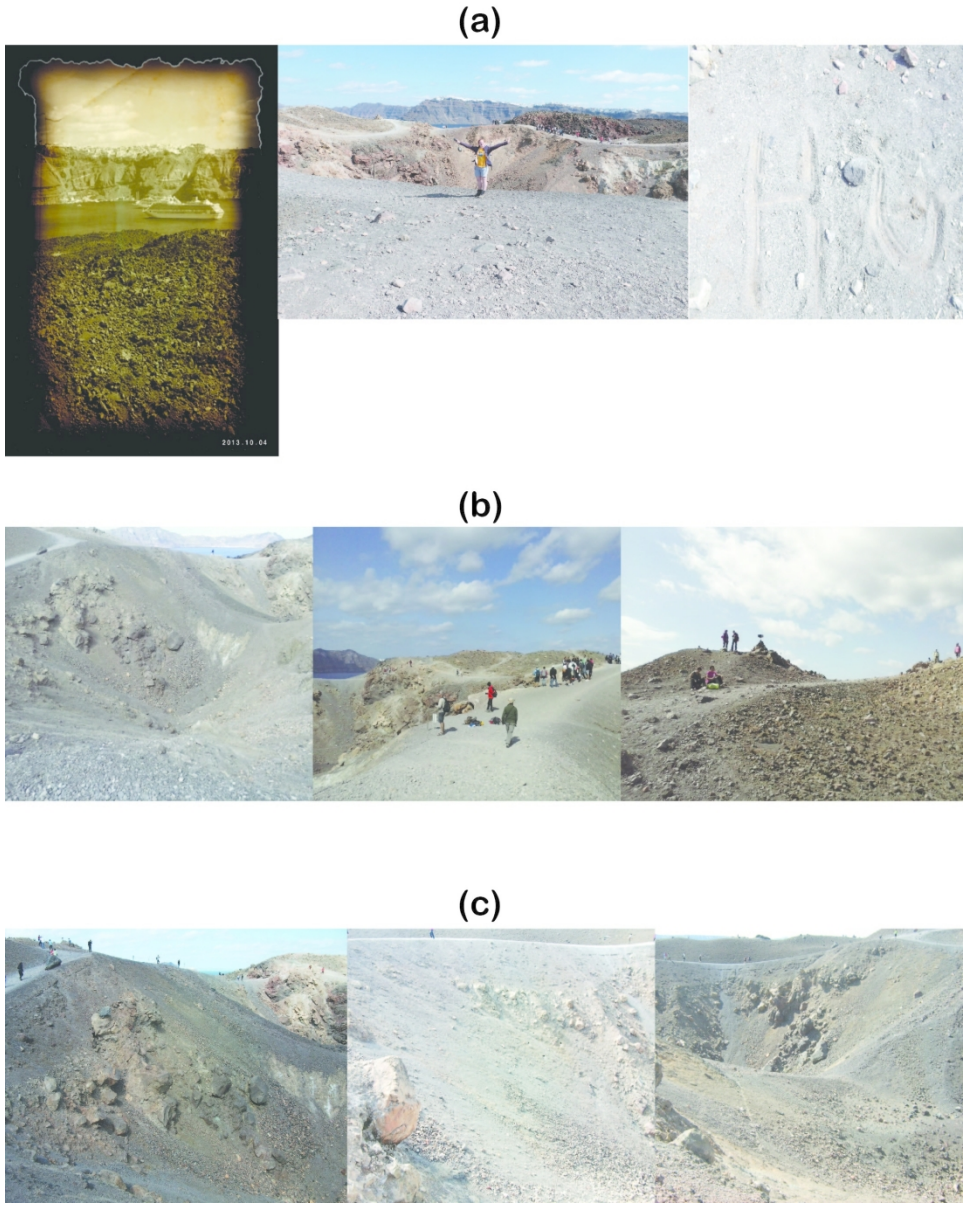
Zastrow M (2014) Crisis mappers turn to citizen scientists. *Nature* 515: 321.



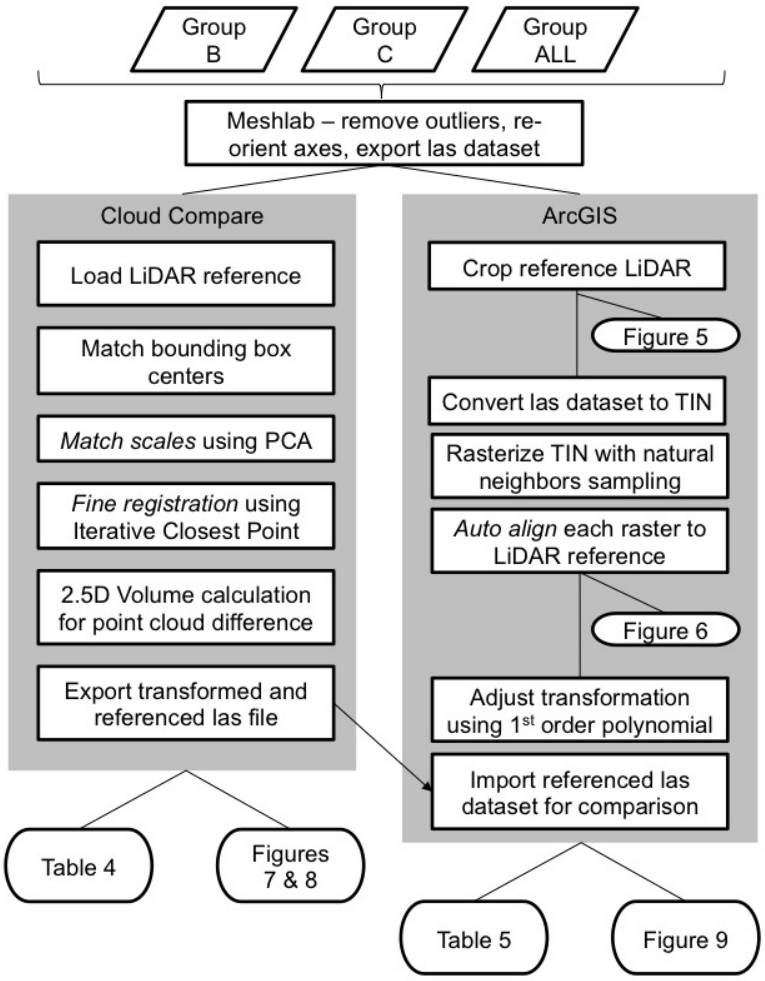
203x292mm (300 x 300 DPI)



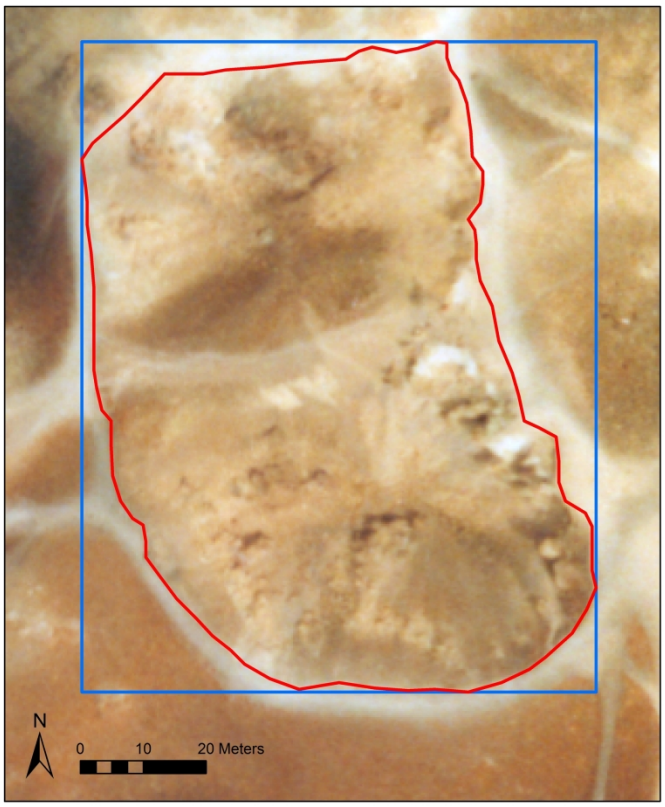
602x233mm (300 x 300 DPI)



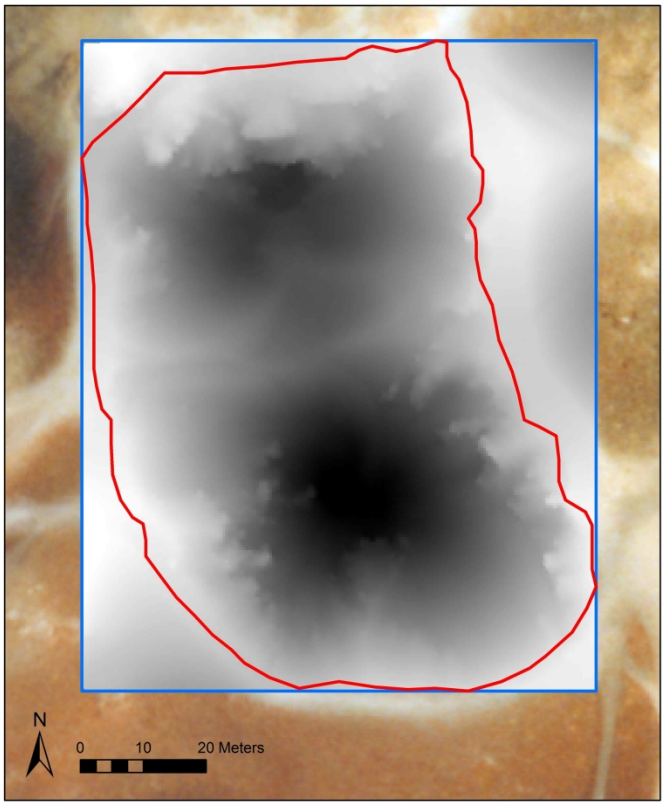
177x222mm (300 x 300 DPI)



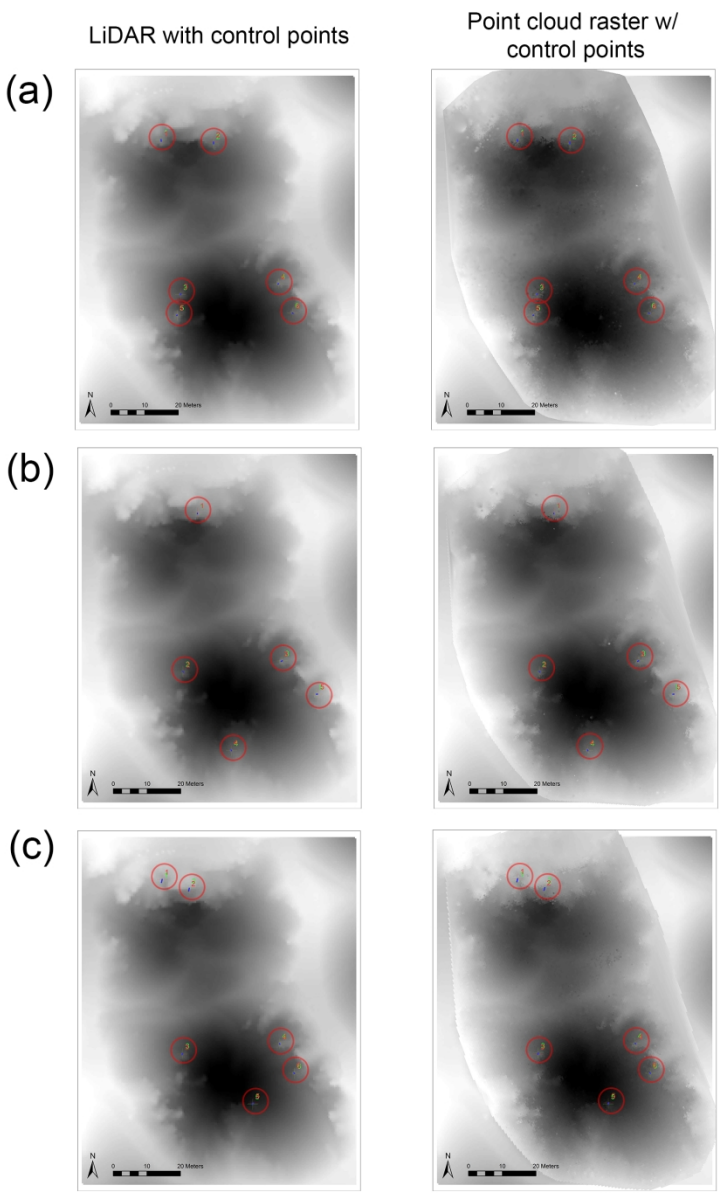
254x338mm (72 x 72 DPI)



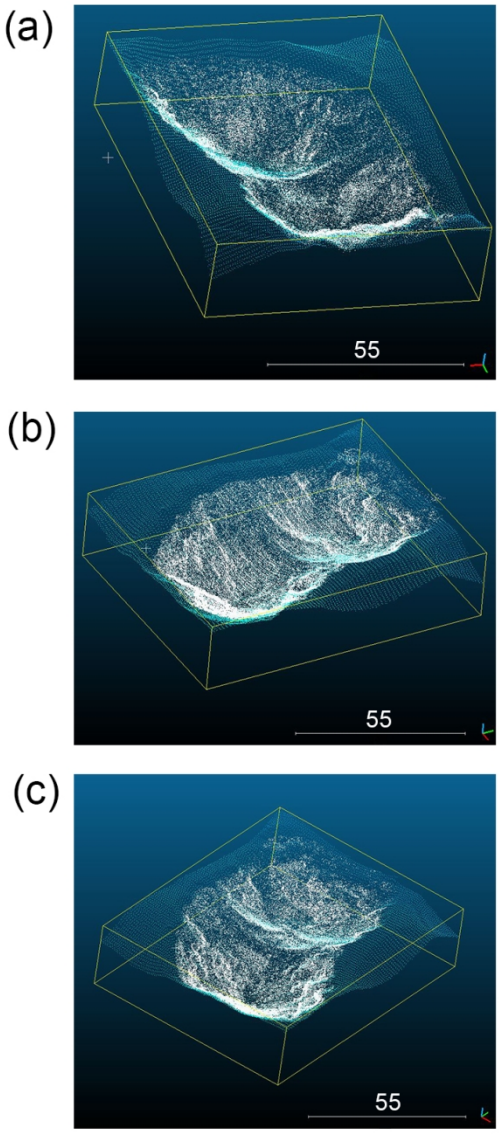
215x279mm (300 x 300 DPI)



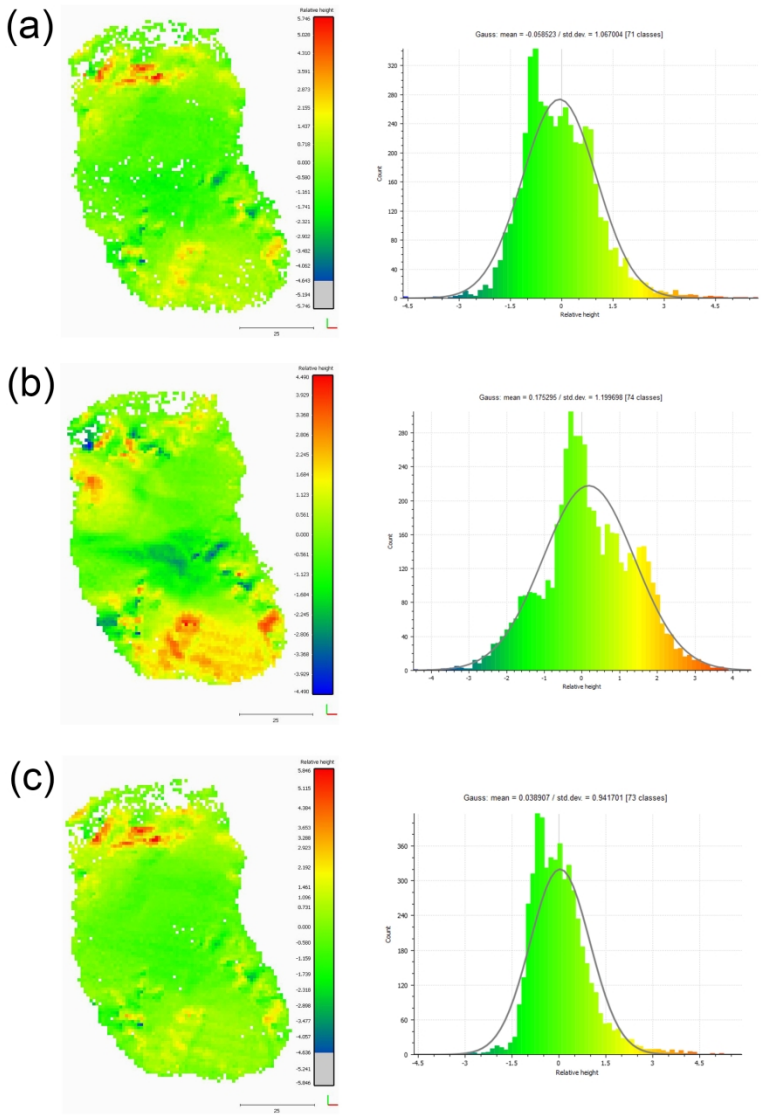
215x279mm (300 x 300 DPI)



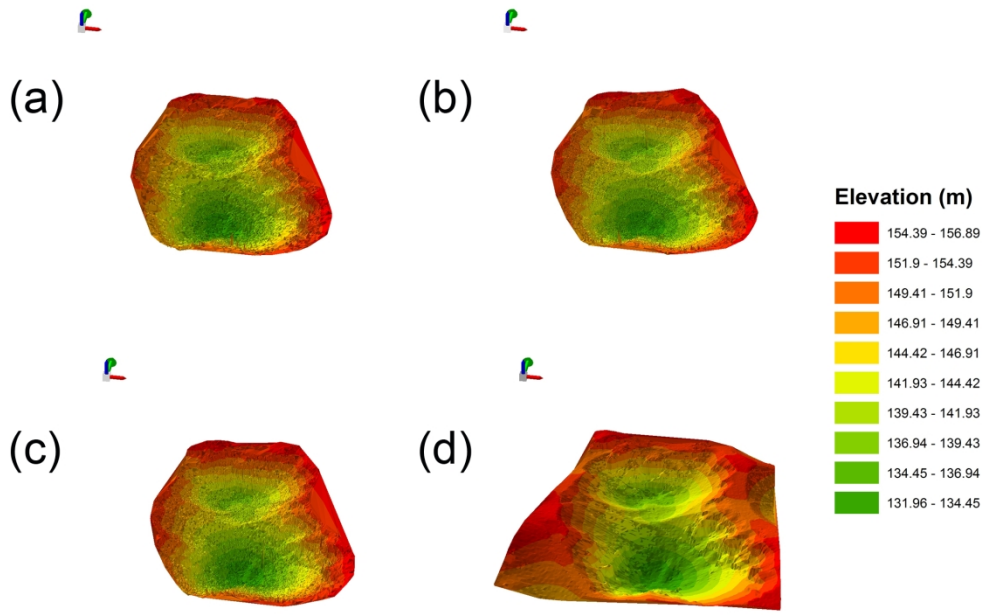
177x254mm (300 x 300 DPI)



65x143mm (300 x 300 DPI)



177x254mm (300 x 300 DPI)



185x130mm (300 x 300 DPI)

Figure 1. Schematic of the structure-from-motion process. Photos are input (Step 1) and scanned for identifiable features (Step 2). Identified features are matched across photos using the SIFT algorithm (Step 3). 'Structure-from-motion' is Step 4: features are simultaneously matched and used to reverse-compute the relative positioning of 'cameras' (in SfM, 'camera' refers to the location from which each individual photo was taken). Figure after (Snavely et al., 2010).

Figure 2. Study area orthophoto (Hellenic Cadastre, 2014). The Agios Georgios crater is located at the center of Nea Kameni Island in the Santorini island group in Greece. The crater is accessible via a network of footpaths (emphasized in figure), popular with tourists.

Figure 3. Examples of photos collected. (a) Group A photos were not ideal for SfM, often featuring large amounts of sky and background noise, humans in the foreground, or cropping and filtering rendering the photos useless. (b) Group B photos often included topography as part of the photo, if not necessarily the exact study area. No or few filters were applied by users in this group, although there are still often large amounts of sky or extraneous objects in the photos. (c) Group C photos minimized background noise and sky to an impressive degree, focused on the crater, and captured all of the area of interest.

Figure 4. Flowchart of data analysis methods.

Figure 5a. For analysis of data, a GIS shapefile was created from a perimeter trace of the study area as seen in the orthophoto (red line). For matching SfM data sets to LiDAR, the LiDAR bounding box was restricted to the study area (blue box).

Figure 5b. The extent trace (red line) and bounding box (blue box) of LiDAR data can be seen in the rasterized TIN, corresponding to the same extent trace and bounding box from the orthophoto.

Figure 6. Georeferencing raster-converted SfM datasets to raster-converted LiDAR in ArcGIS using auto-registration and first-order polynomial transformation generated 5-6 control points per dataset (inside red circles). (a) Group B (b) Group C (c) Group ALL. Readers will please refer to Table 5 for error values in georeferencing.

Figure 7. Alignment, registration, and iterative closest point analysis (ICP) of SfM datasets (white points) with LiDAR (blue points) in CloudCompare. Bounding boxes are included for perspective. (a) Group B (b) Group C (c) Group ALL. Readers will please refer to Table 4 for error values in alignment and registration.

Figure 8. Z-accuracy (elevation) difference maps of SfM datasets as compared against LiDAR on a metric scale in CloudCompare using the 'volume distribution' tool. (a) Group B (b) Group C (c) Group ALL. Green areas show areas of minimal difference, while red indicates the SfM cloud registering above the LiDAR points on the Z-axis, and blue indicates SfM registering below the LiDAR points. On the right hand side next to the maps are corresponding histograms demonstrating elevational accuracies in Gaussian distribution, including standard deviations and means.

Figure 9. Interpolated TIN images for LiDAR and SfM datasets. Without applying a smoothing function, these TINs are a coarse approximation of what a crowd-sourced DTM from SfM data may look like in a real world scenario. (a) Group B (b) Group C (c) Group ALL (d) LiDAR reference. There were no significant differences in value ranges to warrant separate legends for each experimental group.

Acknowledgements

The authors thank colleagues at the National Center for Disaster Preparedness at Columbia University's Earth Institute for support and constructive discussion, with special gratitude to Jeff Schlegelmilch. Additionally, we wish to thank colleagues at STREVA (Strengthening Resilience in Volcanic Areas, see streva.ac.uk) and COMET for support in this project, also the leaders, students, and demonstrators of the 2013 Santorini field trip. Additional thanks to Michelle M. Parks for supplying the LiDAR data, to John Stevenson for valuable feedback, and Greg Yetman for his technical guidance. Finally, to editor Karen Anderson and the anonymous reviewers for shaping this work to a publishable standard.

Funding

D.M. Pyle and T.A. Mather acknowledge funding from NERC COMET and NERC/ESRC STREVA grants, NE/J020052/1 and NE/J020001/1.

Exercise Information Sheet for Trip Leader

Objective

The objective of this exercise is to test the effect of "expertise level" on completeness and coherence of digital photo data collection for rapidly modeled DTMs (digital topography models) using structure-from-motion technology.

Materials Required

The participants of the field trip will require either a **camera**, or a **smart phone with a camera**, as well as the **USB cord used to transfer photos** from the device to a laptop. Alternatively, smart phone users can download the **Dropbox app** and ask for access to the folder.

SFM Methods and Technology Overview

SfM is a computer-vision based technology that takes tens to thousands of digital photographs and uses pixel-mapping algorithms to match points or items found in multiple photos. The result of matching the points in the digital photos is a 3D digital replica of objects featured in the photos. It is a useful technology for any circumstance where a digital model is needed: archaeologists use SfM to model rare artifacts so that they may be studied in detail from the comfort of an office (see Figure 1a); SfM is used in architecture to visualize buildings in 3D without the need for travel; and geoscience applications are currently exploring the usefulness of quick and cheap SfM as an alternative to costly or time consuming topographical mapping technologies like LiDAR or InSAR, which rely on lasers, radars, and satellites for data collection (see Figure 1b).

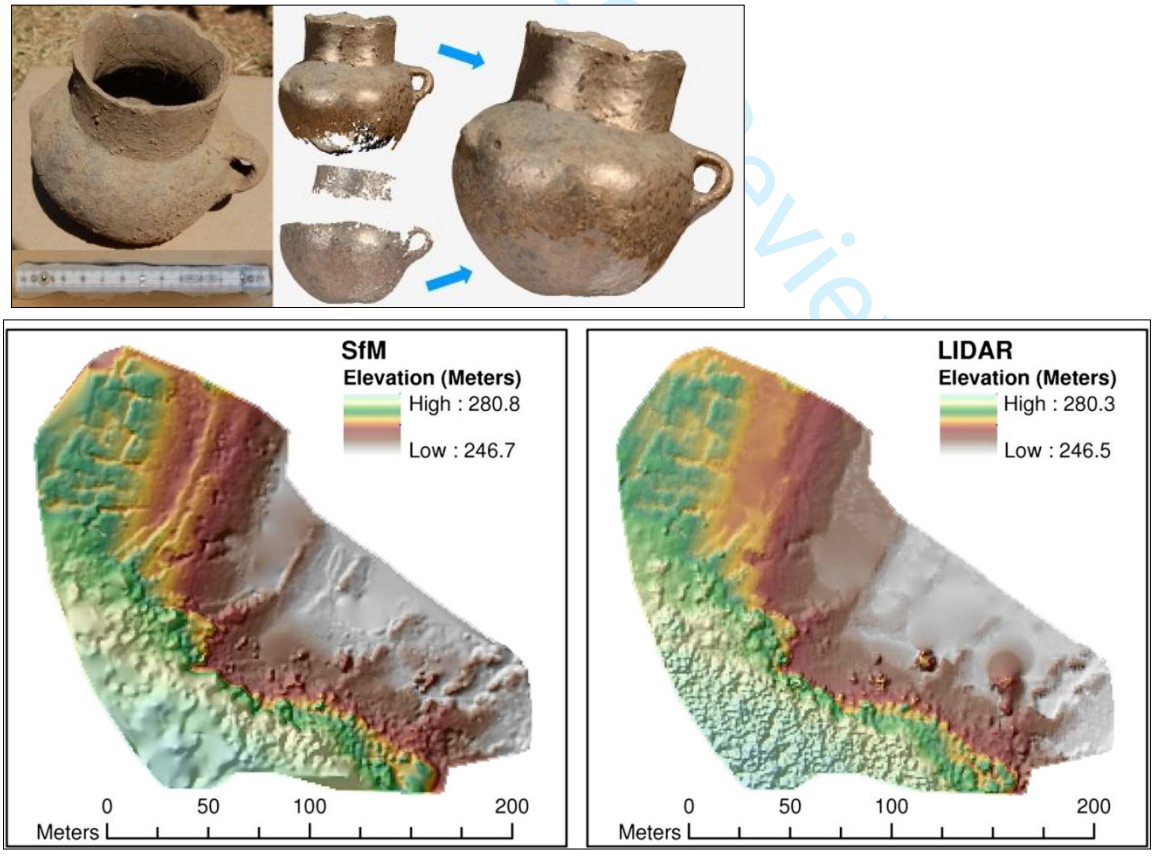


Figure 1. (a) Top: Pottery from Ethiopia reconstructed using SfM package Bundler/PMVS2 (Kersten 2012); **(b)** Bottom: Comparison of elevation measurements from SfM techniques versus LiDAR techniques (Fonstad 2012).

This exercise is part of a broader range of experiments testing the utility of structure-from-motion (SfM) technology in natural hazard and disaster reduction applications. The advantage provided by SfM is that it is capable of producing digital topography maps (DTMs) very rapidly (on a scale of hours) and with very little cost. The only input is tens to thousands of digital photographs of the area of interest. The best output is largely determined by how well matched the photos are- fewer photographs with more matching points often yield a better resulting DTM than many photos that match poorly. Similarly, high-resolution photos do not necessarily equate to better quality DTMs. There are a few approaches to photo collection that can increase the odds of making matches between points in photos:

1. *The rule of 3:* each point of interest must appear in at least 3 different photos, from 3 slightly different perspectives, in order to locate it in 3D space (triangulation).
2. *Overlap*: photos are more likely to match when they are very similar to one another. The best way of assuring this similarity is to make the field of view in the photograph only very slightly offset. As a general rule, a minimum of 60% overlap from one photo to the next yields good results.
3. *Lighting conditions:* although most SfM software packages contain measures to correct for differences in lighting (shadows, reflections, or other visually-altering conditions), it is still the safest bet to collect photos under conditions of strong but diffuse light (e.g. like the light on a slightly overcast day at noon). If not, care must be taken in areas of shadow, and reflections must be avoided.
4. *Unaltered photos:* more and more frequently, it's the norm to take digital photos through "filters" to achieve such effects as "antiqued" looks or more vibrant colors. SfM can correct for color alteration (see previous point about lighting conditions), but best results are achieved when colors, brightness, and contrast have not been skewed. More importantly, SfM calculations rely upon the orientation of the original photo, and so cropped photos will only produce outliers in the search for matching points.

Exercise Overview

4th year undergraduates on a fieldtrip to Greece will be split into three groups at the Nea Kameni Island volcanic crater. Each of the three groups will be given a role meant to emulate real-life actors in natural hazard scenarios.

- *Group A* will represent lay-people (everyday observers, bystanders, tourists, etc.) and will have no knowledge of the data collection technique employed.
- *Group B* will represent actors with a slight scientific knowledge (citizen scientists, volcano observatory volunteers, etc.), and will be given a brief overview of the exercise objective.
- *Group C* will represent experts (scientists, trained local officials, etc.) and will be given a thorough explanation of the technology and methods of data collection.

Groups will be instructed not to discuss their exercise sheets or to share information, except for with members of the same group. All groups will be instructed to use a set amount of time (e.g. a lunch hour) to collect data (digital photos) pertinent to the scenario described on their exercise sheet.

At the end of data collection, all photos will be uploaded into a Dropbox folder specific to that Group (A/B/C); each participant will make their own named folder in the Group folder and upload their photos to that location. As an additional measure of assurance, photos will be transferred to a USB that contains folders for each group's photos and returned to Jackie in Oxford upon the conclusion of the trip.

Group A (8 people)– Exercise Description

Please read the following scenarios, choose one role, and act according to your role using only information presented to you in the exercise. Omit any outside knowledge or experience. You will need a camera or a smart phone with a camera; later you will need a cable or other method of transferring your photos.

Once your time for the exercise is through, it is very important that you upload your photos to your* own named folder *within the* Group A Dropbox folder *and the USB for this experiment.

(1) You are a tourist visiting Greece for the first time and want to take photos that you can share with your friends and family (who unfortunately couldn't be with you), and to serve as memoirs of your holiday. *(Collect a minimum of 50 photos per person)*

(2) You are a local from Santorini who runs a small tourism operation. You have come out to Nea Kameni island to collect promotional photos for your website. *(Collect a minimum of 50 photos per person)*

(3) You are an amateur photographer and have come to visit Nea Kameni to take photos for artistic inspiration. *(Collect a minimum of 50 photos per person)*

(4) You are a travel blogger who has been offered a free trip to Santorini in exchange for publicizing the location and activities on the social media platforms that you use (e.g. Facebook, Instagram, etc.) *(Collect a minimum of 50 photos per person)*

Please note here:
(1) Your name,
(2) Role selected and how this impacted your approach to the exercise,
(3) The make and model of the device you are using,
(4) The quantity of photos you take,
(5) The numbers of the individual photos (if your device numbers them),
(6) Any camera parameters used (such as resolution, etc.).
Thank you for participating in this experiment.

Group B (6 people)- Exercise Description

Please read the following scenarios, choose one role, and act according to your role using only information presented to you on this exercise sheet. Omit any outside knowledge or experience. You will need a camera or a smart phone with a camera; later you will need a cable or other method of transferring your photos.

Once your time for the exercise is through, it is very important that you upload your photos to your* own named folder *within the* Group B Dropbox folder *and the USB for this experiment.

(1) You are a community member from one of the villages perched on the rim of the Santorini caldera. You, like most locals, know that the Kameni islands are part of a volcanic system but you do not understand more than a vague sense of danger from a distant future eruption. You have heard that there is an app that can help predict disasters using digital photos and you have come to the Nea Kameni crater to see if your photos can help. The website of the app gave you the instructions listed below. *(Collect a minimum of 50 photos per person)*

(2) You are an intern at the Institute for the Study and Monitoring of the Santorini Volcano (ISMOSAV) with a basic geologic knowledge of volcanoes. You understand that the topography of the Nea Kameni crater can tell you important things about past and future eruptions. You have been instructed to collect photos for a research project on the topography of the Nea Kameni crater, which is supposed to help with volcanic hazard management, and your supervisor offered the advice below. *(Collect a minimum of 50 photos per person)*

> When taking your photos, use “The Rule of 3” (each point of interest in the photo must appear in a minimum of three photos, from three different perspectives that overlap by at least 60%). Set your camera resolution to 5M or 8M and turn off the image stabilizer setting.

Please note here:
(1) Your name,
(2) Role selected and how this impacted your approach to the exercise,
(3) The make and model of the device you are using,
(4) The quantity of photos you take,
(5) The numbers of the individual photos (if your device numbers them),
(6) Any camera parameters used (such as resolution, etc.).
Thank you for participating in this experiment.

Group C (4 people)- Exercise Description

Please read the following scenarios and act according to your role using only information presented to you on this exercise sheet. Omit any outside knowledge or experience. You will need a camera or a smart phone with a camera; later you will need a cable or other method of transferring your photos.

Once your time for the exercise is through, it is very important that you upload your photos to your* own named folder *within the* Group C Dropbox folder *and the USB for this experiment.

You are a local government official from the municipality of Thera (governing body of all the Santorini islands). You have been trained by scientists from the Institute for the Study and Monitoring of the Santorini Volcano (ISMOSAV) to use a software that allows you to monitor the topography of the volcano without their help, thereby allowing you to track many important changes in volcanic activity and make autonomous decisions about hazard zoning and safety. Digital photos are the input for this software, and an excerpt from your training manual is presented below. *(Collect a minimum of 100 photos per person)*

SfM Methods and Technology Overview

Structure-from-motion (SfM) is a computer-vision based technology that takes tens to thousands of digital photographs and uses pixel-mapping algorithms to match points or items found in multiple photos. The result of matching the points in the digital photos is a 3D digital replica of objects featured in the photos. It is a useful technology for any circumstance where a digital model is needed: archaeologists use SfM to model rare artifacts so that they may be studied in detail from the comfort of an office (see Figure 1a); SfM is used in architecture to visualize buildings in 3D without the need for travel; and geoscience applications are currently exploring the usefulness of quick and cheap SfM as an alternative to costly or time consuming topographical mapping technologies like LiDAR or InSAR, which rely on lasers, radars, and satellites for data collection (see Figure 1b).

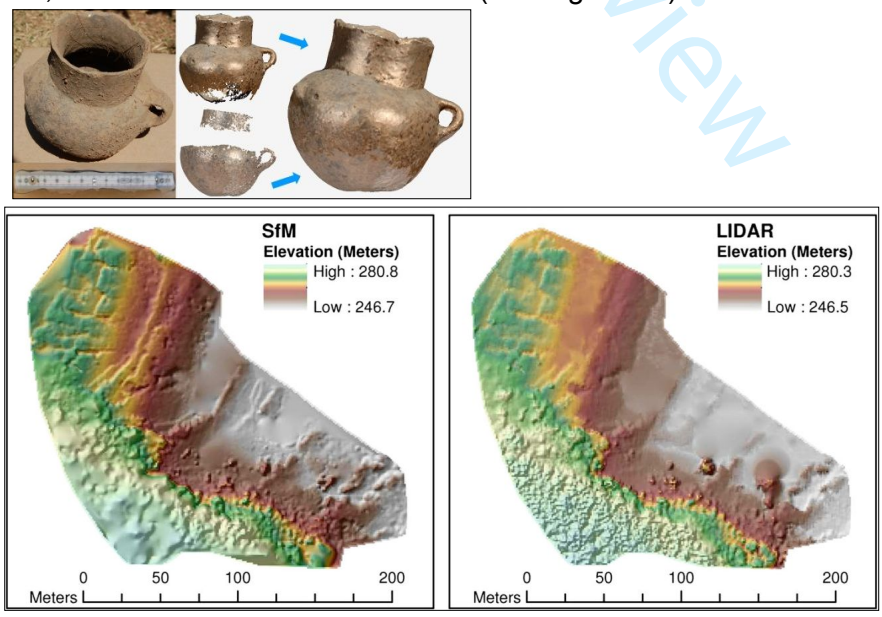
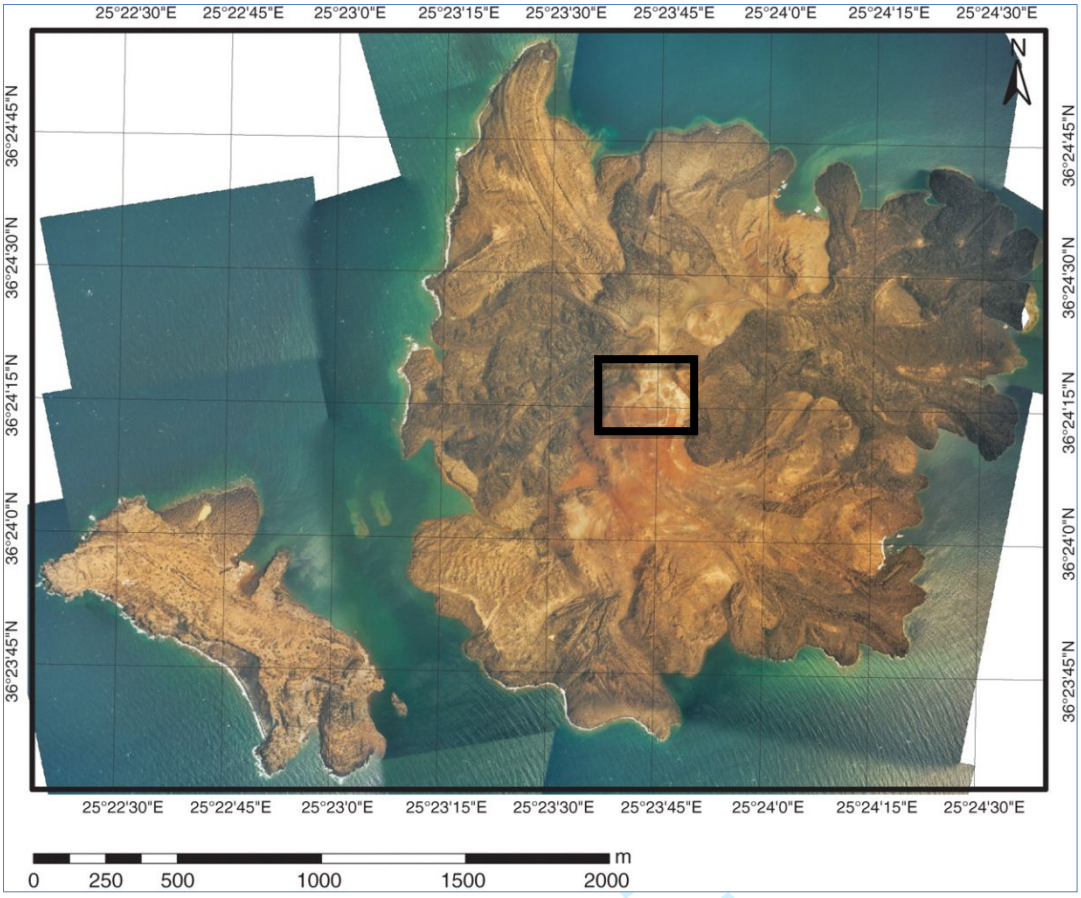


Figure 1. (a) Top: Pottery from Ethiopia reconstructed using SfM package Bundler/PMVS2 (Kersten 2012); **(b)** Bottom: Comparison of elevation measurements from SfM techniques versus LiDAR techniques (Fonstad 2012).

Structure-from-motion (SfM) technology has a particularly important utility in natural hazard and disaster reduction applications. The advantage provided by SfM is that it is capable of producing digital topography maps (DTMs) very rapidly (on a scale of hours) and with very little cost; this promotes an effective response to natural hazards and disasters. The only input is tens to thousands of digital photographs of the area of interest. The best output is largely determined by how well matched the photos are; fewer photographs with more matching points often yield a better resulting DTM than many photos that match poorly. Similarly, high-resolution photos (> 12 megapixels) do not necessarily equate to better quality DTMs; good results are obtained at 5M or 8M. There are a few approaches to photo collection that can increase the odds of making matches between points in photos:

1. *The rule of 3:* each point of interest must appear in at least 3 different photos, from 3 slightly different perspectives, in order to locate it in 3D space (triangulation).
2. *Overlap*: photos are more likely to match when they are very similar to one another. The best way of assuring this similarity is to make the field of view in the photograph only very slightly offset. As a general rule, a minimum of 60% overlap from one photo to the next yields good results. Areas of particular interest can be photographed from a closer distance, and will be incorporated into the model as a whole but to greater detail.
3. *Lighting conditions:* although most SfM software packages contain measures to correct for differences in lighting (shadows, reflections, or other visually-altering conditions), it is still the safest bet to collect photos under conditions of strong but diffuse light (e.g. like the light on a slightly overcast day at noon). If not, care must be taken in areas of shadow, and reflections must be avoided.
4. *Unaltered photos:* more and more frequently, it's the norm to take digital photos through "filters" to achieve such effects as "antiqued" looks or more vibrant colors. SfM can correct for color alteration (see previous point about lighting conditions), but best results are achieved when colors, brightness, and contrast have not been skewed. More importantly, SfM calculations rely upon the orientation of the original photo, and so cropped photos will only produce outliers in the search for matching points.

Additional points for smooth data collection: minimize background noise by filling your frame with your object and setting your focal lock to "infinite", decrease elevation error by taking photos directly facing your object whenever possible (avoid oblique angles, upward angles, or downward angles), and move in small, regular increments between photo collection (panoramas from one location are still from the same perspective and should be avoided).


25°22'30"E
25°22'45"E
25°23'0"E
25°23'15"E
25°23'30"E
25°23'45"E
25°24'0"E
25°24'15"E
25°24'30"E
36°24'45"N
36°24'30"N
36°24'15"N
36°24'0"N
36°23'45"N
N
0
250
500
1000
1500
2000
m

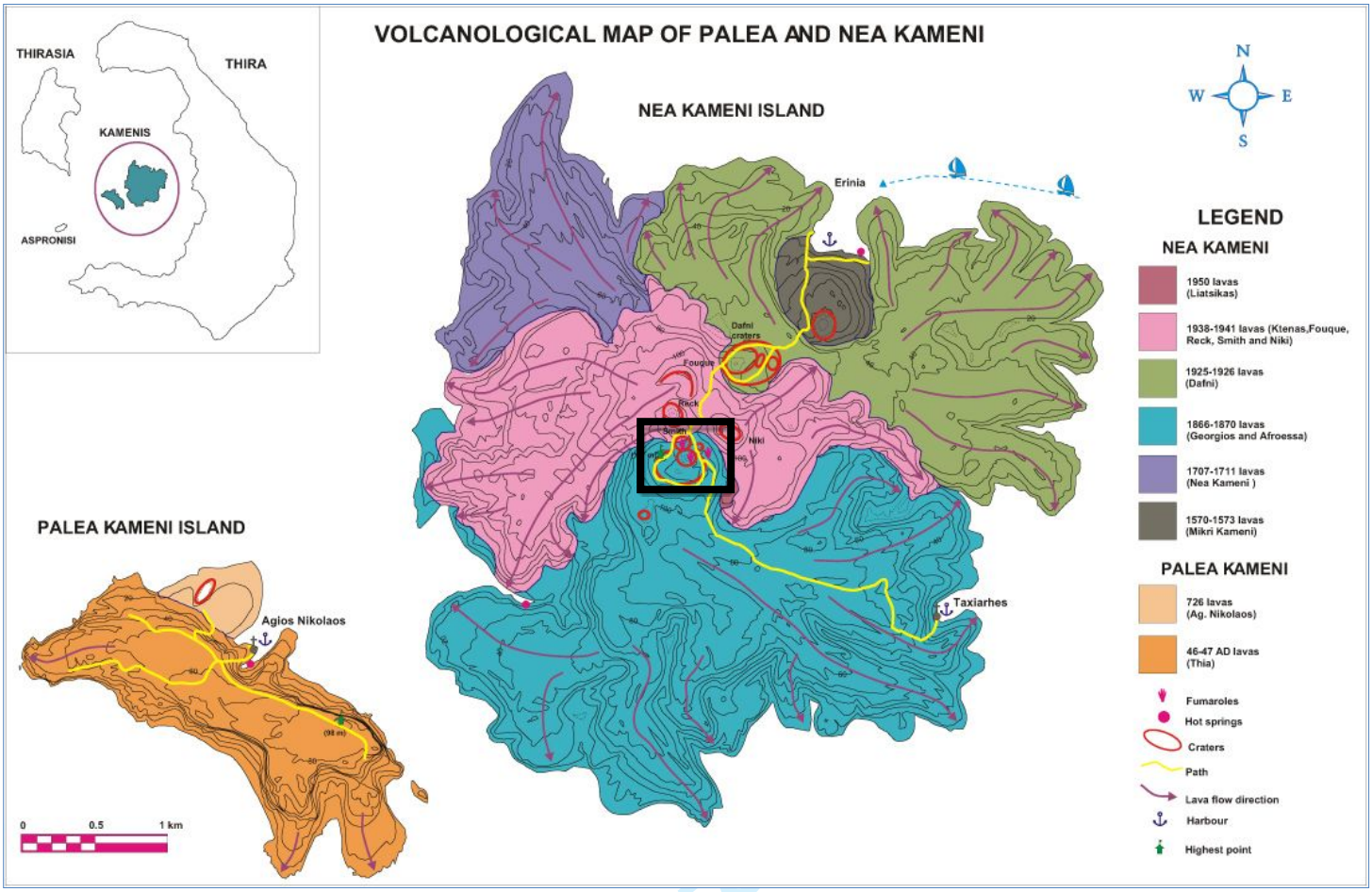


Figure 2. (a) Top: Orthorectified aerial photo mosaic of Nea Kameni island, with Agios Georgios crater in black rectangle. **(b)** Bottom: Geological contour map of Nea Kameni Island, with walking paths in yellow overlain on volcanic craters in red; area of interest in black rectangle.

Location-Specific SfM Methods

For understanding the topography of the Agios Georgios crater on Nea Kameni Island (see Figure 2) a circumnavigational survey of the crater interior should be completed using the footpath.

Photos should be taken at regular intervals (e.g. every 10° or 50 m) along the footpath, facing the interior. The best approach is to sight the opposite wall of the crater, filling the photo frame with the crater and taking multiple photos to mosaic the entire view, remembering that a photo mosaic is *not* the same thing as a panorama because a panorama (Figure 3a) is a wide-angle view with a curved baseline that needs to be orthorectified to relate distances and is therefore not useful for the objective here.

Figure 3. (a) Top: Panorama of Agios Georgios crater, notice the curved baseline that distorts distances (emphasized with yellow curved line). **(b)** Left: Straight baseline photo of Agios Georgios crater (emphasized with straight yellow line).

Ideally, the entire crater should be circumnavigated. If this is not possible, it is best to focus on one area (wall, crater floor, etc.) and image it thoroughly. For added reference between points, turn on the "geotagging" option on cameras and phones with the capability.

Please note here:

(1) Your name,
(2) How your role impacted your approach to the exercise,
(3) The make and model of the device you are using,
(4) The quantity of photos you take,
(5) The numbers of the individual photos (if your device numbers them),
(6) Any camera parameters used (such as resolution, etc.).

Thank you for participating in this experiment.