



Quantifying uncertainty in high-resolution remotely sensed topographic surveys for ephemeral gully channel monitoring

Robert R. Wells¹, Henrique G. Momm², and Carlos Castillo³

¹National Sedimentation Laboratory, Agricultural Research Service, United States Department of Agriculture, Oxford, Mississippi 38655, USA

²Department of Geosciences, Middle Tennessee State University, Murfreesboro, Tennessee 37132, USA

³Department of Rural Engineering, University of Córdoba, Córdoba, Spain

Correspondence to: Henrique G. Momm (henrique.momm@mtsu.edu)

Received: 13 January 2017 – Discussion started: 20 January 2017

Revised: 17 May 2017 – Accepted: 26 May 2017 – Published: 6 July 2017

Abstract. Spatio-temporal measurements of landform evolution provide the basis for process-based theory formulation and validation. Over time, field measurements of landforms have increased significantly worldwide, driven primarily by the availability of new surveying technologies. However, there is no standardized or coordinated effort within the scientific community to collect morphological data in a dependable and reproducible manner, specifically when performing long-term small-scale process investigation studies. Measurements of the same site using identical methods and equipment, but performed at different time periods, may lead to incorrect estimates of landform change as a result of three-dimensional registration errors. This work evaluated measurements of an ephemeral gully channel located on agricultural land using multiple independent survey techniques for locational accuracy and their applicability in generating information for model development and validation. Terrestrial and unmanned aerial vehicle photogrammetry platforms were compared to terrestrial lidar, defined herein as the reference dataset. Given the small scale of the measured landform, the alignment and ensemble equivalence between data sources was addressed through postprocessing. The utilization of ground control points was a prerequisite to three-dimensional registration between datasets and improved the confidence in the morphology information generated. None of the methods were without limitation; however, careful attention to project preplanning and data nature will ultimately guide the temporal efficacy and practicality of management decisions.

1 Introduction

Spatio-temporal measurements of landform evolution provide the basis for process-based theory formulation and validation. Field measurements of landforms have increased significantly worldwide, driven by the availability of new surveying technologies. Recent improvements include, but are not limited to, aerial and terrestrial light detection and ranging (lidar) systems (Kukko et al., 2012; Vinci et al., 2015; Eitel et al., 2016; Hawdon et al., 2016), integrated unmanned aerial vehicles (UAVs) utilizing both photogrammetric and lidar payloads (Bachrach et al., 2012; Bry et al.,

2015; Honkavaara et al., 2016), real-time kinematics (RTK; Rietdorf et al., 2006), terrestrial photogrammetric systems (James and Robson, 2014; Gómez-Gutiérrez et al., 2014; Di Stefano et al., 2016; Marzolf, 2016), and low-cost and/or freeware coupled structure-from-motion (SfM) and multi-view stereo (MVS) photogrammetric software (Castillo et al., 2012, 2015; Smith and Vericat, 2015; Piermattei et al., 2016). However, the buyer must be aware that these systems can be prone to misinterpretation (Wheaton et al., 2010), and even the “high-resolution” equipment can provide misleading information (e.g., Fig. 13b in Vinci et al., 2015). Research efforts should focus on a standardized and/or coordinated ef-

fort within the scientific community to collect morphological data in a dependable and reproducible manner, specifically when performing long-term process investigation studies (Castillo et al., 2016).

Ephemeral gullies are often defined as small channels on the order of a few centimeters in depth, predominantly in agricultural fields (Soil Science Society of America, 2008). The emergence, evolution, and persistence of these concentrated flow path erosion features is controlled by the combined effects of flow, slope, soil properties, topography, and vegetation characteristics (Zevenbergen, 1989; Castillo et al., 2016). The term ephemeral refers to the fact that agricultural producers often erase these channels during regular farming operations (Foster, 2005); flow within these channels is also often cyclical. The combination of a highly dynamic lifespan with the relatively small-scale channel features requires high-accuracy measurements with high temporal and spatial resolution.

Many studies have been conducted to assess the topographical accuracy of ephemeral or classical gully morphological measurements using a wide range of systems (e.g., Casali, et al., 2006; Gómez-Gutiérrez et al., 2014; Di Stefano et al., 2016). Among them, lidar data have been used as the reference for the evaluation of secondary remote sensing systems and physical contact systems. Traditional airborne lidar studies have primarily focused on quantifying locational error from datasets generated by airborne systems, in which locational variations are the result of coalesced errors generated by inaccuracies in the global positional system (GPS), aircraft inertial measurement unit (IMU), and overall timing of the system (Hodgson and Bresnahan, 2004). Lidar positional errors can also be the result of an interaction between the laser pulse and features with sharp relief change or occlusions that result in multiple returns from one laser pulse (Milenković et al., 2015). Evaluations of the accuracy of topographical information using airborne lidar are often compared with discrete sample locations and/or man-made targets with known coordinates (Hodgson and Bresnahan, 2004; Csanyi et al., 2005). Despite the large number of studies and methods developed to quantify positional errors in traditional airborne lidar surveys, this type of survey does not offer the temporal and spatial resolution necessary for the quantitative monitoring of small-scale geomorphological characteristics (i.e., ephemeral gullies) in terms of process description; however, recent developments in UAV lidar systems provide 10 mm of survey-grade accuracy, one million measurements per second, and a 360° field of view (FoV) in < 1.6 kg payloads. These UAV lidar systems can range from USD 100 000 to USD 400 000, depending on the level of accuracy and the data collection rate (see <http://www.rieglusa.com> for an example).

At a finer scale, investigation of ground-based and terrestrial lidar has demonstrated a high locational accuracy (~ 2 mm) and noted the importance of appropriate spatial sampling density for ephemeral and classical gully investi-

gation (Momm et al., 2013a). Topographic representations of gully channels require datasets with a specific minimum sampling density, which is dependent on site-specific topographical characteristics (Castillo et al., 2012; Momm et al., 2013a). Overlapping the same area with multiple scans increases the overall sampling density and assists in occlusion and shadow avoidance while normalizing spatial resolution.

Studies involving various surveying techniques of concentrated flow paths have revealed a wide range of quality, accuracy, cost, and field campaign effort (Momm et al., 2011, 2013b; Castillo et al., 2012; Wells et al., 2016). Among the surveying techniques considered, photogrammetry has been shown to provide simple but robust measurements of small-scale changes in geomorphologic characteristics within agricultural fields (Castillo et al., 2012; Gesch et al., 2015; Wells et al., 2016). Further, a wide variety of platforms and techniques have been used to capture images, including kites (Marzloff et al., 2003), backpacks (Wells et al., 2016) and UAVs (Ries and Marzloff, 2003; Bachrach et al., 2012; James and Robson, 2014; Cook, 2017). Erosion monitoring programs based on photogrammetry have several advantages compared to other surveying techniques. Photogrammetric field surveys do not interfere with farming operations, as they are nonobstructive; field campaigns are also extremely efficient and often do not require specialized technical skill sets to implement (James and Robson, 2012). However, photogrammetric results can vary as a function of the controlling parameters used during data collection and processing (Eltner et al., 2016).

A particular point of interest is the general query posed by Wheaton et al. (2010) concerning real geomorphic change. With these evolving technologies, our ability to collect topographical information is seemingly limitless. At what point can we agree that the results describe “real” change over noise? The alignment of temporal topographical elements is the most critical step when planning small-scale erosion studies (Smith and Vericat, 2015). Reliance on control points is the foundation of classical surveying. All surveys must close with a shot back to the initial occupation point. This is also the initiation of error propagation. A multitude of solutions exist for each set of photos and/or lidar points; however, the unique solution is bounded by the spatial and vertical positioning of the control points (Micheletti et al., 2015). Provided that alignment can be controlled, the next operation typically involves a culling process of some sort as the data shift into organized units.

The conversion of irregularly sampled point clouds into regular grids, referred to as digital elevation models (DEMs), is extremely common as most flow routing algorithms and soil erosion modeling technologies based on a geographic information system (GIS) are designed to work using these digital representations. As a result, a large number of studies have been conducted to evaluate DEM representation as affected by sampling intervals, interpolation algorithms (Aguilar et al., 2005; Ziadat, 2007; Bater and Coops, 2009;

Table 1. Datasets generated by three distinct surveying methods for the purpose of quantifying locational uncertainty in gully studies.

Dataset identification	Dataset description
Ground_2A	Channel left and right photo pair
Ground_2B	Upstream and downstream photo pair
Ground_4A	Channel left and right with corner left and right photo pair
Ground_4B	Upstream and downstream with corner left and right photo pair
Ground_4C	Upstream and downstream with channel left and right photo pair
Ground_6A	Upstream, downstream, channel left and right with corner left and right photo pair
Ground_8A	Upstream, downstream, channel left and right with both corner photo pairs
Quad_20m	Quadrotor flight at 20 m above ground surface
Quad_35m	Quadrotor flight at 35 m above ground surface
Fixed_61m	Fixed-wing flight at 61 m above ground surface
Fixed_122m	Fixed-wing flight at 122 m above ground surface
Lidar	Terrestrial lidar survey (reference)

* Dashed lines represent the delineation between survey modes.

James and Robson, 2012, 2014), and DEM spatial resolution (Zhang and Montgomery, 1994; Kienzle, 2004; Momm et al., 2013a).

The majority of previous studies have focused on accuracy evaluation of a specific photogrammetric survey method at a single time period. Varying sensors, platforms, and processing methods can yield different results (variations in sampling densities, gaps, and noise). Furthermore, measurements of the same site using identical methods and equipment, but performed at different time periods, can also lead to three-dimensional registration errors. Therefore, the scope of this work was to evaluate multiple survey techniques and provide a framework for temporal studies of ephemeral gully channels. Three surveying platforms with varying parameters were independently evaluated for locational accuracy and applicability in generating information for model development and validation. The objectives of this study are twofold: to quantify the overall accuracy of the different survey configurations and to develop practical guidelines for the design and implementation of future ephemeral gully monitoring studies.

2 Methods

2.1 Study site

The study site was located in the northwest corner of Webster County, Iowa, USA (Fig. 1). Farming is the dominant enterprise in Webster County. The crop rotation was a corn–soybean rotation. Total annual precipitation is about 873 mm, 70 % of which usually falls between April and September. The area of interest (AOI) within the field survey was a small reach (1.9×1.3 m; 2.47 m²) of a 150 m long ephemeral gully oriented south to northwest with eroding Clarion loam (fine-loamy, mixed, superactive, mesic Typic Hapludolls) at the upper (south) extent, Terril loam (fine-loamy, mixed, su-

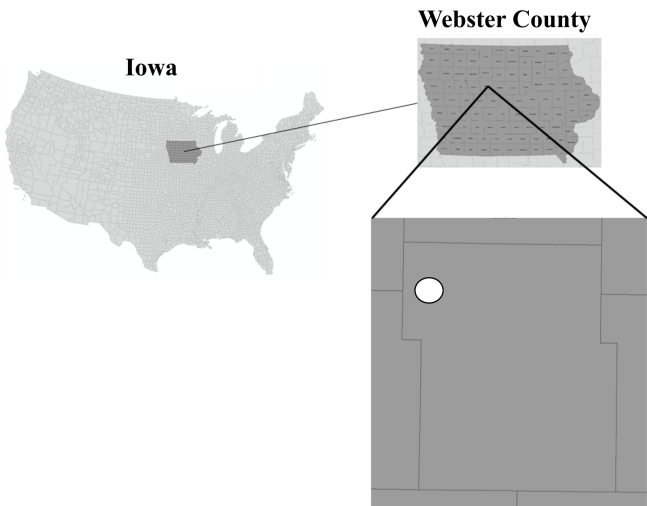


Figure 1. Study site location used in the evaluation of close-range photogrammetric surveys of ephemeral gully channels.

peractive, mesic Cumulic Hapludolls) on the intermediate slopes, and Webster clay loam (fine-loamy, mixed, superactive, mesic Typic Endoquolls) within the lower relief section of the field (Fig. 2). Within the AOI, the soil was Clarion loam (Fig. 2b).

2.2 Field survey

Field surveys were conducted using three independent modes: ground-based and terrestrial lidar, ground-based and terrestrial photogrammetry, and airborne photogrammetry. The surveys yielded 12 datasets (Table 1; dashed lines represent the delineation between survey modes). The terrestrial lidar was considered the reference dataset due to perceived superior accuracy. All surveys were run independently of

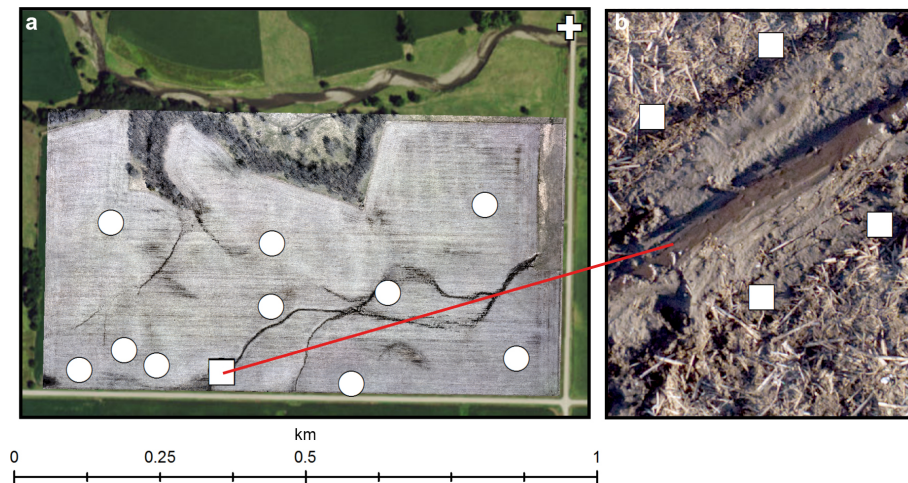


Figure 2. (a) Study site with field ground control points (GCPs; circles) and a state monument (cross) on the bridge in the upper right corner. (b) Selected AOI with channel GCPs (squares) for detailed surveys and comprehensive evaluation.

each other and completed on the same day. Each dataset was represented using the NAD83 UTM 15N coordinate system.

In this study, the terrestrial lidar point cloud was generated using Topcon ScanMaster software (<https://www.topconpositioning.com/software/mass-data-collection/scanmaster>), all terrestrial photogrammetric point clouds were generated using PhotoModeler Scanner software (www.photomodeler.com/products/scanner/default.html), and all airborne photogrammetric point clouds were generated using Pix4Dmapper Pro software (<https://pix4d.com/pix4dmapper-pro/>). It is acknowledged that the selection of input parameters influences the sampling intensity and local elevation variance; however, the quantification of the influence of input parameters on the output is beyond the scope of this study. Here, similar survey methods used the same input parameters to generate point clouds.

2.2.1 Differential global positioning system (DGPS) for ground control points (GCPs)

Site preparation began by locating a state monument point (Fig. 2a) and laying out 406 × 406 mm quad-triangle, white-on-black sheet GCPs and considering the long and short axes of the field as well as the high and low elevations within the field boundary, herein considered to be the field GCPs (10 total). One additional set of GCPs with RAD coded targets was arranged along the gully channel (four pins at the location; Wells et al., 2016), herein considered to be the channel GCPs (four total). All GCPs were surveyed using Topcon GR-3 DGPS survey equipment (Topcon Corporation, Tokyo, Japan; 10 mm of horizontal and 15 mm of vertical kinematic accuracy) to obtain relative position in reference to the state monument point. A static occupation (6 h; 3 mm of horizontal and 5 mm of vertical accuracy) was initiated with the base

station, then all GCPs (field, channel, and state monument) were surveyed with the rover (20 s collection interval). All survey data were corrected using an OPUS (National Geodetic Survey) solution for the base station location, processed with reference to the state monument point (6 mm of overall vertical accuracy). Both (field and channel) GCP positions were used to adjust point clouds from the lidar and photogrammetric surveys of the site.

2.2.2 Terrestrial lidar survey

The terrestrial lidar survey was conducted using a Topcon GLS 1500 (Topcon Corporation, Tokyo, Japan; 4 mm of single point and 2 mm of surface accuracy with a spot size < 6 mm). The system operates in a similar fashion to standard total stations. For each laser pulse, the system records *x*, *y*, and *z* coordinate values with respect to the position of the scanner sensor (local coordinate system), the intensity of the returned signal (reflectance), and spectral information from an integrated digital camera within the instrument. Local coordinates are transformed into global coordinates during postprocessing by entering the external geometry coordinates (i.e., absolute position determined from a postprocessed kinematic survey) of the GCPs. The AOI (demarcated by channel GCPs; 2.47 m²; Fig. 2b) within the field boundaries covering the gully channel was surveyed with one scan resulting in a total of 5 613 334 scan points.

Given the level of user control over the input parameters and the high locational accuracy of terrestrial lidar systems, this survey method was selected as the reference to which all other survey methods were compared. However, it is important to acknowledge that this survey method does have limitations.

In surveys with a high sampling intensity, it is common for the same location on the ground to be hit by multiple

laser pulses. This yields datasets with a high sampling intensity but a range of elevation values for the same location (i.e., fluff) given the vertical accuracy of the system. In this study, this elevation variability is estimated to be approximately ± 2 mm. The sensor operates in the near-infrared portion of the electromagnetic spectrum (1535 nm) and in this spectral range; electromagnetic energy is absorbed by water (Aronoff, 2005). In locations with water features, the sensor emits laser pulses, but no laser pulse is reflected back to the sensor; this prevents the range calculation for that particular pulse. During data collection for this survey, a shallow film of water was present in the gully channel (Fig. 3a). As a result, no points were recorded in the water-covered region (Fig. 3b; e.g., Gómez-Gutiérrez et al., 2014). Sampling gaps in lidar surveys can also be attributed to vertical features that limit the sensor line-of-sight, which is referred to as shadowing (Fig. 3c). The basic principle of lidar technology is to measure the time needed for an individual laser pulse to travel from the transmitter to the target and back to the receiver, allowing the range distance to be calculated (Wehr and Lohr, 1999). However, in certain situations, as the scanner moves along the scan arc, the laser footprint hits an area just past the edge of a surface where the next return appears to be from a distance greater than expected (i.e., occlusion). In this case, the gap is linearly filled by equally spaced points (Fig. 3c, highlight). A shadow of the obstruction appears in the dataset. These artificial points may be filtered by intensity, and multiple scan positions may be used to discriminate the features. Here, the AOI was slightly decreased in areal size to omit GCP occlusions from the dataset.

2.2.3 Terrestrial photogrammetric survey

Terrestrial photogrammetry was conducted using a Nikon D7000 16.2 MP camera (Nikon Inc., Melville, NY) with a calibrated 20 mm lens (Gesch et al., 2015; Wells et al., 2016). The camera was mounted to a backpack frame connected to an iPad mini (Apple, Cupertino, CA) through a WiFi CamRanger hub (Camranger LLC; <http://www.camranger.com>) (Wells et al., 2016). Multiple images were collected around the channel GCPs, including views from each corner and all sides. Still images captured by the camera were transformed into point clouds using PhotoModeler Scanner photogrammetric software. Initial data processing included aerial triangulation and bundle adjustment, camera position, and orientation. Following initial processing, the channel GCP positions (i.e., global external geometry) were included to optimize point cloud accuracy.

2.2.4 UAV-based photogrammetric survey

Two UAV platforms were used to collect airborne photography (<https://www.sensefly.com/home.html>): fixed wing (eBee) and quadrotor (albris). The fixed-wing platform had

a 12 MP nadir camera (i.e., belly mount; Canon S110 RGB) and was deployed (eMotion2 v2.4.10) to capture the entire field boundary (Fig. 2a) by throwing the craft in the air; the craft flies, captures images, and then lands itself. The deployment software parameters were altitude (117 m, 58 m), resolution (41 mm, 20 mm), latitude overlap (80 %, 50 %), longitude overlap (80 %, 50 %), image collection (356, 569), and image format (CR2 RAW) for the two respective flights. The quadrotor platform had a 38 MP camera mounted within a 180° vertical range head and was deployed (eMotionX v1.3.0) to capture both the extent of the gully within the field and specific points of interest (i.e., AOI; Gesch et al., 2015; Wells et al., 2016). The quadrotor was deployed through mission planning software. The craft takes off, flies and captures images, and then lands itself. The deployment software parameters were altitude (35 m, 20 m), resolution (7 mm, 5 mm), latitude overlap (75 %, 75 %), longitude overlap (80 %, 80 %), image collection (96, 146), and image format (DNG RAW) for the two respective flights. During the flights, winds from the southeast ranged from 7 to 10 m s⁻¹ and skies were clear.

Still images captured by the UAVs were transformed into point clouds using Pix4DMapper Pro photogrammetric software. Initial data processing included camera calibration, aerial triangulation and bundle adjustment, camera position, and orientation. Following initial processing, field GCP positions (i.e., global external geometry) were included to optimize point cloud accuracy.

Both fixed-wing and quadrotor UAV systems were deployed with a fixed path and common photograph overlap percentage. All missions and deployments were preplanned using flight planning and control software provided by the manufacturer. A mission block and a specific area or point of interest were selected, including preferred ground resolution, camera head angle (quadrotor only), and flight altitude. Flight lines for aerial coverage, circular paths with a horizontal plane around objects of interest (quadrotor only), image capture points, and waypoints were then generated prior to deployment. Key flight parameters were displayed in real time, along with the battery level and image acquisition progress, while the autopilot continuously analyzed onboard control data to optimize the flight.

2.3 Dataset alignment

To ensure the highest three-dimensional alignment among all point clouds and consistent spatial coverage by all methods, a three-step preprocessing approach was developed. First, the set of four channel GCPs (white square targets in Fig. 2b) were used to generate a rectangular polygon to subset the point clouds in all surveys and ensure that all surveys cover exactly the same ground position. Second, a smaller polygon subset was created by generating a polygon with a 50 mm reduction on all sides. This was performed to exclude areas close to the channel GCPs and ensure the elimination

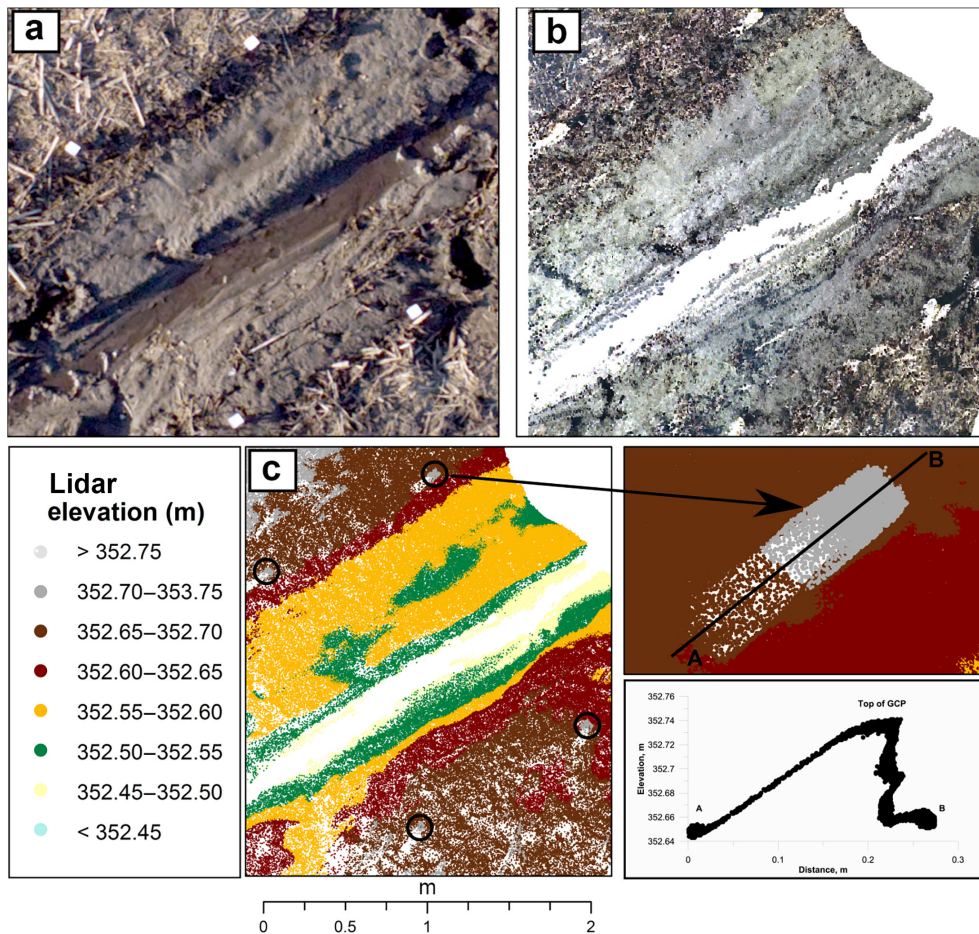


Figure 3. Limitations of the terrestrial lidar survey used herein as a reference dataset. (a) Photograph of AOI showing water in the channel, which limits laser pulse return to the sensor, causing sampling gaps in the point cloud (b). The presence of high relief features (GCPs) in the DEM (c) with sharp edges that cause the generation of multiple laser pulse returns due to the split footprint effect.

of shadowing and occlusion created by the slightly elevated channel GCPs. Third, the sampling void within the lidar dataset, created by the presence of a thin film of water within the channel (Fig. 3a), was manually digitized into another polygon and used to remove points from all photogrammetrically generated datasets to ensure uniformity among all datasets. Essentially, instead of using interpolated data within this void in the reference dataset, we simply placed a void in all datasets; therefore, we do not introduce bias into the calculations with regard to the water film void within the lidar data (e.g., Gómez-Gutiérrez et al., 2014).

Subsequently, since each surveying method was performed using the same set of field and channel GCPs, a manual inspection of measured points located coincident with channel GCPs (white square targets in Fig. 2b) was used to generate planes (10 total), one for each dataset with the exception of the Fixed_61m and Fixed_122m datasets, in which no points were located on top of the channel GCPs. The four GPS-surveyed coordinates of the center location of the channel GCPs were used to fit a reference plane to be matched

by all surveys (black squares in Fig. 4). Three-dimensional locational differences between the reference plane generated using the GPS survey (black squares in Fig. 4) and the planes of each surveyed dataset (lidar and photogrammetry) were calculated using the iterative closest point (ICP) algorithm (Besl and McKeay, 1994; James and Robson, 2012; Micheletti et al., 2015) implemented in Matlab (MathWorks Inc., Natick, Massachusetts). Since no scale issues were observed, no scaling factor was implemented in the ICP. The ICP algorithm minimizes the locational differences between two sets of three-dimensional point clouds and outputs a 3×3 rotation angle matrix, \mathbf{R} , and a 1×3 translation vector, \mathbf{T} (Eq. 1). These matrices were used to three-dimensionally transform, through rotation and translation, the measured point clouds to best match the reference plane:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{ref}} = \mathbf{T} + \mathbf{R}_{\omega\phi\kappa} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{meas}} \quad (1)$$

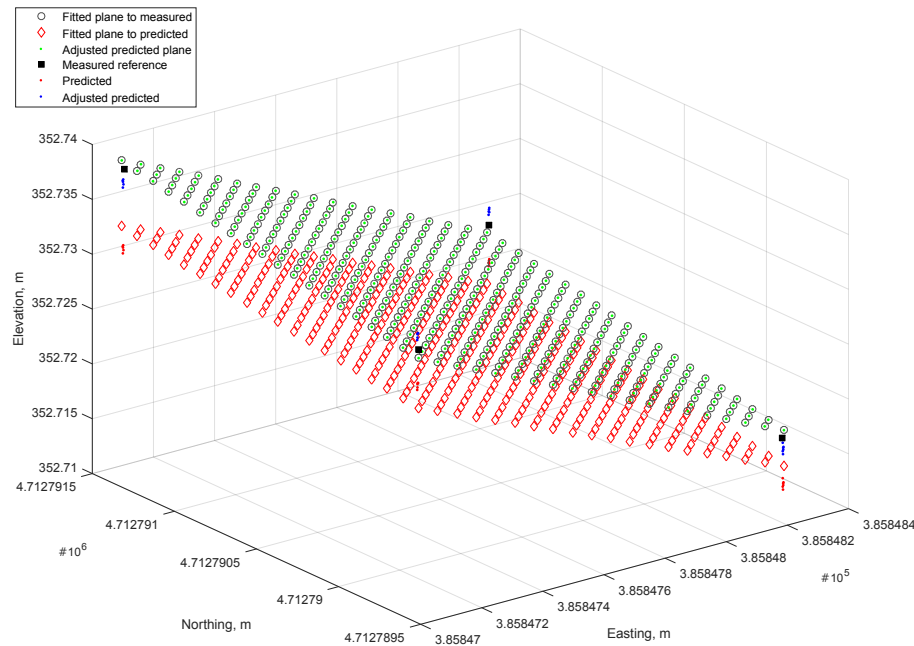


Figure 4. Determination of affine transformation matrices using iterative closest point (ICP) methodology.

2.4 Error metrics

One of the problems in looking at the data in the original point cloud format was the large difference between the total number of points within datasets (i.e., hundreds to millions), which tends to bias the results; therefore, in the sections that follow, the investigation of point cloud data is complemented by an analysis of gridded data. In the point cloud analysis, each point within the photogrammetry surveys was compared to that within the lidar survey. The analysis is carried out in two ways: point normal to the plane (each photogrammetry point was projected normal to a fitted plane of lidar points at the nearby position) and spot elevation to triangular irregular network (TIN; each photogrammetry point was projected up or down to intersect the TIN surface of the lidar points). In the gridded data analysis, a volume difference and cross-sectional assessment are performed. This type of data structure (i.e., raster grid) is a common format used to estimate soil loss volumes and generate cross sections for modeling exercises (Dabney et al., 2014). The gridded data introduce a common means of discussing differences between the survey methods.

2.4.1 Sampling intensity, local variance, and spatial pattern

Sampling intensity is defined as the number of points per unit of area. Investigation of the sampling intensity spatial variation can reveal oversampled or undersampled locations. Undersampled locations may be potential sources of error in quantifying geomorphologic change (i.e., cross-sectional

areas or volumes), especially in surfaces with high relief. Sampling intensity was evaluated using the quadrant method (Dodd, 2011), in which a virtual regular grid of 1 cm was imposed on each dataset and the number of lidar and photogrammetry points falling within each grid was counted and recorded. Similarly, the local elevation variance was evaluated by calculating the elevation range (difference between the maximum and minimum) within each grid. The local elevation variance is a function of the terrain characteristics, sampling intensity, survey method, and postprocessing parameters.

Two metrics were used to quantify the spatial pattern distribution: distances between events and between events and random points not in the pattern (void space). The G function, $G(r)$, defined as the cumulative frequency distribution of nearest-neighbor distances (Lloyd, 2010), provides the conditional probability that the distance between points (event–event) is less than the point distance threshold (r). The empirical distribution is obtained for each distance r by counting the number of points at distances less than or equal to r from each point within the AOI. The theoretical distribution is obtained by assuming a completely random pattern with density λ (estimated by the ratio of the total number of points divided by the area of the AOI), modeled as a Poisson process. Empirical values closer to the theoretical values indicate a random distribution, empirical values above the theoretical values indicate clustering, and empirical values below the theoretical values indicate a more regular distribution (Bivand et al., 2008). The F function, $F(r)$, defined as the cumulative frequency distribution of the distance to the nearest point in the AOI from random locations not represented

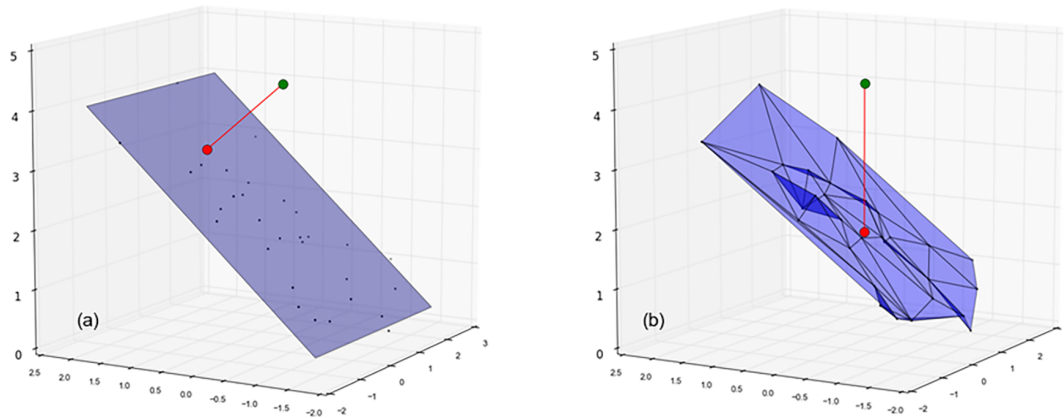


Figure 5. Schematic representation of the positional accuracy analysis. Black dots represent the reference dataset (lidar) and the green circle represents the point being evaluated from photogrammetry. (a) Red circles represent the normal projection of the green point onto the tangential plane fitted to the reference dataset and (b) the vertical projection of the green point into the three-dimensional reference triangular irregular network (TIN).

within the AOI (Lloyd, 2010), provides the probability of observing at least one point (event) closer than r to an arbitrary point within the AOI (“empty space” or “void” distances). Estimated and theoretical distributions are obtained in a way similar to the G function. The interpretation of the graphed observed versus the theoretical values indicates a regular pattern when the observed is above the theoretical values and clustering when it is below (Bivand et al., 2008). These point pattern analyses were performed using the spatstat package in the R software package (Baddeley and Turner, 2005; Bivand et al., 2008).

Furthermore, points with the same x , y , and z to the fifth decimal place were removed from the lidar dataset, as they indicate the collection of redundant information.

2.4.2 Vertical and horizontal displacement using point tangential projection to plane

The vertical and horizontal displacement between the lidar and photogrammetry measurements in each dataset was quantified using the normal projection of each photogrammetry point into a plane fitted to the nearest lidar points (Fig. 5a). For each photogrammetry point (green circle in Fig. 5a), the nearest (within a 25 mm sphere) lidar points were selected (black dots in Fig. 5a), a plane was fitted to the selected lidar point cloud points, the photogrammetry point was normally projected onto the plane, the coordinates of the intersection point (red circle in Fig. 5a) were recorded, and statistics were generated. This analysis was performed for all datasets using an in-house-developed Python script, where

$$\begin{aligned} D_x &= (Z_x^{\text{plane}} - Z_x^{\text{photo}}) \\ D_y &= (Z_y^{\text{plane}} - Z_y^{\text{photo}}) \\ D_z &= (Z_z^{\text{plane}} - Z_z^{\text{photo}}). \end{aligned} \quad (2)$$

2.4.3 Vertical displacement using spot elevation

The three-dimensional point cloud representing the reference dataset (lidar) was converted into a TIN (Fig. 5b). Each point in the photogrammetry dataset (green circle in Fig. 5b) was compared to the lidar TIN by fixing the photogrammetry x and y coordinates while varying the z coordinate up or down until the point intersected the TIN (red circle in Fig. 5b). The z coordinate at intersection was recorded. This analysis was performed using ArcGIS (ESRI, 2011):

$$D_z = (Z_z^{\text{TIN}} - Z_z^{\text{photo}}). \quad (3)$$

The vertical displacement at a location is calculated by subtracting the photogrammetry elevation from the lidar elevation of the TIN. Descriptive elevation statistics were generated based on all raster grid cells.

2.4.4 Gridded surface assessment

All point clouds were converted into 5×5 mm regular raster grids using linear interpolation. The two upstream channel GCPs served as the base of the rectangular grid; the roughness of individual datasets was highly dependent on the sampling intensity. Volume difference calculations were performed between the lidar raster grid and the photogrammetry raster grid. Two metrics were calculated, volume difference and absolute volume difference (Eqs. 4 and 5):

$$V_{\text{diff}} = \sum_{i=1}^n [(Z_i^{\text{lidar}} - Z_i^{\text{photo}}) \cdot ca], \quad (4)$$

$$V_{\text{acc}} = \sum_{i=1}^n [\|Z_i^{\text{lidar}} - Z_i^{\text{photo}}\| \cdot ca], \quad (5)$$

where Z^{lidar} is the reference elevation, Z^{photo} is photogrammetry elevation, ca is the raster grid cell area (0.000025 m^2), and n is the total number of raster grid cells.

Table 2. Metrics used in the ranking analysis of the photogrammetric measurements.

Metric	Analysis	Description
M1	Volume calculations from gridded data	Absolute volume difference between photogrammetric and lidar-generated raster grids.
M2	Volume calculations from gridded data	Iterative accumulation of individual raster grid cell elevation differences using absolute values.
M3	Cross-sectional comparisons	Averaged coefficient of correlation between photogrammetric and lidar cross-sectional elevation values derived from raster grid analysis.
M4	Cross-sectional comparisons	Averaged standard error between photogrammetric and lidar cross-sectional elevation values derived from raster grid analysis.
M5	Cross-sectional comparisons	Averaged area percent difference between photogrammetric and lidar cross-sectional elevation values derived from raster grid analysis.
M6	Spot elevation	Range of elevation difference between lidar three-dimensional irregular mesh and gridded photogrammetry.
M7	Spot elevation	Variance of elevation difference between lidar and photogrammetry elevation values.
M8	Spot elevation	Mean elevation difference between lidar and photogrammetry elevation values.
M9	Spot elevation	Coefficient of correlation between lidar and photogrammetry elevation values.
M10	Spot elevation	Standard error of linear regression between lidar and photogrammetry elevation values.
M11	Normal to plane	Range of elevation difference between fitted plane to nearest-neighbor lidar points and gridded photogrammetry.
M12	Normal to plane	Variance of elevation difference between lidar and photogrammetry elevation values.
M13	Normal to plane	Mean elevation difference between lidar and photogrammetry elevation values.
M14	Normal to plane	Coefficient of correlation between lidar and photogrammetry elevation values.
M15	Normal to plane	Standard error of linear regression between lidar and photogrammetry elevation values.

2.4.5 Gridded cross-sectional assessment

Gully modeling technologies often use cross sections as basic modeling units. With the objective of assessing the error introduced by each survey to a cross-sectional analysis, the raster grid surfaces were used to generate nine cross sections. For each cross section, various assessments were conducted, including minimum elevation, maximum elevation, mean elevation, variance, linear modeling (R^2 , p , SEM), and area calculations. The area above the curve was selected as one of the metrics to quantify cross-sectional accuracy, given as

$$A_c = \sum_{i=1}^n [(353.00 - Z_i) \cdot \Delta_{\text{dist}}], \quad (6)$$

where A_c is the area for cross section c in square meters, Z_i is the elevation at point i in the cross section, and Δ_{dist} is the distance between points in the cross section (0.005 m). An elevation constant (353.00) was used to adjust all cross-sectional elevations due to local elevation relation to mean sea level, thereby truncating the area values. The deviation of the area estimates from the lidar were calculated using

$$A_{\text{dev}} = \left(\frac{A_{\text{obs}} - A_{\text{est}}}{A_{\text{obs}}} \right) \cdot 100. \quad (7)$$

2.5 Dataset scoring

Since there is no true standard of judging the performance of the measurements provided herein, a system of scoring was developed to grade the photogrammetry data with regard to the lidar data (Table 2). Scores between 1 and 11 were assigned to each evaluation category using both point cloud and gridded data. For example, the difference in absolute volume was assigned decreasing scores ($1 \rightarrow 11$) for increasing volume difference, and correlation coefficients were assigned decreasing scores ($1 \rightarrow 11$) for decreasing correlation. Put simply, if a variable had a positive impact, it received a higher score and all variables were equally weighted. Each score is defined in Table 2.

3 Results

3.1 Sampling intensity and point pattern evaluations

The point clouds evaluated here had a large variability in sampling intensity (from 1 to >250 points cm^{-2} ; Fig. 6). The difference in point sampling influences micro-topography and apparent roughness and may lead to bias in volume estimation. If the surface is rough, the effect will be greater (Fig. 7). Point counts are very low for the fixed-wing flights

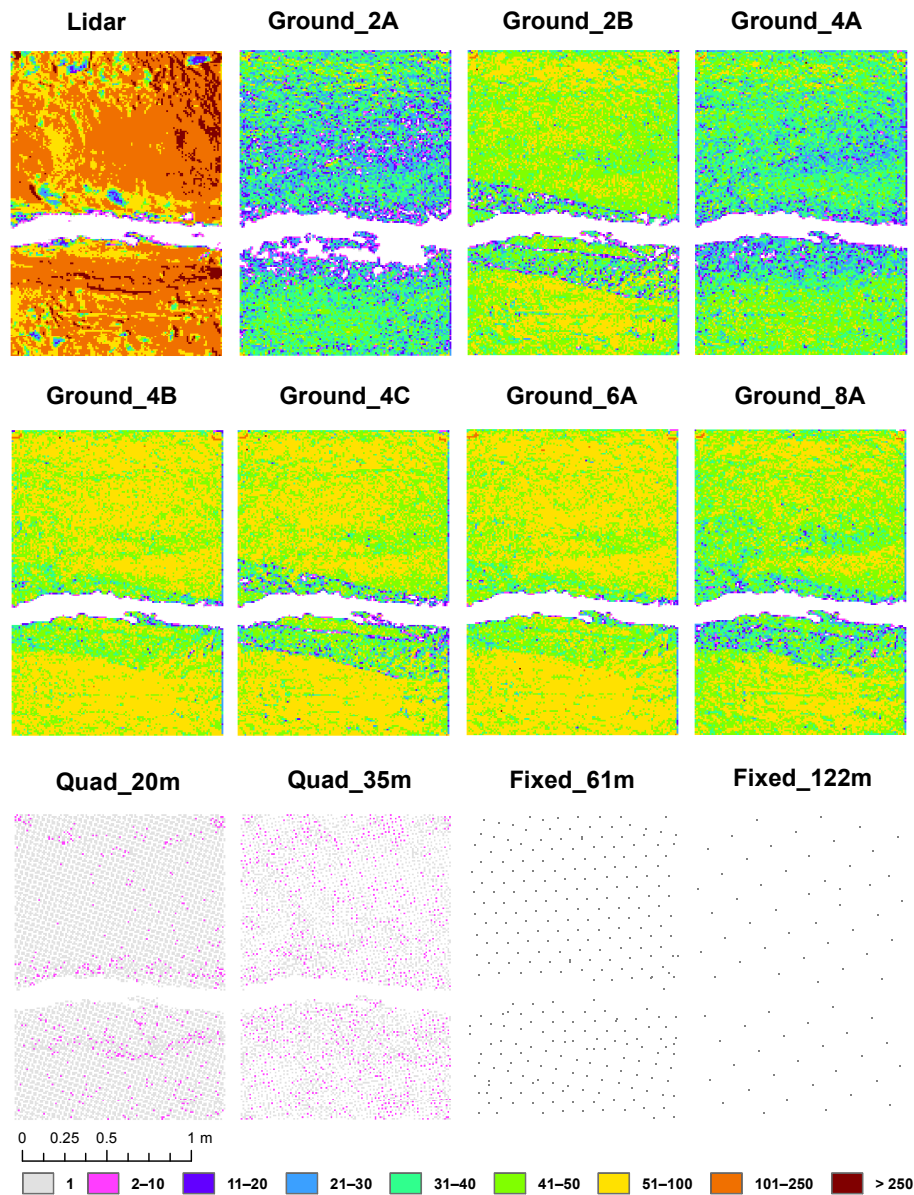


Figure 6. Sampling intensity using the quadrat method for each dataset considered. The individual colors represent point sampling count intervals within a 1×1 cm virtual grid. Points located in the channel were removed to match the area covered by the lidar dataset (herein considered as a reference).

in comparison to the other methods, and the sparse point count leads to interpolation (filling) during raster gridding (Fig. 6), while the elevation range is very similar for all methods with the exception of the fixed-wing datasets (Fig. 7).

Point pattern analysis was examined using the G and F functions (Fig. 8). Each analysis tests through an assumption of complete spatial randomness (homogeneous Poisson process), although interpretations for clustering and regularity are in opposition for each test (i.e., the regular point spacing outcome for the G function is below the gray confidence bounds, and for the F function it is above them). At first, the confidence bounds (gray envelope bounding the

theoretical values; red dashed line) show that the Fixed_122 has a sparse point count. Looking at the G function results, the data are clustered at distances of 0.02 m (Fixed_122), 0.002 m (Quad_20), and 0.001 m (Ground_8A) and then regularly distributed. This indicates that small distances occur less often than expected under the assumption of spatial randomness. The F function results indicate that the data are randomly distributed to 0.06 m (Fixed_122) and 0.025 m (Quad_20), and clustered for Ground_8A (i.e., at short distances, fewer points are encountered than for a random pattern); however, the scale of r should be acknowledged (< 3 mm). All 12 datasets yielded observed G func-

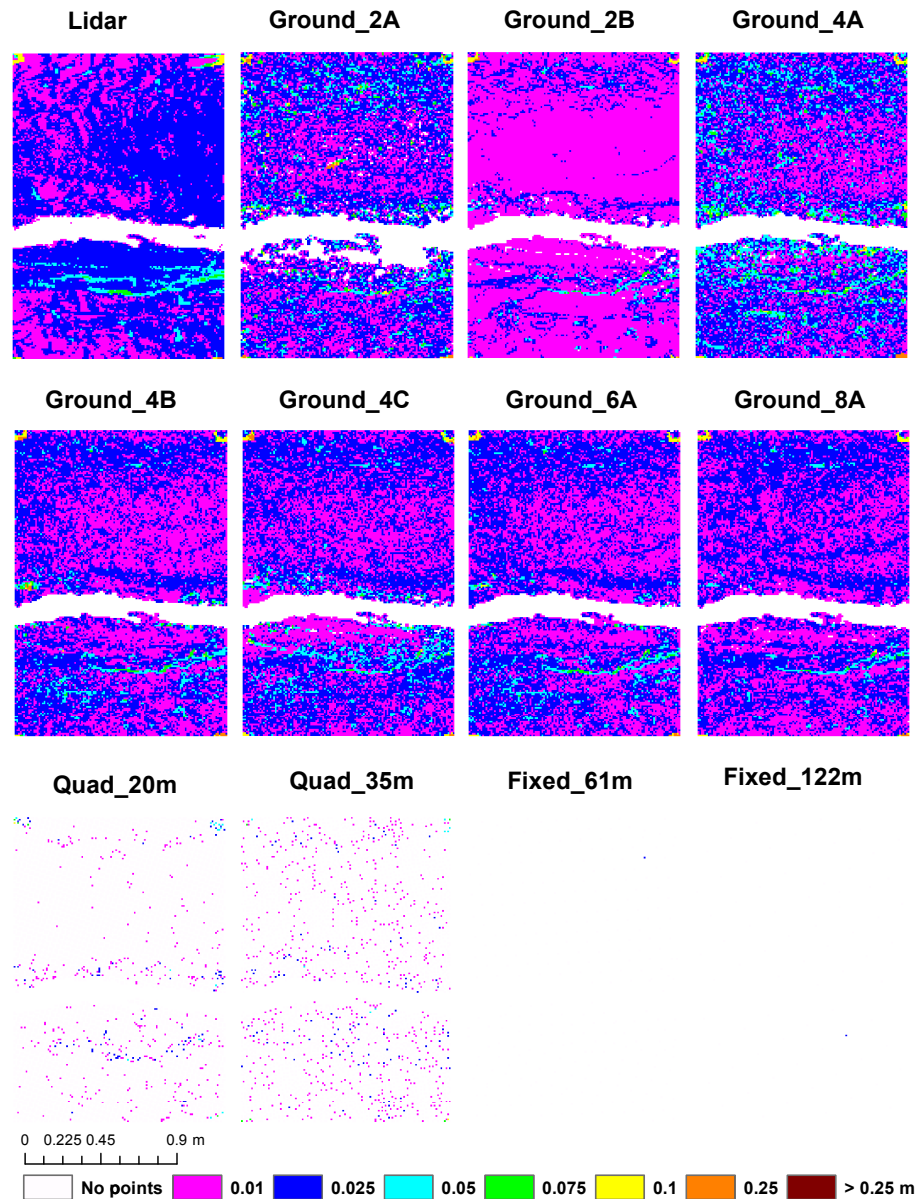


Figure 7. Elevation range (difference between the minimum and maximum elevation) represented as individual colors within a 1×1 cm virtual grid. Points located in the channel were removed to match the area covered by the lidar dataset (herein considered as a reference).

tion values below the theoretical values, indicating a regular sampling pattern. The terrestrial photogrammetric surveys showed slight clustering at small distances (< 3 mm). Therefore, based on these metrics, a regular sampling pattern was observed, indicating that all locations within the study area were sampled with a similar sampling pattern (no areas were oversampled or undersampled) for all 12 datasets in this study.

3.2 Vertical and horizontal displacement evaluations

Graphical representations of both spot elevation to TIN and normal to fitted plane for the fixed-wing flight at 122 m

of altitude (Fixed_122; Fig. 9) and the four-photo pair (Ground_8A; Fig. 10) are provided for comparative purposes. In the Fixed_122 spot and normal analysis (Fig. 9), the range was larger for the spot (± 0.06 m) than for the normal (± 0.02 m); the residuals suggest a nonlinear response, potentially attributed to over-smoothing of the surface and lack of preprocessing (slope of the blue line). Residuals for the Ground_8A (Fig. 10) have a constant variance and do not show an x or y axis bias. Clearly, outliers can be identified (Fig. 10; ~ 0.01 m) and it seems as though the spot analysis provides a larger variance (i.e., amplified residual signature) than the normal analysis, which may be attributed to the

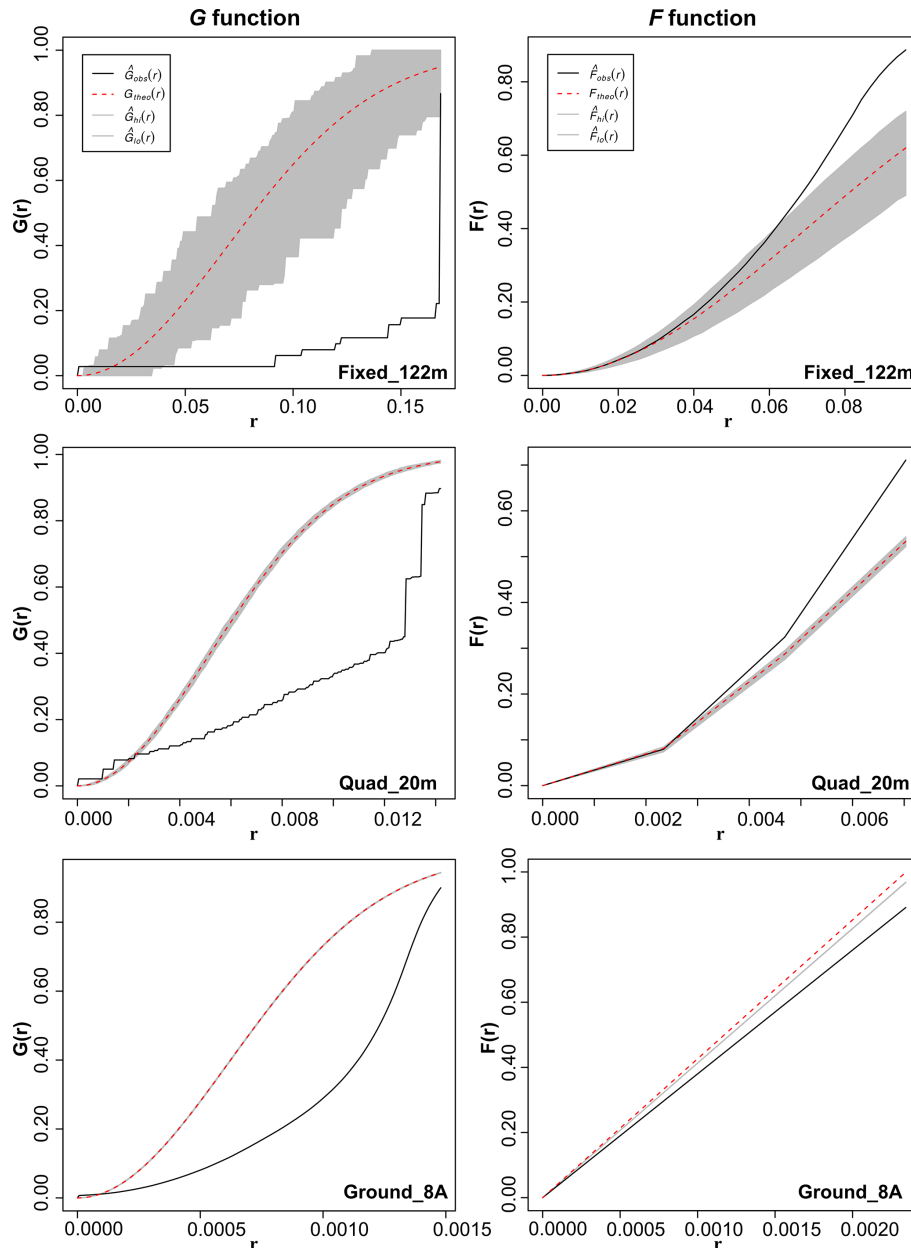


Figure 8. Results of the G function and F function analysis for the Fixed_122m, Quad_20m, and Ground_8A datasets.

2.5 cm sphere used to define the plane in the normal analysis (i.e., smoothing).

Very similar results were obtained with these two methods. All datasets had negligible mean displacement in the x and y directions (Table 3). The standard error of Z in the normal to plane analysis (Table 3) ranged from 0.3 mm (Ground_8A, Ground_6A, Ground_4B, Ground_4C, Ground_2B) to 2.9 mm (Fixed_122). The mean displacement for the normal to plane analysis in the z direction (Table 3) ranged from 0.2 mm (Quad_35) to 7.4 mm (Fixed_61). For the vertical spot to TIN analysis (Table 4), the standard error ranged from 0.4 mm (Ground_8A, Ground_6A,

Ground_4B, Ground_2B) to 2.1 mm (Fixed_122), and the mean displacement ranged from 0.1 mm (Ground_8A) to 35 mm (Fixed_122). With the spot to TIN analysis, the lidar point cloud was evaluated, resulting in a standard error of 0.3 mm and a mean displacement of 0.1 mm; both results are similar to Ground_6A and Ground_8A for standard error and Quad_35 and Ground_4A for mean displacement. The mean elevation difference between photogrammetry and lidar was approximately 5.3 mm; between the quadrotor and lidar it was 3 mm, and between the fixed wing and lidar it was 25 mm. Perhaps the 10-fold increase derived from the fixed-wing flights was simply because they were not prepro-

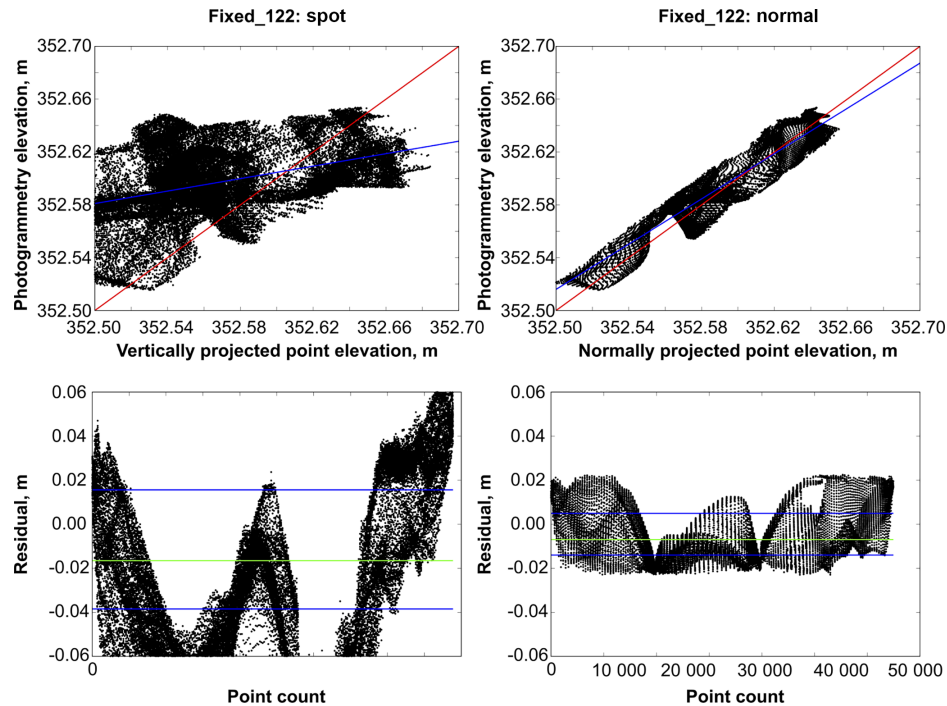


Figure 9. At 122 m of flight altitude, fixed-wing spot elevation comparison (left column) and normal to plane comparison (right column) of photogrammetry and lidar point cloud data with a fitted line through the elevations (blue; spot) and the 1 : 1 line (red; spot); green is the mean residual and the blue lines are the 25th and 75th percentiles.

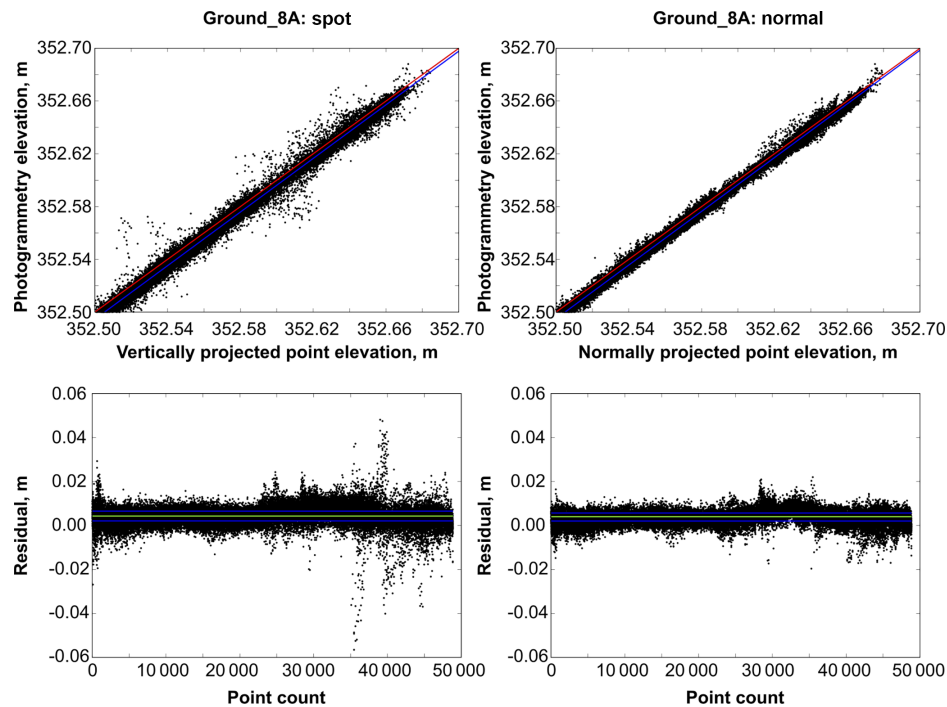


Figure 10. Four-photo pair (Ground_8A) spot elevation comparison (left column) and normal to plane comparison (right column) of photogrammetry and lidar point cloud data with a fitted line through the elevations (blue; spot) and the 1 : 1 line (red; spot); green is the mean residual and the blue lines are the 25th and 75th percentiles.

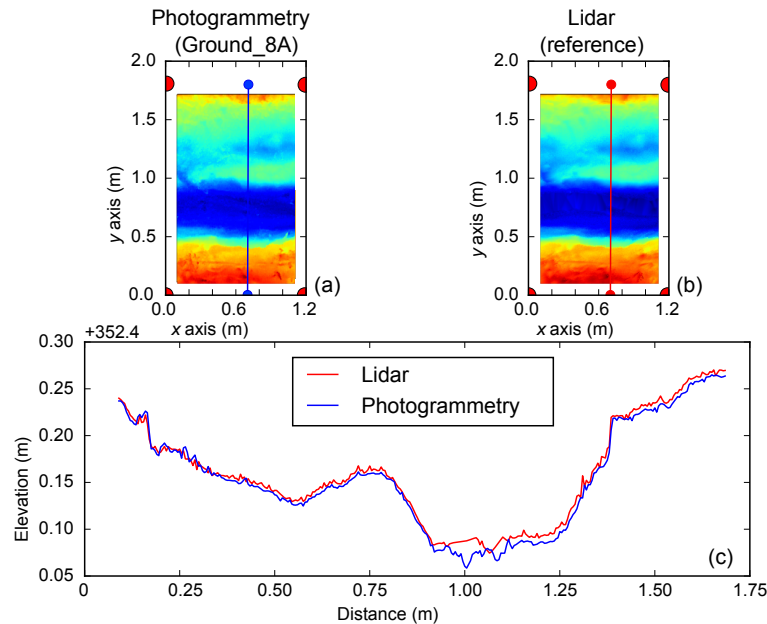


Figure 11. Illustration of three-dimensional point cloud interpolation into raster grids for volume and cross-sectional analysis of gully monitoring and geomorphologic quantification. Direct comparison of Ground_8A photogrammetry (a) and lidar (b) raster grid data with a highlight of one specific cross section (c). The cross section can be realized anywhere within the scene.

cessed (affine transformation was not applied); however, this result shows the importance of common reference points between surveys.

Most of what is reported here is due to point cloud alignment during the preprocessing step discussed earlier (i.e., GCP alignment for each, except fixed-wing flights). The mean elevation difference was approximately 5 mm for all datasets (Tables 3 and 4). Positive values indicate that the lidar data was, on average, higher than the photogrammetry data and negative values indicate that the lidar data was, on average, lower than the photogrammetry data. Similarly, when contrasting the photogrammetry elevation with the elevation of the normal point through linear regression (Table 3), the slope of the fitted line is very close to unity for all methods except Fixed_122, indicating spatially variable discrepancies (higher elevation differences at one region than the rest of the study site).

3.3 Gridded surface evaluations

The conversion of point clouds with irregularly spaced points and spatially varying sampling intensity point clouds into regular raster grids affected each dataset differently (Fig. 11). For example, the lidar dataset contained a high sampling intensity (>100 points cm^{-2}) with relatively large elevation variability in the points within a raster grid cell. Therefore, the interpolation procedure generated a significantly smoothed surface (e.g., Eitel et al., 2011). Conversely, the fixed-wing surveys had a low sampling intensity and the interpolation procedure linearly filled the gaps, potentially

generating a significantly smoothed surface that differs from the “natural” surface. The variance shows that roughness is similar for most of the surveys, with the exception of Fixed_122 for which the variance is extremely low (0.0005; Table 5). This result merely points out that the Fixed_122 is extremely smooth in comparison to the other surveys due to the low sampling density (Fig. 6) and enhanced interpolation between points for the high spatial resolution raster grid (0.005 m cell size).

One of the most important measurements for gully monitoring is the volume difference between surfaces (Table 6). Given the small scale of this type of erosional feature (on the order of a few centimeters), it is vital to have a good understanding of the expected error for each method. Among similar collection methods (terrestrial photogrammetry), the absolute volume difference (Table 6) ranged from 1 to 52 % in comparison to Ground_8A, although these differences were extremely small in reality (i.e., a range of 0.0062 to 0.0105 m^3). The performance ranking, in terms of absolute volume difference, was Quad_35, Ground_8A, Ground_4A, Ground_6A, Quad_20, Ground_4C, Ground_4B, Ground_2B, Ground_2A, Fixed_61, and Fixed_122 (10 to 164 % absolute volume difference for Ground_8A and Fixed_122, respectively, in comparison to Quad_35; Table 6). The variances in elevation difference between the lidar and photogrammetry data were all quite similar (Table 5), with the exception of the fixed-winged flights (effect of interpolation). In terms of the elevation range of the data (Table 5), the terrestrial photogrammetry and quadrotor flights were within 1.15 % of

Table 3. Statistics comparing photogrammetry and terrestrial lidar using the point normal projected into the fitted plane analysis. Residual values were calculated based on the coordinate difference of all raster grid cells.

Dataset	Minimum z axis	Maximum z axis difference (m)	Variance z axis difference (m)	Mean z axis difference (m)	Fitting linear model between photogrammetry and lidar in z axis				
					Slope	Intercept	r value	p value	Standard error (m)
Fixed_122	−0.0227	0.0231	0.0001	−0.0038	0.8564	50.6434	0.9332	0.0000	0.0029
Fixed_61	−0.0229	0.0229	0.0001	−0.0074	0.9284	25.2422	0.9793	0.0000	0.0010
Quad_35	−0.0205	0.0216	0.0000	0.0002	0.9935	2.2965	0.9959	0.0000	0.0004
Quad_20	−0.0148	0.0193	0.0000	0.0053	1.0315	−11.1177	0.9970	0.0000	0.0004
Ground_2A	−0.0226	0.0225	0.0000	0.0059	1.0239	−8.4214	0.9937	0.0000	0.0005
Ground_2B	−0.0192	0.0216	0.0000	0.0061	1.0258	−9.0910	0.9985	0.0000	0.0003
Ground_4A	−0.0219	0.0229	0.0000	0.0007	1.0027	−0.9458	0.9949	0.0000	0.0005
Ground_4B	−0.0212	0.0228	0.0000	0.0059	1.0203	−7.1574	0.9984	0.0000	0.0003
Ground_4C	−0.0218	0.0223	0.0000	0.0055	1.0178	−6.2732	0.9976	0.0000	0.0003
Ground_6A	−0.0204	0.0215	0.0000	0.0052	1.0219	−7.7293	0.9984	0.0000	0.0003
Ground_8A	−0.0197	0.0219	0.0000	0.0040	1.0177	−6.2473	0.9982	0.0000	0.0003

Dataset	Minimum y axis	Maximum y axis difference (m)	Variance y axis difference (m)	Mean y axis difference (m)	Fitting linear model between photogrammetry and lidar in y axis				
					Slope	Intercept	r value	p value	Standard error (m)
Fixed_122	−0.0169	0.0164	0.0000	0.0004	1.0001	−511.3216	1.0000	0.0000	0.0001
Fixed_61	−0.0180	0.0165	0.0000	0.0002	0.9974	12 467.67	1.0000	0.0000	0.0000
Quad_35	−0.0138	0.0158	0.0000	0.0004	0.9998	917.4185	1.0000	0.0000	0.0000
Quad_20	−0.0112	0.0153	0.0000	0.0004	1.0016	−7635.98	1.0000	0.0000	0.0000
Ground_2A	−0.0161	0.0164	0.0000	0.0002	1.0020	−9613.20	1.0000	0.0000	0.0000
Ground_2B	−0.0166	0.0162	0.0000	0.0003	1.0025	−11 557.52	1.0000	0.0000	0.0000
Ground_4A	−0.0165	0.0157	0.0000	0.0003	1.0004	−1749.80	1.0000	0.0000	0.0000
Ground_4B	−0.0165	0.0154	0.0000	0.0003	1.0023	−10 955.43	1.0000	0.0000	0.0000
Ground_4C	−0.0165	0.0134	0.0000	0.0002	1.0021	−9880.48	1.0000	0.0000	0.0000
Ground_6A	−0.0163	0.0155	0.0000	0.0003	1.0021	−9835.37	1.0000	0.0000	0.0000
Ground_8A	−0.0172	0.0143	0.0000	0.0003	1.0016	−7588.89	1.0000	0.0000	0.0000

Dataset	Minimum x axis	Maximum x axis difference (m)	Variance x axis difference (m)	Mean x axis difference (m)	Fitting linear model between photogrammetry and lidar in x axis				
					Slope	Intercept	r value	p value	Standard error (m)
Fixed_122	−0.0111	0.0098	0.0000	0.0000	1.0002	−88.1715	1.0000	0.0000	0.0001
Fixed_61	−0.0126	0.0094	0.0000	−0.0001	1.0005	−176.6591	1.0000	0.0000	0.0000
Quad_35	−0.0079	0.0042	0.0000	−0.0001	0.9999	21.3139	1.0000	0.0000	0.0000
Quad_20	−0.0055	0.0055	0.0000	−0.0001	0.9998	90.0612	1.0000	0.0000	0.0000
Ground_2A	−0.0113	0.0065	0.0000	−0.0001	0.9999	49.4894	1.0000	0.0000	0.0000
Ground_2B	−0.0114	0.0048	0.0000	0.0000	0.9999	28.8352	1.0000	0.0000	0.0000
Ground_4A	−0.0110	0.0061	0.0000	0.0000	1.0000	−7.5965	1.0000	0.0000	0.0000
Ground_4B	−0.0123	0.0047	0.0000	0.0000	0.9999	35.4773	1.0000	0.0000	0.0000
Ground_4C	−0.0111	0.0052	0.0000	0.0000	0.9999	41.1742	1.0000	0.0000	0.0000
Ground_6A	−0.0111	0.0037	0.0000	0.0000	0.9999	21.1864	1.0000	0.0000	0.0000
Ground_8A	−0.0118	0.0031	0.0000	0.0000	1.0000	13.6678	1.0000	0.0000	0.0000

Table 4. Statistics comparing photogrammetry and terrestrial lidar using the point vertical spot to TIN approach (z direction). Residual values were calculated based on the elevation difference of all raster grid cells.

Dataset	Minimum elevation difference (m)	Maximum elevation difference (m)	Variance elevation difference (m)	Mean elevation difference (m)	Fitting linear model between photogrammetry and lidar				
					Slope	Intercept	r value	p value	Standard error (m)
Fixed_122	−0.1388	0.0784	0.0019	−0.0346	0.236	269.429	0.449	<.0001	0.00212
Fixed_61	−0.0910	0.0506	0.0003	−0.0144	0.849	53.133	0.939	<.0001	0.00141
Quad_35	−0.0401	0.0538	0.0000	0.0001	0.987	4.589	0.992	<.0001	0.00057
Quad_20	−0.0275	0.0438	0.0000	0.0058	1.029	−10.133	0.994	<.0001	0.00051
Ground_2A	−0.0906	0.0659	0.0001	0.0066	1.023	−8.144	0.985	<.0001	0.00080
Ground_2B	−0.0534	0.0535	0.0000	0.0070	1.024	−8.505	0.996	<.0001	0.00040
Ground_4A	−0.0765	0.0723	0.0001	0.0001	0.985	5.428	0.986	<.0001	0.00074
Ground_4B	−0.0520	0.0509	0.0000	0.0066	1.019	−6.531	0.996	<.0001	0.00040
Ground_4C	−0.0541	0.0517	0.0000	0.0061	1.014	−4.810	0.994	<.0001	0.00049
Ground_6A	−0.0582	0.0506	0.0000	0.0059	1.020	−7.193	0.996	<.0001	0.00040
Ground_8A	−0.0566	0.0486	0.0000	0.0045	1.016	−5.689	0.996	<.0001	0.00039
Lidar	−0.0420	0.0394	0.0000	−0.0001	1.001	−0.257	0.998	<.0001	0.00029

Table 5. Simple statistics of comparative cross-sectional elevations generated using different surveying methods. Values were calculated for nine cross sections individually and then averaged.

Dataset	Minimum elevation (m)	Maximum elevation (m)	Mean elevation (m)	Variance elevation (m)
Fixed_122	352.551	352.628	352.596	0.0005
Fixed_61	352.492	352.673	352.576	0.0019
Quad_35	352.474	352.659	352.556	0.0025
Quad_20	352.474	352.659	352.556	0.0025
Ground_2A	352.463	352.667	352.555	0.0025
Ground_2B	352.473	352.659	352.554	0.0024
Ground_4A	352.474	352.666	352.561	0.0022
Ground_4B	352.472	352.660	352.555	0.0024
Ground_4C	352.472	352.658	352.555	0.0024
Ground_6A	352.472	352.660	352.555	0.0024
Ground_8A	352.475	352.660	352.557	0.0024
lidar	352.472	352.660	352.555	0.0024

the lidar range and appeared to be very similar (Fig. 6), with the exceptions of Ground_2A (too rough; 8.2 % roughness increase) and Fixed_122 (too smooth; 84 % roughness decrease).

3.4 Gridded cross-sectional evaluations

In the gridded elevation evaluations (min, max, mean; Table 5; Fig. 12), the absolute difference from lidar was less than 0.02 %, and these differences were only seen in the fixed-wing flights. The variance (i.e., roughness), however, shows that the absolute differences from lidar were 131 %

Table 6. Volume difference between photogrammetry and terrestrial lidar raster grids generated from three-dimensional point clouds.

Dataset	Volume difference (m ³)	Cut volume (m ³)	Fill volume (m ³)	Absolute volume difference* (m ³)
Fixed_122	−0.0422	0.0078	−0.0499	0.0577
Fixed_61	−0.0175	0.0021	−0.0196	0.0217
Quad_35	0.0002	0.0029	−0.0027	0.0056
Quad_20	0.0072	0.0079	−0.0007	0.0085
Ground_2A	0.0081	0.0093	−0.0012	0.0105
Ground_2B	0.0086	0.0088	−0.0002	0.0090
Ground_4A	0.0003	0.0032	−0.0029	0.0062
Ground_4B	0.0082	0.0084	−0.0002	0.0086
Ground_4C	0.0076	0.0080	−0.0005	0.0085
Ground_6A	0.0073	0.0076	−0.0002	0.0078
Ground_8A	0.0056	0.0059	−0.0003	0.0062

* Iterative accumulation of individual raster grid cell elevation differences using absolute values.

(Fixed_122), 23 % (Fixed_61), and 9 % (Ground_4A). A comparison of the elevation information (Table 7; Fig. 12) between photogrammetric cross sections and lidar cross sections through linear regression indicates a coefficient of determination larger than 0.98 for all datasets, excluding the two fixed-wing flights. The standard error for this regression was less than 10 mm for Quad_20, Quad_35, Ground_2B, Ground_4B, Ground_4C, Ground_6A, and Ground_8A. Following that, Ground_2A and Ground_4A had a standard error of approximately 17 mm and the two fixed-wing flights had a standard error of 25 mm. The average area percent differences for all cross sections were within 1.5 %, while the two fixed-wing flights had 3 % (Fixed_61) and 8 % (Fixed_122). It is important to mention that the range of area

Table 7. Cross-sectional evaluation comparison between photogrammetry and lidar.

Dataset	Fitting linear model between photogrammetry and lidar					Mean area* percent difference	Minimum area* percent difference	Maximum area* percent difference
	Slope	Intercept	<i>r</i> value	<i>p</i> value	Standard error			
Fixed_122	0.311	242.937	0.618	<0.0001	0.025	7.79 %	4.78 %	15.04 %
Fixed_61	0.747	89.336	0.882	<0.0001	0.025	3.18 %	1.77 %	4.44 %
Quad_35	1.008	−2.745	0.989	<0.0001	0.009	−0.05 %	−0.23 %	0.43 %
Quad_20	1.048	−17.059	0.990	<0.0001	0.009	−1.36 %	−1.70 %	−0.64 %
Ground_2A	0.969	11.030	0.972	<0.0001	0.015	−1.57 %	−2.49 %	−0.90 %
Ground_2B	1.019	−6.554	0.994	<0.0001	0.007	−1.63 %	−1.84 %	−1.43 %
Ground_4A	0.923	27.163	0.958	<0.0001	0.018	−0.03 %	−0.74 %	0.69 %
Ground_4B	1.028	−9.963	0.994	<0.0001	0.007	−1.55 %	−1.71 %	−1.38 %
Ground_4C	1.028	−9.737	0.993	<0.0001	0.008	−1.44 %	−1.61 %	−1.24 %
Ground_6A	1.030	−10.502	0.994	<0.0001	0.007	−1.38 %	−1.53 %	−1.08 %
Ground_8A	1.031	−10.823	0.996	<0.0001	0.006	−1.06 %	−1.26 %	−0.76 %

* Area calculations used a horizontal reference elevation of 353 m.

Table 8. Results from the ranking analysis based on the difference metrics of multiple photogrammetric surveys applied to gully channel monitoring.

Volume and cross section (M1–M5)		Spot elevation (M6–M10)		Normal to plane (M11–M15)		Combined (M1–M15)	
Dataset	Points	Points	Dataset	Points	Dataset	Points	Dataset
Ground_8A	49	Ground_8A	50	Ground_2B	44	Ground_8A	140
Quad_35	42	Ground_4B	41	Ground_6A	42	Ground_6A	121
Ground_6A	38	Ground_6A	41	Ground_8A	41	Ground_4B	114
Ground_4A	37	Quad_20	37	Ground_4B	41	Quad_35	111
Quad_20	35	Quad_35	36	Quad_20	35	Ground_2B	107
Ground_4B	32	Ground_4C	34	Quad_35	33	Quad_20	107
Ground_4C	32	Ground_2B	33	Ground_4C	31	Ground_4C	97
Ground_2B	30	Ground_4A	25	Ground_4A	26	Ground_4A	88
Ground_2A	20	Ground_2A	16	Ground_2A	15	Ground_2A	51
Fixed_61	9	Fixed_61	12	Fixed_122	13	Fixed_61	30
Fixed_122	6	Fixed_122	5	Fixed_61	9	Fixed_122	24

percent difference is within ± 2 %, while the fixed-wing systems had up to 15 % difference. The error is huge, for instance, if this dataset was intended to be used for the development, calibration, and validation of a soil erosion model.

3.5 Dataset scoring evaluations

The gridded data performance was led by Ground_8A and Quad_35. Combined category score points ranged from 49 (Ground_8A) to 6 (Fixed_122), terrestrial photogrammetry ranged from 49 to 20, quadrotor ranged from 42 to 35, and the fixed wing ranged from 9 to 6 (Table 8). For the point cloud analysis, the scoring results were, for the most part, very similar. The terrestrial photogrammetry surveys all score very high, with the quadrotor falling in the middle and the fixed wing at the bottom. Overall, scoring ranged from

140 to 24 with the terrestrial photogrammetry leading the group. As the number of photos increased, so did the sample density; however, the four-photo pair (Ground_8A) was less dense than the six-photo pair (Ground_6A) or the two-photo pair (Ground_4B, Ground_4C), which may be associated with higher accuracy in pixel matching or the addition of inferior images to the project. However, it is noteworthy to add that sampling intensity increases as the UAV altitude decreased, although the Quad_35 outperformed the Quad_20 in a number of categories.

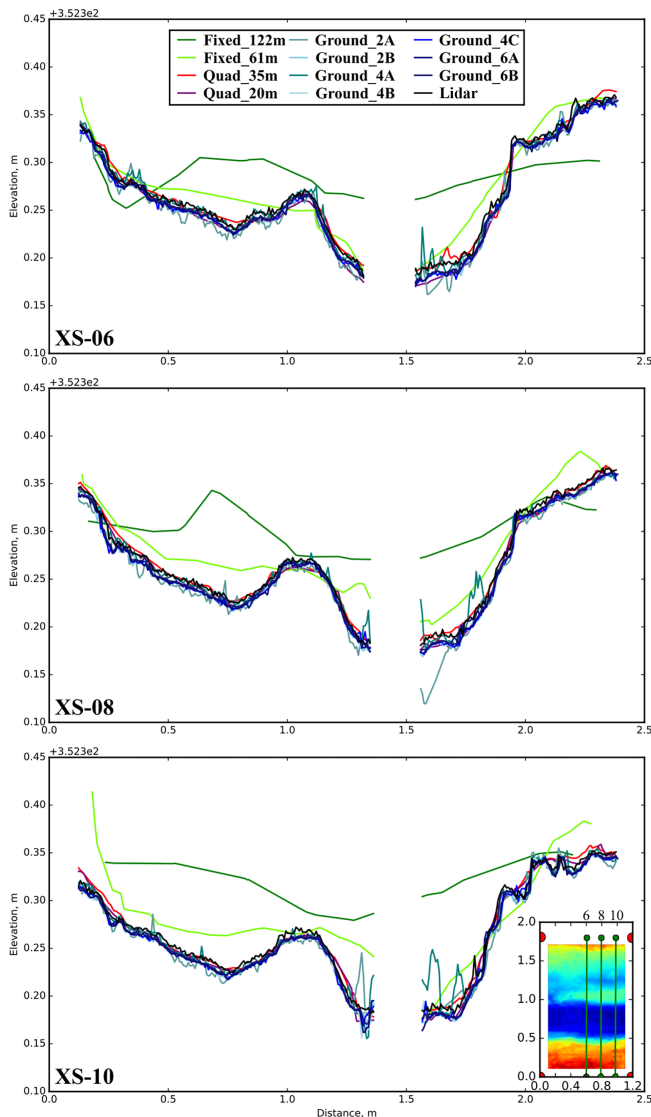


Figure 12. Selected cross sections generated from interpolating point clouds into a 5×5 mm raster grid file.

4 Discussion

4.1 Method comparison

Two photogrammetric software packages (Pix4DMapper Pro and PhotoModeler Scanner) were used to generate solutions for the UAV platform and terrestrial photogrammetry surveys. Pix4DMapper Pro uses a larger number (≥ 3) of overlapping photos, while PhotoModeler Scanner can offer solutions with only two overlapping photos. These software packages differ in the level of user control options for processing and point cloud generation. Point clouds processed by different software packages and/or users could yield very different solutions; however, this aspect was not investigated here.

An alarming concern in this analysis was the realization that rotation and translation (Fig. 4) were required to ensure that all data were properly aligned. The lidar global coordinates were the same as those used for the fixed-wing and quadrotor flights (i.e., field GCPs). The channel GCPs were also utilized to optimize the lidar point cloud solution. In all terrestrial photogrammetry point cloud solutions, the same set of global coordinates (channel GCPs) were used. One might expect the solutions to converge without the need to manipulate the point clouds in postprocessing; however, not one of the solutions contained the exact positions of the channel GCPs, including the solutions generated using the same platform but with varying processing parameters. For example, three-dimensional registration discrepancies were detected between lidar solutions and solutions from the quadrotor platform at 20 and 35 m, the fixed-wing platform at 61 and 122 m, and the terrestrial photogrammetry surveys. This realization presents extreme difficulty for temporal studies of ephemeral erosion processes, no matter the choice of resolution, platform, or processing parameters.

Initially, an attempt was made to analyze all datasets in their original form; however, two limitations to the approach were noted: the lack of three-dimensional registration between the datasets skewed efforts to quantify individual point accuracy and, more importantly, reduced confidence in the geomorphology information generated. The difference in point sampling density, ranging from hundreds (fixed-wing platform) to millions (lidar), biased the results. Therefore, the discussion presented herein relates to solutions that have been altered from the original solutions produced by the respective software packages. A comparison of the measured F and G functions with the estimated theoretical spatial distributions under the complete spatial randomness assumption suggested that all datasets did not present any spatial clustering, therefore indicating that the study site was sampled uniformly (regular spatially distributed data throughout the study site; Fig. 8). The main difference between datasets was the scale, at which terrestrial photogrammetry and quadrotor airborne photogrammetry yielded sub-centimeter distances as a result of the large number of samples when compared to the fixed-wing airborne photogrammetry. The results from the point evaluations suggest that the Fixed_122 data are clustered below $r = 20$ mm; however, the very same point cloud was interpreted to a 5 mm raster, so part or all of the metrics associated with these flights may be biased.

The normal projection and vertical spot analysis place the mean elevation for quadrotor flights at 2.9 mm below the lidar, for terrestrial photogrammetry at 5.0 mm below the lidar, and for the fixed-wing flights at 16 mm above the lidar. The range of cross-sectional elevation from all terrestrial photogrammetry was within 14 % of the lidar and, if we drop the rough sample (Ground_2A), the difference falls below 0.24 % (e.g., Gómez-Gutiérrez et al., 2014; Di Stefano et al., 2016). This is all seemingly acceptable for terrain mapping and perhaps even process development; how-

ever, if it is assumed that the bulk density of the soil is 1500 kg m^{-3} , then the soil mass difference for the quadrotor flights is 5.6 kg (erosion and elevation depletion mass). For terrestrial photogrammetry it is 9.8 kg (erosion and elevation depletion mass), and for the fixed-wing flights it is 44.8 kg (deposition and elevation enhancement mass). Furthermore, if the size of the study area (2.47 m^2) is projected onto a 1 ha field, the elevation distortion is anywhere from 23 000 to 181 000 kg of material, which is substantial. Another way to visualize these data would be to look at the area calculations, from which there is a 1.064 % decrease in cross-sectional area (Fig. 10) for Ground_8A over that of the lidar. One percent is a very low difference and amounts to an impacted area of 0.03 m^2 . Again, when considering the site projected onto the 1 ha field, the impacted area is on the order of 50 times the original measurement area (121.5 m^2). These findings are reflective of the decision to keep all collected data and further promote the importance of data uncertainty analysis (Wheaton et al., 2010).

Another interesting finding was the difference between solutions from terrestrial photogrammetry (varying number and/or orientation of photo pairs). The solutions from the Ground_2A and Ground_2B datasets both used only one photo pair; however, the results from the analysis indicate a superior solution generated from the Ground_2B pairing (i.e., upstream or downstream orientation; Table 1). This could be potentially attributed to the orientation of the images in relation to the channel, in which differences in illumination could hamper the photogrammetric process of automated pixel matching between each photo pair (Marzolf and Poesen, 2009). Additionally, increasing the number of photo pairs used in the solution seems to yield improved solutions. Results from the volumetric analysis show that the Quad_35 was a very close approximation (0.0002 m^3) to the lidar, and Ground_4A (i.e. channel left and right with corner left and right photo pair; Table 1) was within 0.0001 m^3 of the Quad_35. Solutions obtained with the “corner left and right photo pair” tended to improve the estimates. However, within the overall assessment of data performance, the Ground_4A data finished eighth, and small differences between Ground_6A and Ground_8A suggest a potential threshold in the number of photo pairs to which including additional photo pairs adds marginally to the final quality of the solution. Whether or not the datasets were adjusted spatially in accordance with the channel GCP positions, the absolute volume differences were similarly ranked between photogrammetry datasets. Point clouds built from higher photo pairs and flights at lower altitudes produced better results when compared to terrestrial lidar.

4.2 Monitoring guidelines

The long-term photogrammetric monitoring of ephemeral gullies should be performed with systems and procedures that strongly consider the following.

1. Provide a minimum sampling density to capture the overall and local terrain characteristics based on the study objectives (i.e., a temporal headcut migration process understanding may require data with sub-centimeter resolution, while a temporal channel meander process understanding may only require decimeter resolution; James and Robson, 2012; Gómez-Gutiérrez et al., 2014). The planning phase of the project must consider the physical characteristics of the process to be investigated, the study site physical and environmental variables, and the available hardware and software.
2. Utilize static ground control points visible in comparable photo pairs in all time-step surveys (i.e., fixed known points within the scene provide checks to ensure proper three-dimensional registration of temporal data; e.g., Smith and Vericat, 2015). An organized scheme for control points must be realized for a detailed multi-temporal quantitative assessment. Small variations in alignment within temporal surveys will introduce error into length, width, cross-sectional area, and volume estimates (e.g., Casalí et al., 2015). Repeated realizations of GCP coordinates will always reduce error in survey solutions.
3. Collect the same number of photo pairs using the same sensor and with the same orientation in all time-step surveys (i.e., data collection strategies should not vary temporally and new sensors must be carefully calibrated to preexisting datasets). Consistency in photo collection (i.e., scheduling and number of photo pairs) will enhance the comparison of temporal solutions (Gómez-Gutiérrez et al., 2014). Also, consider site visits at a particular time of day.
4. Process and generate photogrammetric solutions using the same software package and similar input processing parameters. A calibrated camera will always yield better solutions.

5 Conclusions

Comparative evaluations were completed using terrestrial lidar and photogrammetry, both terrestrial and aerial (UAV). None of these methods were without limitation, and the ultimate goal of the data collection effort should guide the planning phase of the project. One cautionary note: without GCP there is no reasonable expectation that temporal activities will be successful. Although GCP may increase the workload during data acquisition, this is the only realization that will ensure global alignment, minimize project error, and enhance process theory development. While adherence to conventional ground methods for GCP establishment is essential for accurate temporal terrain characterization, the results presented herein are transferrable to larger survey areas with different terrain and surface characteristics. In terms of survey

choice, all results point to financial and temporal questions. What is the project goal? What are the data expectations? A temporal assessment of gully channels and most geomorphic process descriptions can be accomplished with a camera and a few GCPs, whether on the ground or airborne. Each of the survey methods provided herein performed very well; although the scoring was not spectacular, the Fixed_61 data would be satisfactory for most static model evaluations. As expectations rise, so will the planning and technology.

Data availability. The data is available at: http://capone.mtsu.edu/hmomm/data_dist.html

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. The use of trade, firm, or corporation names in this paper is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

Acknowledgements. This research was partially funded by grant 1359852 from the US National Science Foundation. The authors would like to acknowledge the support provided by Nathan Stein (remote pilot for the quadcopter UAV), Daniel Murphy (remote pilot for the fixed-wing UAV), Justin Hobart, Tom Buman, Bob Buman, Sarah Anderson, and Rick Cruse.

Edited by: Anette Eltner

Reviewed by: Álvaro Gómez-Gutiérrez and one anonymous referee

References

- Aguilar, F. J., Agüera, F., Aguilar, M. A., and Carvajal, F.: Effects of terrain morphology, sampling density, and interpolation methods, *Photogram. Eng. Remote Sens.*, 71, 805–816, 2005.
- Aronoff, S.: *Remote Sensing for GIS Managers*, ESRI Press, Redlands, California, USA, 63–65, 2005.
- Bachrach, A., Prentice, S., He, R., Henry, P., Huang, A. S., Krainin, M., Maturana, D., Fox, D., and Roy, N.: Estimation, planning, and mapping for autonomous flight using an RGB-D camera in GPS-denied environments, *Int. J. Rob. Res.*, 31, 1320–1343, 2012.
- Baddeley, A. and Turner, R.: Spatstat: An R package for analyzing metric estimation of interaction in inhomogeneous point pattern, *Stat. Neerl.*, 54, 329–350, 2005.
- Bailey, T. C. and Gatrell, A. C.: *Interactive Spatial Data Analysis*, Longman, Harlow, 83–90, 1995.
- Bater, C. W. and Coops, N. C.: Evaluating error associated with lidar-derived DEM interpolation, *Comp. Geosci.*, 35, 289–300, 2009.
- Besl, P. and McKay, N.: A Method for Registration of 3-D Shapes, *IEEE T. Pattern Anal.*, 14, 239–256, 1992.
- Bivand, R. S., Pebesma, E. J., and Gomez-Rubio, V.: *Applied Spatial Data Analysis with R*, Springer, New York, New York, USA, 160–163, 2008.
- Bry, A., Richter, C., Bachrach, A., and Roy, N.: Aggressive flight of fixed-wing and quadrotor aircraft in dense indoor environments, *Int. J. Rob. Res.*, 37, 969–1002, 2015.
- Casalí, J., Loizu, J., Campo, M. A., De Santisteban, L. M., and Álvarez-Mozos, J.: Accuracy of methods for field assessment of rill and ephemeral gully erosion, *Catena*, 67, 128–138, 2006.
- Casalí, J., Giménez, R., and Campo-Bescos, M. A.: Gully geometry: what are we measuring?, *SOIL*, 1, 509–513, <https://doi.org/10.5194/soil-1-509-2015>, 2015.
- Castillo, C. and Gómez, J. A.: A century of gully erosion research: urgency, complexity and study approaches, *Earth-Sci. Rev.*, 160, 300–319, 2016.
- Castillo, C., James, M. R., Redel-Macías, M. D., Pérez, R., and Gómez, J. A.: SF3M software: 3-D photo-reconstruction for non-expert users and its application to a gully network, *SOIL*, 1, 583–594, <https://doi.org/10.5194/soil-1-583-2015>, 2015.
- Castillo, C., Pérez, R., James, M. R., Quinton, J. N., Taguas, E. V., and Gómez, J. A.: Comparing the accuracy of several field methods for measuring gully erosion, *Soil Sci. Soc. Am. J.*, 76, 1319–1332, <https://doi.org/10.2136/sssaj2011.0390>, 2012.
- Cook, K. L.: An evaluation of the effectiveness of low-cost UAVs and structure from motion for geomorphic change detection, *Geomorph.*, 278, 195–208, <https://doi.org/10.1016/j.geomorph.2016.11.009>, 2017.
- Csanyi, N., Toth, C. K., Grejner-Brzezinska, D., and Ray, J.: Improvement of LiDAR data accuracy using LiDAR specific ground targets, *Proceedings of the 2005 ASPRS Annual Conference*, Baltimore, Maryland, USA, 7–11 March, 2005.
- Dabney, S. M., Vieira, D. A. N., and Yoder, D. C.: Predicting ephemeral gully erosion with RUSLER and EphGEE, *Proceedings of Sediment Dynamics from the Summit to the Sea symposium*, New Orleans, Louisiana, USA, 11–14 December, 2014.
- Di Stefano, C., Ferro, V., Palmeri, V., and Pampalona, V.: Testing the use of an image-based technique to measure gully erosion at Sparacia experimental area, *Hydrol. Process.*, 31, 573–585, <https://doi.org/10.1002/hyp.11048>, 2017.
- Dodd, M.: Where are my quadrats? Positional accuracy in fieldwork, *Methods Ecol. Evol.*, 2, 576–584, <https://doi.org/10.1111/j.2041-210x.2011.00118x>, 2011.
- Eitel, J. U. H., Williams, C. J., Vierling, L. A., Al-Hamdan, O. Z., and Pierson, F. B.: Suitability of terrestrial laser scanning for studying surface roughness effects on concentrated flow erosion processes in rangelands, *Catena*, 87, 398–407, 2011.
- Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., Glennie, C. L., Joerg, P. C., LeWinter, A. L., Magney, T. S., Mandlbürger, G., Morton, D. C., Müller, J., and Vierling, K. T.: Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences, *Remote Sens. Environ.*, 186, 372–392, 2016.
- Eltner, A., Kaiser, A., Castillo, C., Rock, G., Neugirg, F., and Abellán, A.: Image-based surface reconstruction in geomorphometry – merits, limits and developments, *Earth Surf. Dynam.*, 4, 359–389, <https://doi.org/10.5194/esurf-4-359-2016>, 2016.

- ESRI: ArcGIS Desktop: Release 10, Redlands, CA, Environmental Systems Research Institute, available at: <http://resources.arcgis.com/en/help/main/10.2/index.html#//00q900000016000000>, last access: June 2016, 2011.
- Foster, G. R.: Modeling ephemeral gully erosion for conservation planning, *Int. J. Sed. Res.*, 20, 157–175, 2005.
- Gesch, K. R., Wells, R. R., Cruise, R. M., Momm, H. G., and Dabney, S. M.: Quantifying uncertainty of measuring gully morphological evolution with close-range digital photogrammetry, *Soil Sci. Soc. Am. J.*, 79, 650–659, 2015.
- Gómez-Gutiérrez, A., Schnabel, S., Berenguer-Sempere, F., Lavado-Contador, F., and Rubio-Delgado, J.: Using 3D photo-reconstruction methods to estimate gully headcut erosion, *Catena*, 120, 91–101, 2014.
- Hawdon, A. A., Wilkinson, S. N., Kinsey-Henderson, A. E., Goodwin, N., Bartley, R., and Baker, B. W.: Rapid acquisition of gully topography using a mobile, handheld laser scanner, *Proceedings of the 7th International Symposium on Gully Erosion*, West Lafayette, Indiana, USA, 23–27 May, 2016.
- Hodgson, M. E. and Bresnahan, P.: Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget, *Photogram. Eng. Remote Sens.*, 70, 331–339, 2004.
- Honkavaara, E., Eskelinen, M. A., Pölonen, I., Saari, H., Ojanen, H., Mannila, R., Holmlund, C., Hakala, T., Litkey, P., Rosnell, T., Vijanen, N., and Pulkkanen, M.: Remote sensing of 3-D geometry and surface moisture of a peat production area using hyperspectral frame cameras in visible to short-wave infrared spectral ranges onboard a small unmanned airborne vehicle (UAV), *IEEE T. Geosci. Remote*, 54, 5440–5454, 2016.
- James, M. R. and Robson, S.: Straightforward reconstruction of 3-D surfaces and topography with a camera: Accuracy and geoscience application, *J. Geophys. Res.*, 117, F03017, <https://doi.org/10.1029/2011JF002289>, 2012.
- James, M. R. and Robson, S.: Mitigating systematic error in topographic models derived from UAV and ground-based image networks, *Earth Surf. Proc. Land.*, 39, 1413–1420, 2014.
- Kienzle, S.: The effect of DEM raster resolution on first order, second order, and compound terrain derivatives, *Trans. GIS*, 8, 83–111, 2004.
- Kukko, A., Kaartinen, H., Hyypä, J., and Chen, Y.: Multiplatform mobile laser scanning: usability and performance, *Sensors*, 12, 11712–11733, 2012.
- Lloyd, C. D.: *Spatial Data Analysis – An Introduction for GIS Users*, Oxford University Press, New York, NY, USA, 2010.
- Marzolf, I.: Advances in image-based measuring and monitoring of gully erosion, *Proceedings of the 7th International Symposium on Gully Erosion*, West Lafayette, Indiana, USA, 23–27 May, 2016.
- Marzolf, I. and Poesen, J.: The potential of 3D gully monitoring with GIS using high-resolution aerial photography and a digital photogrammetry system, *Geomorphology*, 111, 48–60, 2009.
- Marzolf, I., Ries, J. B., and Albert, K.-D.: Kite aerial photography for gully monitoring in Sahelian landscapes, *Proceedings of the Second Workshop of the EARSel, Special Interest Group on Remote Sensing for Developing Countries*, 18–20 September, 2002, Bonn, Germany, CD-ROM publication, 2003.
- Micheletti, N., Chandler, J. H., and Lane, S. N.: Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smart phone, *Earth Surf. Process. Land.*, 40, 473–486, 2015.
- Milenković, M., Pfeifer, N., and Glira, P.: Applying terrestrial laser scanning for soil surface roughness assessment, *Remote Sens.*, 7, 2007–2045, 2015.
- Momm, H. G., Bingner, R. L., Wells, R. R., and Dabney, S. M.: Methods for Gully Characterization in Agricultural Croplands using Ground-Based Light Detection and Ranging, *Sediment Transport – Flow and Morphological Processes*, edited by: ABM Faruk Bhuiyan, InTech, 101–124, 2011.
- Momm, H. G., Bingner, R. L., Wells, R. R., Dabney, S. M., and Frees, L. D.: Effect of Terrestrial LiDAR Point Sampling Density in Ephemeral Gully Characterization, *Open J. Modern Hydrol.*, 3, 38–49, 2013a.
- Momm, H. G., Bingner, R. L., Wells, R. R., Rigby, J. R., and Dabney, S. M.: Effect of topographic characteristics on compound topographic index for identification of gully channel initiation locations, *Trans. ASABE*, 56, 523–537, 2013b.
- Piermattei, L., Carturan, L., de Blasi, F., Tarolli, P., Dalla Fontana, G., Vettore, A., and Pfeifer, N.: Suitability of ground-based SfM-MVS for monitoring glacial and periglacial processes, *Earth Surf. Dynam.*, 4, 425–443, <https://doi.org/10.5194/esurf-4-425-2016>, 2016.
- Ries, J. B. and Marzolf, I.: Monitoring of gully erosion in the Central Ebro Basin by large-scale aerial photography taken from a remotely controlled blimp, *Catena*, 50, 309–328, 2003.
- Rietdorf, A., Daub, C., and Loef, P.: Precise positioning in real-time using navigation satellites and telecommunication, *Proceedings of the 3rd Workshop on Positioning and Communication*, Hannover, Germany, 16 March 2006, 2006.
- Smith, M. W. and Vericat, D.: From experimental plots to experimental landscapes: topography, erosion and deposition in sub-humid badlands from structure-from-motion photogrammetry, *Earth Surf. Process. Land.*, 40, 1656–1671, 2015.
- Soil Science Society of America: Glossary of soil science terms, Madison: Soil Science Society of America, 19 pp., 2008.
- Vinci, A., Brgante, R., Todisco, F., Mannocchi, F., and Radicioni, F.: Measuring rill erosion by laser scanning, *Catena*, 124, 97–108, <https://doi.org/10.1016/j.catena.2014.09.003>, 2015.
- Wehr, A. and Lohr, U.: Airborne laser scanning – An introduction and overview, *ISPRS J. Photogram.*, 54, 68–82, 1999.
- Wells, R. R., Momm, H. G., Bennett, S. J., Gesch, K. R., Dabney, S. M., Cruise, R., and Wilson, G. V.: A measurement method for rill and ephemeral gully erosion assessments, *Soil Sci. Soc. Am. J.*, 80, 203–3214, 2016.
- Wheaton, J. M., Brasington, J., Darby, S. E., and Sear, D. A.: Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets, *Earth Surf. Process. Land.*, 35, 136–156, 2010.
- Zevenbergen, L.: Modelling erosion using terrain analysis. Ph.D. thesis. London, UK, Department of Geography and Earth Science, Queen Mary College, University of London, 1–345, 1989.
- Zhang, W. and Montgomery, D. R.: Digital elevation model grid size, landscape representation, and hydrologic simulations, *Water Resour. Res.*, 30, 1019–1028, 1994.
- Ziadat, F. M.: Effect of contour intervals and grid cell size on the accuracy of DEMs and slope derivatives, *Trans. GIS*, 11, 67–81, 2007.