Quantifying uncertainty of remotely sensed topographic surveys for ephemeral gully channel monitoring

Robert R. Wells[1], Henrique G. Momm[2], and Carlos Castillo[3]

[1]National Sedimentation Laboratory, Agricultural Research Service, United States Department of Agriculture, Oxford, Mississippi,
5 38655, U.S.A.
[2]Department of Geosciences, Middle Tennessee State University, Murfreesboro, Tennessee, 37132, U.S.A.
[3]Department of Rural Engineering, University of Córdoba, Córdoba, Spain.

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

Abstract. Spatio-temporal measurements of landform evolution provide the basis for process-based theory formulation and
10 validation. Overtime, field measurement of landforms has increased significantly worldwide, driven primarily by the availability
of new surveying technologies. However, there is not a standardized and/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 can lead to incorrect estimates of landform change as a result of three-dimensional registration errors. This work evaluated
15 measurements of an ephemeral gully channel located in agricultural land using multiple independent survey techniques for
locational accuracy and their applicability to generate information for model development/validation. Terrestrial photogrammetry
and un-maned aerial vehicle-based platforms were compared to terrestrial LiDAR, defined herein as the reference dataset. Given
the small-scale of the landform measured, alignment and ensemble equivalence between data sources were addressed sine qua non
through post-processing. Utilization of ground control points were prerequisite to three-dimensional registration between datasets
20 and improved confidence in the morphology information generated. Not one of the methods was without limitation; however,
careful attention to project pre-planning and data nature will ultimately guide temporal efficacy and practicality of management
decisions.

Keywords: close-range photogrammetry, un-manned aerial vehicles, ephemeral gully, soil erosion, remote sensing.

1 Introduction

25 Spatio-temporal measurements of landform evolution provide the basis for process-based theory formulation and validation. Field
measurement of landforms has increased significantly worldwide, driven by the availability of new surveying technologies. Recent
improvements to existing systems, such as mobile light detection and ranging (LiDAR) systems (Kukko et al., 2012; Eitel et al.,
2016; Hawdon et al., 2016), integrated Un-manned Aerial Vehicles (UAV) (Bachrach et al., 2012; Bry et al., 2015; Honkavaara et
al., 2016), Real Time Kinematics (RTK) (Rietdorf et al., 2006) and photogrammetry techniques (James and Robson, 2014;
30 Marzolff, 2016), and low-cost and/or freeware structure from motion (SfM) photogrammetric software (Castillo et al., 2016;
Piermattei et al., 2016). However, there is not a standardized and/or coordinated effort within the scientific community to collect
morphological data in a dependable and reproducible manner, specifically when performing long-term process investigation
studies.

Ephemeral gullies are often defined as small channels, on the order of a few centimeters in depth, predominantly in agricultural
35 fields (SSSAJ, 2008). The emergence, evolution and persistence of these concentrated flow path erosion features are controlled by
the combined effects of flow, slope, soil properties, topography and vegetation characteristics (Zevenbergen, 1989). The term

ephemeral refers to the fact that agricultural producers often erase these channels during regular farming operations (Foster, 2005). The combination of a highly dynamic lifespan with the relatively small-scale channel features require high accuracy measurements with high temporal and spatial resolution.

Many studies have been conducted to assess topographic accuracy of ephemeral/classical gully morphological measurements using a wide range of systems (i.e. Casalli, et al., 2004). Among them, LiDAR studies primarily focused on quantifying locational error from datasets generated by air-borne systems, where locational variations are the result of coalesced errors generated by inaccuracies of Global Positional Systems (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 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 topographic information using airborne LiDAR are often compared with discrete sample locations and/or man-made targets with known coordinates (Hodgson and Bresnahan, 2004; Csany 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 quantitative monitoring of small-scale geomorphological characteristics (i.e. ephemeral gullies) necessary for process description.

At a finer scale, investigation of ground-based/terrestrial LiDAR has demonstrated a high locational accuracy (~2mm) and noted the importance of appropriate spatial sampling density for ephemeral and classical gully investigation (Momm et al., 2013a). Topographic representation of gully channels requires datasets with a specific minimum sampling density, which is dependent on site-specific topographic characteristics (Castillo et al., 2012; Momm et al., 2013a). Overlapping the same area by multiple scans increases the overall sampling density to overcome this limitation.

Studies involving various surveying techniques of concentrated flow paths 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 measurement 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 (Marzolff, et al., 2003), backpacks (Wells et al., 2016) and UAVs (Ries and Marzolff, 2003; James and Robson, 2014; Bachrach et al., 2012; Cook, 2017). Erosion monitoring programs based on photogrammetry have several advantages when compared to other surveying techniques. Photogrammetric field surveys do not interfere with farming operations, as it is non-obstructive; field campaigns are extremely efficient; and often they do not require specialized technical skillsets to implement. However, photogrammetric results can vary as a function of controlling parameters used during data collection and processing (Eltner et al., 2016).

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 Geographic Information Systems (GIS)-based soil erosion modeling technology 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 and interpolation algorithms (Aguilar et al., 2005; Ziadat, 2007; Bater and Coops, 2009; James and Robson, 2012; James and Robson, 2014), and by 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 for measuring ephemeral gully channels. Three surveying methods with varying parameters were conducted

independently and evaluated for locational accuracy and their applicability in generating information for model development/validation. The objectives of this study are twofold: to quantify the overall accuracy of the different survey methods with varying configurations and to develop practical guidelines for the design and implementation of future ephemeral gully monitoring studies.

5 2 Methods

2.1 Study Site

The study site was located in the northwest corner of Webster County, Iowa, U.S.A. (Figure 1). Farming is the dominant enterprise in Webster County. The crop rotation was a corn-soybean rotation. Total annual precipitation is about 873 mm, of which, 70 percent usually falls between April and September. The area of interest (AOI) within the field survey was a small reach (1.9 m x

10 1.3 m; 2.47m^2) of a 150 m long ephemeral gully, oriented south to northwest, eroding Clarion loam (fine-loamy, mixed, superactive, mesic Typic Hapludolls) at the upper (south) extent, Terril loam (fine-loamy, mixed, superactive, 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. Within the AOI, the soil was Clarion loam.

2.2 Field Survey

15 Field surveys were conducted using three independent modes: ground-based/terrestrial LiDAR, ground-based photogrammetry, and airborne photogrammetry. The surveys yielded twelve datasets (Table 1; dashed lines represent delineation between survey modes). The terrestrial LiDAR was considered the reference dataset, due to perceived superior accuracy. All surveys were run independently of each other and completed on the same day. Each dataset was represented using the NAD83 UTM15N coordinate system.

20 In this study, the terrestrial LiDAR point cloud was generated using TopCon ScanMaster software (https://www.topconpositioning.com/software/mass-data-collection/scanmaster), all ground-based 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

25 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 GPS of Ground Control Points

Site preparation began by locating a state monument point (Figure 2) and laying out 406 x 406 mm quad-triangle, white-on-black sheet GCPs, considering the long and short axes of the field as well as the high and low elevations within the field boundary, herein

30 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 (4 total). All GCPs were surveyed using TopCon GR-3 dGPS survey equipment (TopCon Corporation, Tokyo, Japan) to obtain relative position in reference to the state monument point. A static occupation (6hrs) was initiated with the base station, then all GCPs (field, channel and state monument) were surveyed with the rover (20 sec collection interval). All survey data was corrected using an OPUS (National Geodetic Survey)

solution of the base station location, processed with reference to the state monument point. 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). The system operates
5 in 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 post-processing by entering the external geometry coordinates (i.e. absolute position determined from Post-Processed Kinematic survey) of GCPs. The AOI (demarked by channel GCPs; 2.47 m^2) within the field boundaries, covering the gully
10 channel, was surveyed with one scan resulting in a total of 5,613,334 scan points.

Given the level of user control of 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 high sampling intensity, it is common for the same location on the ground to be hit by multiple laser pulses. This
15 yields datasets with high sampling intensity but with a range of elevation values for the same location (i.e. fluff) given the vertical accuracy of the system. In this study, this locational variability is estimated to be approximately ±2 mm. The sensor operates in the near-infrared portion of the electromagnetic spectrum (1,535 ηm) and in this spectral range; electromagnetic energy is absorbed by water (Aronof, 2005). In locations with water features, the sensor emits laser pulses but no laser pulse is reflected back to the sensor, preventing the range calculation for that particular pulse. During data collection for this survey, a shallow film of water
20 was present in the gully channel (Figure 3A) and, as result, no points were recorded in the water-covered region (Figure 3B). Sampling gaps in LiDAR surveys can also be attributed to vertical features that limit the sensor line-of-site, referred to as shadowing (Figure 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
25 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 (Figure 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 to omit GCP occlusions from the dataset.

2.2.3 Ground-based Photogrammetric Survey

30 Ground-based/terrestrial photogrammetry was conducted using a Nikon D7000 16.2 MP camera (Nikon Inc., Melville, NY) with calibrated 20mm 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 to point clouds using PhotoModeler Scanner photogrammetric software. Initial data processing
35 included: aerial triangulation and bundle adjustment, camera position and exterior orientation. Following initial processing, channel GCP positions were included to optimize point cloud accuracy.

2.2.4 UAV-based Photogrammetric Survey

Still images captured by the UAVs were transformed to point clouds using Pix4DMapper Pro photogrammetric software. Initial data processing included: camera calibration, aerial triangulation and bundle adjustment, camera position and exterior orientation. Following initial processing, field GCP positions were included to optimize point cloud accuracy.

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

10 with battery level and image acquisition progress, while the autopilot continuously analyzed onboard control data to optimize the flight.

2.2.4.1 Fixed-wing UAV

The fixed-wing platform had an 18.2 MP nadir camera (i.e. belly mount) and was deployed to capture the entire field boundary (Figure 2) by throwing the craft in the air, then the craft flies, captures images and lands itself. Two altitudes were flown: 122 m

15 and 61 m.

2.2.4.2 Quadrotor UAV

The quadrotor platform had a 38 MP camera mounted within an 180° vertical range head and was deployed to capture both, the extent of the gully within the field and specific points of interest (i.e. gully cross sections; Gesch et al. 2015; Wells et al. 2016). The quadrotor was deployed through mission planning software. The craft takes off, flies and captures images, then lands itself.

20 Two altitudes were flown: 35 m and 20 m.

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 Figure 2B) were used to generate a rectangular polygon to subset the point clouds in all surveys, assuring that all surveys cover exactly the same ground

25 position. Second, a smaller polygon subset was performed by generating a polygon with a 5 cm reduction on all sides. This was performed to exclude areas close to the channel GCPs and assure the elimination of shadowing and occlusion created by the slightly elevated channel GCPs. Third, the sampling void within the LiDAR dataset, created from the presence of a thin film of water within the channel (Figure 3A), was manually digitized into another polygon and used to remove points from all photogrammetric-generated datasets to assure uniformity between all datasets. Essentially, instead of using interpolated data within this void in the

30 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 in the LiDAR data.

Subsequently, since each surveying method was performed using the same set of GCPs, manual inspection of measured points located coincident with channel GCPs (white square targets in Figure 2B) were used to generate planes, yielding ten planes, 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

35 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 Figure 4). Three-dimensional locational differences between the reference plane generated using the GPS survey (black squares in Figure 4) and the planes of each surveyed dataset (LiDAR and photogrammetry) were calculated using the Iterative Closest Point (ICP) algorithm (Besl and Mckey, 1994) implemented in MATLAB (MathWorks Inc., Natick, Massachusetts). The ICP algorithm minimizes the locational differences between two sets of three-dimensional point

5   clouds and outputs a 3x3 rotation angle matrix, R, and a 1x3 translation vector, 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.

$$
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{ref} = T + R_{\omega\varphi\kappa} * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{meas}
\tag{1}
$$

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
10   points within datasets (hundreds to millions), which tends to bias the results; therefore, the photogrammetry data were gridded to introduce a common means of discussing differences between the survey methods. Furthermore, 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).

In the point cloud analysis, each point within the photogrammetry surveys was compared to that within the LiDAR survey. The
15   analysis was 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 TIN (each photogrammetry point was projected up or down to intersect a TIN surface of the LiDAR points).

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
20   reveal over/under sampled locations. Under sampled locations can potentially be sources of positional errors, especially in surfaces with high relief. Sampling intensity was evaluated using the quadrant method (Dodd, 2011), where 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. Point pattern analysis was performed using the G Function (distribution of nearest neighbor distance differences) and F Function (distribution of nearest neighbor to virtual grid centroid distance differences) to evaluate potential irregularities in
25   sampling, such as over sampling one area while under sampling another (Bailey and Gatrell, 1995), and overall spatial distribution of the data sampling (i.e. clustered or disperse).

Similarly, the local elevation variance was evaluated by calculating the elevation range (difference between maximum and minimum) within each grid. The local elevation variance is a function of the terrain characteristics, sampling intensity, survey method, and post-processing parameters.

2.4.2   Vertical and Horizontal Displacement using Point Tangential Projection to Plane

30

Vertical and horizontal displacement between the LiDAR and photogrammetry measurements in each dataset were quantified using the normal projection of each photogrammetry point into a plane fitted to the nearest LiDAR points (Figure 5A). For each photogrammetry point (green circle in Figure 5A), the nearest (within 2.5cm sphere) LiDAR points were selected (black dots in Figure 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 Figure 5A) recorded and statistics were generated (Table 2). This analysis was performed for all datasets using an in-house developed software package, where

$$D_x = \left(Z_x^{plane} - Z_x^{photo}\right)$$
$$D_y = \left(Z_y^{plane} - Z_y^{photo}\right) \tag{2}$$
$$5 \quad D_z = \left(Z_z^{plane} - Z_z^{photo}\right)$$

2.4.3 Vertical Displacement using Spot Elevation

The three-dimensional point cloud representing the reference dataset (LiDAR) was converted into a Triangular Irregular Network (TIN) (Figure 5B). Each point in the photogrammetry (photogrammetry) dataset (green circle in Figure 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
10 intersected the TIN (red circle in Figure 5B). The Z-coordinate at intersection was recorded. This analysis was performed using ArcGIS (Esri, 2011).

$$D_z = \left(Z_z^{TIN} - Z_z^{photo}\right) \tag{3}$$

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 (Table 3).

2.4.4 Gridded Surface Assessment

All point clouds were converted into 5x5 mm regular raster grids using linear interpolation. The two upstream channel GCPs served as the base of the rectangular grid; roughness of individual datasets was highly dependent upon 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 & 5; Table 4).

$$20 \quad V_{diff} = \sum_{i=1}^{n}\left[\left(Z_i^{lidar} - Z_i^{photo}\right) * cz\right] \tag{4}$$
$$V_{acc} = \sum_{i=1}^{n}\left[\left|Z_i^{lidar} - Z_i^{photo}\right| * cz\right] \tag{5}$$

where Z^{lidar} is the reference elevation, Z^{photo} is photogrammetry elevation, cs is the raster grid cell size (0.005), and n is the total number of raster grid cells.

2.4.5 Gridded Cross-Section Assessment

25 Gully modeling technologies often use cross-sections as basic modeling units. With the objective to assess the error introduced by each survey to cross-section analysis (Figure 10), the raster grid surfaces were used to generate eleven (11) cross-sections. Area above the curve was selected as one of the metrics to quantify cross-section accuracy, given as

$$A_c = \sum_{i=1}^{n}\left[(353.00 - Z_i) * \Delta_{dist}\right] \tag{6}$$

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
30 between points in the cross-section (0.005 m). The constant (353.00) was used to adjust all cross-section elevations, due to local elevation relation to mean sea level, thereby truncating area values. The deviation of the area estimates from the LiDAR were calculated using:

$$A_{dev} = \left(\frac{A_{obs} - A_{est}}{A_{obs}}\right) * 100 \tag{7}$$

Quantification of cross-section elevation descriptive statistics (Table 5; Figure 11) and fitted linear models between LiDAR and photogrammetry elevations for individual cross-sections were also used as comparison metrics (Table 6).

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 7). 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→11) for increasing volume difference and, correlation coefficients were assigned decreasing scores (1→11) for decreasing correlation. Simply, if a variable had a positive impact, it received a higher score. Each score is defined in Table 7.

3 Results

3.1 Sampling Intensity Evaluations

The point clouds presented here had a large variability in sampling intensity (from 1 to > 250 points per cm^2). 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 (Figure 11). In Figure 6, point counts are very low for the fixed-wing flights in comparison to the other methods and the sparse point count leads to interpolation (filling) during raster gridding. Variance shows roughness is similar for most of the surveys, with the exception of Fixed_122 where 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 (Figure 8) and enhanced interpolation between points for the high spatial resolution raster grid (0.005m cell size).

3.2 Vertical and Horizontal Displacement Evaluations

Graphical representation of both spot elevation to TIN and normal to fitted plane for the fixed-wing flight at 122 m altitude (Fixed_122 (Figure 6)), and the four photo-pair (Ground_8A (Figure 7)) are provided for comparative purposes.

Very similar results were obtained with these two methods. All datasets had negligible mean displacement in the X and Y-directions (Table 2). The standard error of the normal to plane analysis (Table 2) 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 2) ranged from 0.2 mm (Quad_35) to 7.4 mm (Fixed_61). For the vertical spot to TIN analysis (Table 3), standard error ranged from 0.4 mm (Ground_8A, Ground_6A, Ground_4B, Ground_2B) to 2.1 mm (Fixed_122), and mean displacement ranged from 4.5 mm (Ground_8A) to 35 mm (Fixed_122). The mean elevation difference between photogrammetry and LiDAR was approximately 5mm, with the exceptions of the fixed-wing surveys (10 and 35 mm). Perhaps this was simply because they were not pre-processed (affine transformation was not applied); however, this result does show 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 5mm for all datasets (Tables 2 and 3). 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 2), 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
5   raster grids affected each dataset differently. For example, the LiDAR dataset contained a high sampling intensity (> 100 points per cm^2) with relatively large elevation variability of the points within a raster grid cell. Therefore, the interpolation procedure generated a significantly smoothed surface. Conversely, the fixed–wing surveys had a low sampling intensity and the interpolation procedure linearly filled the gaps, potentially generating a surface that differs from natural-looking surfaces.

One of the most important measurements for gully monitoring is the volume difference between surfaces (Table 4). Given the
10   small scale of this type of erosional feature (on the order of few cm), it is vital to have a good understanding of the expected error for each method. Among similar collection methods (ground-photogrammetry), the absolute volume difference (Table 4) ranged from 0 to 52%, although these differences were extremely small in reality (i.e. range of 0.0062 to 0.0105 m^3). 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 165% absolute volume difference for Ground_8A and Fixed_122,
15   respectively, in comparison to Quad_35; Table 4). Variance values 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, most were smaller than 2.5 cm per cm^2 and appeared to be very similar (Figure 7), with the exceptions of Ground_2A (too smooth) and LiDAR (too rough).

3.4 Gridded Cross-section Evaluations

20   Comparing elevation information (Tables 5 and 6) between photogrammetric cross-sections and LiDAR cross-sections through linear regression, indicates a coefficient of determination larger than 0.95 for all datasets, except the two fixed-wing systems. The standard error of 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 20 mm and the two fixed wings of 25 mm. The average wet area percent difference for all cross-sections were within 1.5%, while the two fixed wings had
25   3 and 7%, for Fixed_61 and Fixed_122, respectively. It is important to mention that the range of wet area percent difference is within ± 2%, while the fixed-wing systems had up to 15% percent difference. The error is huge for instance if this dataset was intended to be used for the purpose of development/calibration/validation of a soil erosion model.

3.5 Dataset Scoring Evaluations

Gridded data performance was led by Ground_8A and Quad_35. Combined category score points ranged from 49 (Ground_8A)
30   to 6 (Fixed_122), ground-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, scoring results were, for the most part, very similar. The ground-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 ground-photogrammetry leading the group. As the number of photos increased, so too did the sample density; however, the four photo-pair (Ground_8A) was less dense than either the six photo-pair (Ground_6A) or two photo-pair
35   (Ground_4B, Ground_4C), which may be associated with higher accuracy in pixel matching or the addition of inferior image/s to

the project. However, it is noteworthy to add that sampling intensity did increase as the UAV altitude decreased, although the Quad_35 did outperform the Quad_20 in a number of categories.

4 Discussion

4.1 Methods Comparison

5 An alarming concern in this analysis was the realization that rotation and translation (Figure 4) was required to ensure that all data was properly aligned. The LiDAR global coordinates were the same as those used for the fixed-wing and quadrotor flights (i.e. field GCPs) and, the channel GCPs were also utilized to optimize the LiDAR point cloud solution. In all ground-photogrammetry point cloud solutions, the same set of global coordinates (channel GCPs) were used. One might expect that the solutions would converge without the need to manipulate the point clouds in post-processing; however, not one of the solutions contained the exact

10 positions of the channel GCPs, including solutions generated using the same platform but with varying processing parameters. For example, three-dimensional registration discrepancies were LiDAR between solutions from the quadrotor platform at 20 and 35 meters, from the fixed-wing platform at 61 and 122 meters, and from the terrestrial photogrammetry surveys. This realization does offer extreme difficulty in temporal studies of ephemeral erosion processes, no matter the choice of resolution, platform, and/or processing parameters.

15 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; and the difference in point sampling density, ranging from hundreds (fixed-wing platform) to millions (LiDAR), biased results. Therefore, the discussion presented herein relates to solutions that have been altered from the original solutions produced by the respective software packages. And, two point

20 pattern analysis techniques (F and G functions; see Sec. 4.2.1) were completed to evaluate the spatial distribution of the data sampling (clustered or disperse). Comparison of measured F and G functions with estimated theoretical spatial distribution 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 (regularly spatially distributed throughout the study site). These point pattern analyses were performed using the spatstat package in the R software package (Bivand et al., 2008, Baddley and Turner, 2005). Furthermore,

25 points with the same X, Y and Z to the fifth decimal place were removed from the LiDAR dataset, indicating collection of redundancy information.

Another interesting finding was the difference between solutions from ground-photogrammetry (varying number and/or orientation of photo pairs). Solutions from Ground_2A and Ground_2B datasets both used only one photo pair, however, results from the ranking analysis indicate a superior solution generated from the Ground_2B (i.e. upstream/downstream orientation; Table 1)

30 pairing. 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 (Marzoff and Poesen, 2009). Additionally, increasing the number of photo pairs used in the solution, does seem to yield improved solutions; however, the 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.

4.2 Monitoring Guidelines

Two photogrammetric software packages (Pix4DMapper Pro and PhotoModeler Scanner) were used to generate solutions for the UAV platform and ground-based 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

5   user control options for processing and point cloud generation and, point clouds processed by different software packages and/or users could yield very different solutions.

Long-term photogrammetric monitoring of ephemeral gullies should be performed with systems and procedures that: (1) provide a minimum sampling density to capture the overall and local terrain characteristics based upon study objectives (i.e. temporal headcut migration process understanding may require sub-centimeter resolution data, while temporal channel meander process

10   understanding may only require decimeter resolution), (2) utilize static ground control points visible in comparable photo pairs in all time-step surveys (fixed known points within the scene provide checks to assure proper three-dimensional registration of temporal data), (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), and (4) process and generate photogrammetric solutions using the same software package and similar input processing

15   parameters (perhaps a topic for further study; however, temporal studies should be processed using the same input parameters).

5 Conclusions

Comparative evaluations were completed using terrestrial LiDAR and photogrammetry, both ground and aerial (UAV). None of these methods are 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

20   GCP may increase the workload during data acquisition, this is the only realization that will assure global alignment, minimize project error and enhance process theory development. In terms of survey choice, all results point to a financial and temporal question. What is the project goal? What are the data expectations? 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 and, although the scoring was not very spectacular, the Fixed_61 data would be

25   satisfactory for most static model evaluations. As expectations rise, so too will the planning and technology.

Acknowledgements

This research was partially funded by grant 1359852 from the U.S. National Science Foundation. Authors would like to acknowledge the support provided by Nathan Stein, Daniel Murphy, Tom Buman and Justin Hobart.

References

Aguilar, F.J., Aguera, F., Aguilar, M.A., and Carvajal, F.: Effects of terrain morphology, sampling density, and interpolation methods, Photogram. Eng. and Remote Sens., 71(7): 805-816, 2005.

Aronoff, S. Remote Sensing for GIS Managers, ESRI Press. Redlands, California, USA, 2005.

5 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(11): 1320-1343, 2012.

Baddeley, A. and Turner, R.: Spatstat: An R package for analyzing metric estimation of interaction in inhomogeneous point pattern, Statistica Neerlandica, 54: 329-350, 2005.

10 Bailey, T.C. and Gatrell, A.C.: Interactive Spatial Data Analysis. Longman, Harlow, 1995.

Bater, C.W. and Coops, N.C.: Evaluating error associated with lidar-derived DEM interpolation, Computer & Geosicences, 35: 289-300, 2009.

Besl, P. and McKay, N.: A Method for Registration of 3-D Shapes, IEEE Trans. Pattern Analysis and Machine Intelligence, 14: 239–256, 1992.

15 Bivand, R.S., Pebesma, E.J. and Gomez-Rubio, V.: Applied Spatial Data Analysis with R, Springer, New York, New York, USA, 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(7): 969-1002, 2015.

Castillo, C. and Gómez, J.A.: A century of gully erosion research: urgency, complexity and study approaches, Earth-Science Reviews, 160: 300-319, 2016.

20 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, doi: 10.21.2136/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. doi:10.1016/j.geomorph.2016.11.009. 2017.

25 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, U.S.A., March 7-11. 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, U.S.A., 11–14 December. 2014.

Dodd, M.: Where are my quadrats? Positional accuracy in fieldwork, Methods in Ecology and Evolution, 2: 576-584. doi: 30 10.1111/j.2041-210x.2011.00118x. 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., Mandlburger, 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 Sensing of Environment, 186: 372-392, 2016.

Eltner, A., A. Kaiser, C. Castillo, G. Rock, F. Neugirg, and A. Abellá. 2016. Image-based surface reconstruction in 35 geomorphemetry – merits, limits and developments. Earth Surf. Dynam., 4: 359-389. doi: 10.5194/esurf-4-359-2016.

ESRI.: ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute, 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. of Am. J., 79(2): 650-659, 2015.

40 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, May 23-27, 2016.

Hodgson, M.E. and Bresnahan, P.: Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget, Photogram. Eng. and Remote Sens., 70(3): 331–339, 2004.

45 Honkavaara, E., Eskelinen, M.A., Pölönen, I., Saari, H., Ojanen, H., Mannila, R., Holmlund, C., Hakala, T., Litkey, P., Rosnell, T., Vijanen, N., and Pulkanen, 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 Trans. On Geoscience and Remote Sensing, 54(9): 5440-5454, 2016.

James, M.R. and Robson, S.: Straightforward reconstruction of 3-D surfaces and topography with a camera: Accuracy and 50 geoscience application, J. Geophys. Res., 117, F03017, doi: 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. Process. Landforms, 39: 1413-1420, 2014.

Kienzle, S.: The effect of DEM raster resolution on first order, second order, and compound terrain derivatives, Trans. in GIS, 8(1): 83-111, 2004.

55 Kukko, A., Kaartinen, H., Hyyppä, J. and Chen, Y.: Multiplatform mobile laser scanning: usability and performance, Sensors, 12: 11712-11733, 2012.

Marsolff, I.: Advances in image-based measuring and monitoring of gully erosion, Proceedings of the 7th International Symposium on Gully Erosion, West Lafayette, Indiana, U.S.A., May 23-27, 2016.

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

Marzolff, 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.

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, ABM Faruk Bhuiyan (Ed.), ISBN 978-953-307-374-3, InTech, 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 Journal of Modern Hydrology, 3(1): 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. of ASABE, 56(2): 523-537, 2013b.

Piermattei, L., Carturan, L., de Blasi, F., Tarrolli, P., Fontana, G.D., Vettore, A., and Pfeifer, N.: Suitability of ground-based SfM-MVS for monitoring glacial and periglacial processes. Earth Surf. Dynam., 4: 425-443. doi:10.5194/esurf-4-425-2016, 2016.

Ries, J.B. and Marzolff, 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 (WPNC '06; http://wpnc.net/proceedings06.html), Hannover, Germany, 16 March, 2006.

Soil Science Society of America: Glossary of soil science terms. Madison: Soil Science Society of America. 2008.

Wehr, A. and Lohr, U.: Airborne laser scanning – An introduction and overview, ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3): 68-82, 1999.

Wells, R.R., Momm, H.G., Bennett, S.J., Gesch, K.R., Dabney, S.M., Cruse, R., and Wilson, G.V.: A measurement method for rill and ephemeral gully erosion assessments, Soil Sci. Soc. Am. J., 80: 203-3214, 2016.

Zevenbergen, L.: Modelling erosion using terrain analysis. Ph.D. thesis. London, UK: Department of Geography and Earth Science, Queen Mary College, University of London, 1989.

Zhang, W. and Montgomery, D.R.: Digital elevation model grid size, landscape representation, and hydrologic simulations, Water Resourc. Res., 30(4): 1019-1028, 1994.

Ziadat, F.M.: Effect of contour intervals and grid cell size on the accuracy of DEMs and slope derivatives, Trans. in GIS, 11(1): 67-81. 2007.

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

Dataset Identification	Dataset Description
Grd_2A	Channel left and right photo pair
Grd_2B	Upstream and downstream photo pair
Grd_4A	Channel left and right with corner left and right photo pair
Grd_4B	Upstream and downstream with corner left and right photo pair
Grd_4C	Upstream and downstream with channel left and right photo pair
Grd_6A	Upstream, downstream, channel left and right with corner left and right photo pair
Grd_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 delineation between survey modes.

Table 2: Statistics comparing photogrammetry and terrestrial LiDAR using the point normal projected into fitted plane analysis. Residual values were calculated based on coordinate difference of all raster grid cells.

Dataset	Minimum Z axis difference (m)	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 difference (m)	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	12467.67	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	-11557.52	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	-10955.43	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 difference (m)	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 3: 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 4: Volume difference between photogrammetry and terrestrial LiDAR raster grids generated from three-dimensional point clouds.

Dataset	Volume difference (m³)	Cut volume (m³)	Fill volume, (m³)	Absolute volume difference* (m³)
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

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

Dataset	Minimum elevation (m)	Maximum elevation (m)	Average 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

Table 6: Cross-section evaluation comparison between photogrammetry and LiDAR.

Dataset	Slope	Intercept	r-value	p-value	Standard error	Wet area* percent difference	Minimum wet area* percent difference	Maximum wet area* percent difference
	Fitting linear model between photogrammetry and LiDAR							
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%

* Wet area calculations used a horizontal reference elevation of 353 meters.

Table 7: Metrics used in the ranking analysis of 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-section comparisons	Averaged coefficient of correlation between photogrammetric and LiDAR cross-section elevation values derived from raster grid analysis.
M4	Cross-section comparisons	Averaged standard error between photogrammetric and LiDAR cross-section elevation values derived from raster grid analysis.
M5	Cross-section comparisons	Averaged wet area percent difference between photogrammetric and LiDAR cross-section 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 neighbors 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.

Table 8: Results from the ranking analysis based on 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
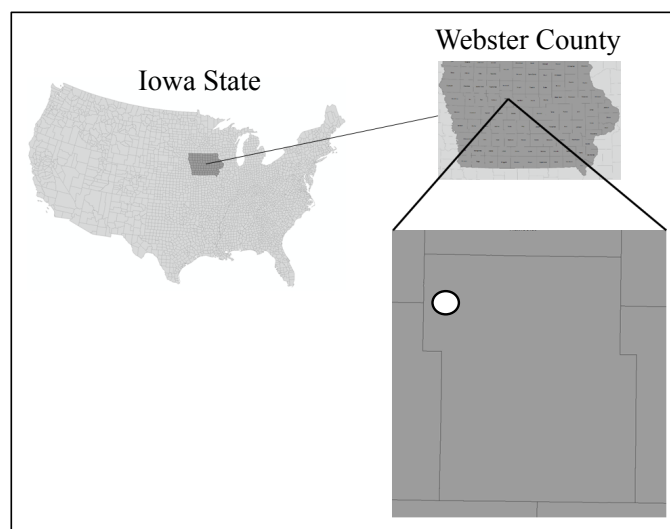
Figure 1: Study site location used in the evaluation of close-range photogrammetric surveys of ephemeral gully channels.
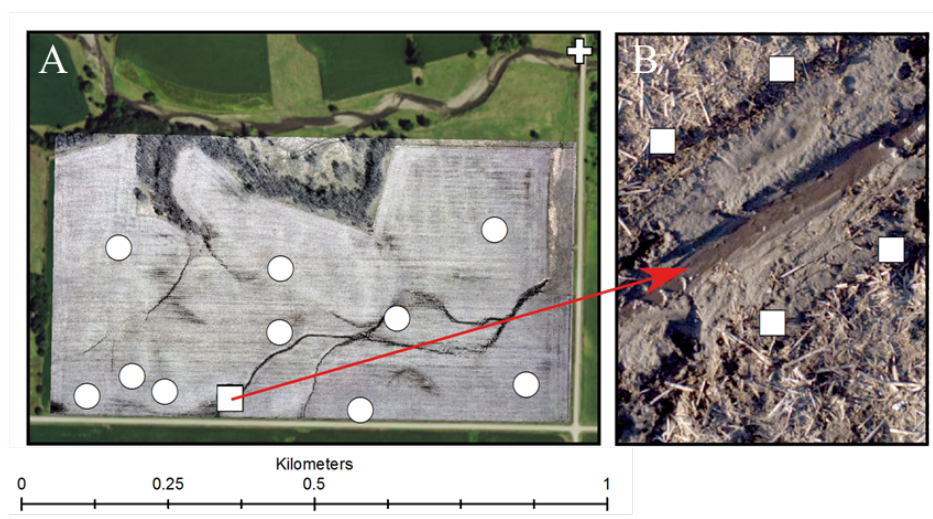
Figure 2: Study site with field ground control points (GCPs; circles), channel GCPs (square) and state monument (cross) on bridge in upper right corner (A). Selected area for detailed surveys and comprehensive evaluation (B).
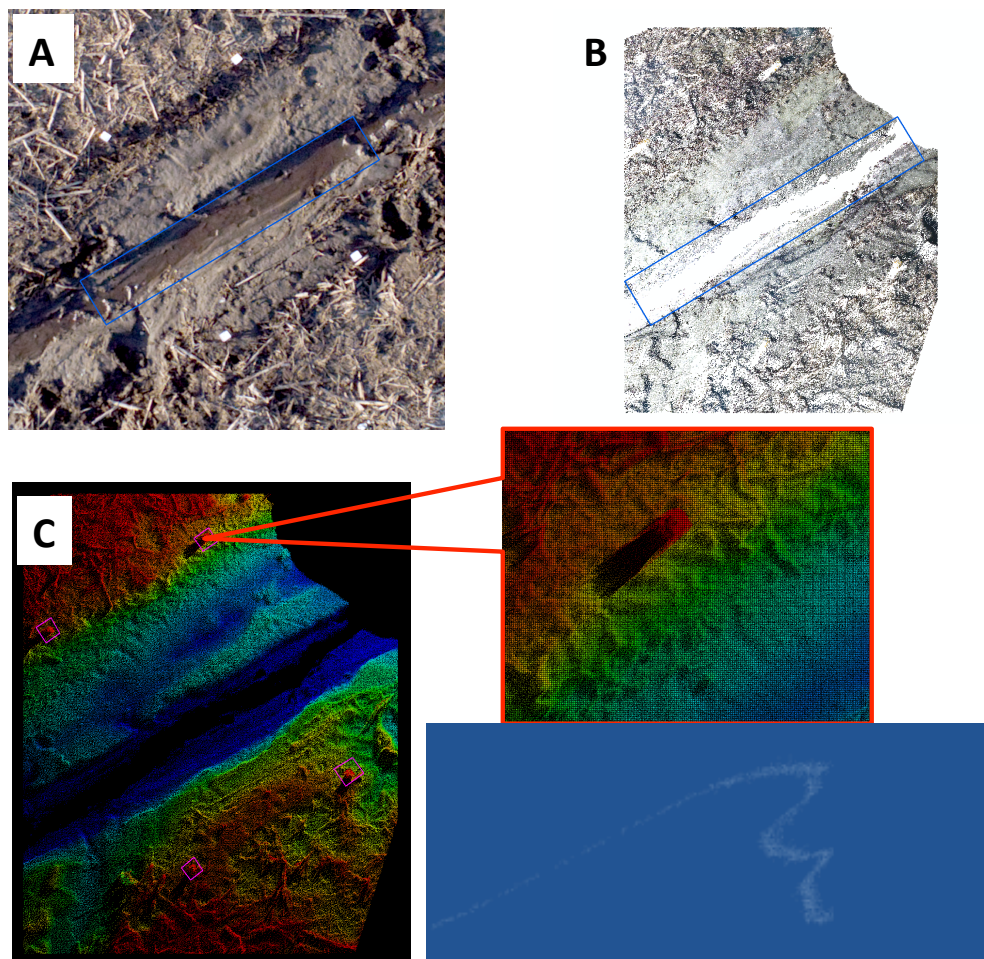
Figure 3: Limitations of ground-based LiDAR survey used as reference dataset. (A) Photograph of AOI shows water in the channel, which limits laser pulse return to the sensor causing sampling gaps in the point cloud (B), and presence of high relief features (GCPs) in the DEM (C) where sharp edges that cause the generation of multiple laser pulse returns due to split footprint effect.
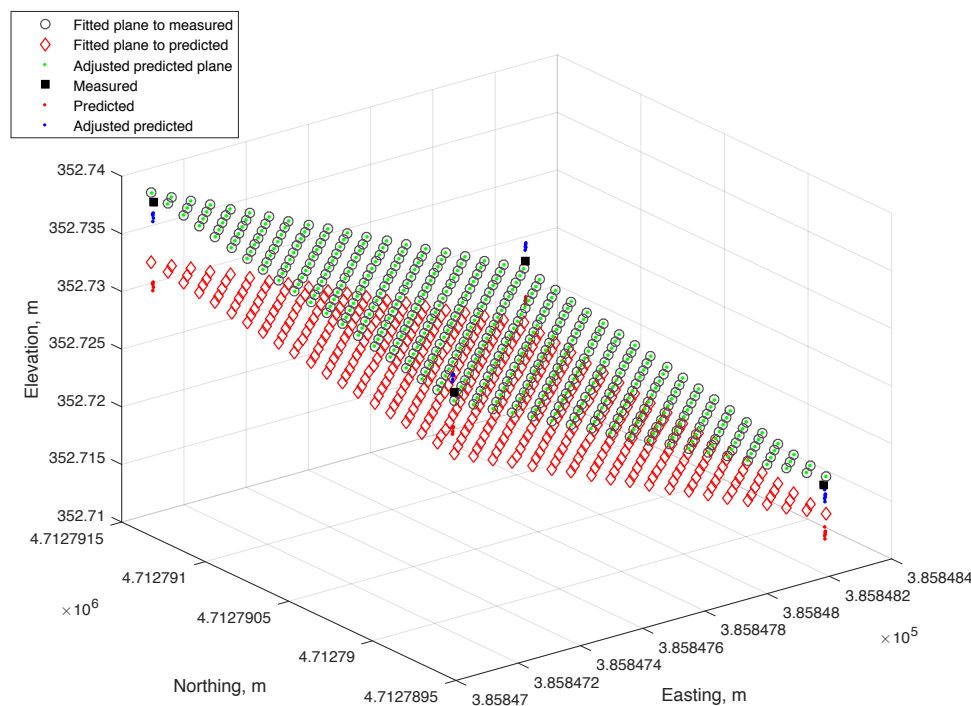
Figure 4: Determination of Afine transformation matrices using Iterative Closest Point (ICP) methodology.
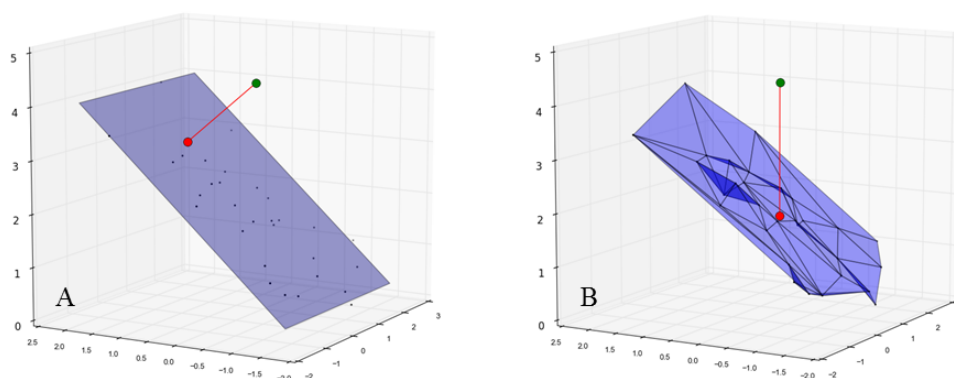
Figure 5: Schematic representation of positional accuracy analysis. Black dots represent reference dataset (LiDAR) and green circle represents point being evaluated. Red circles represent the normal projection of green point onto the tangential plane fitted to the LiDAR dataset (A) and the vertical projection of the green point into the three-dimensional LiDAR-generated triangular irregular network (TIN) (B).
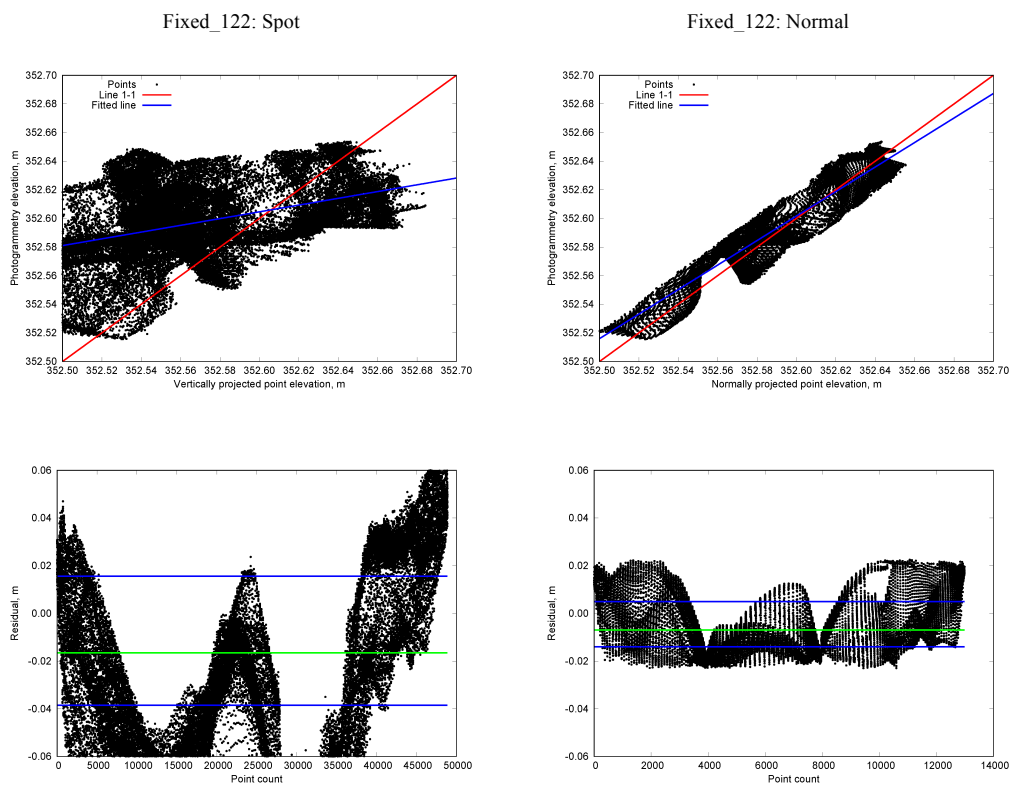
Figure 6: Fixed-wing, at 122 m flight altitude, spot elevation comparison (left column) and normal to plane comparison (right column) of photogrammetry and LiDAR point cloud data with fitted line through the elevations (blue; Spot) and 1:1 line (red; Spot), along with the residuals (green is mean residual and blue lines are 25 and 75 percentiles).
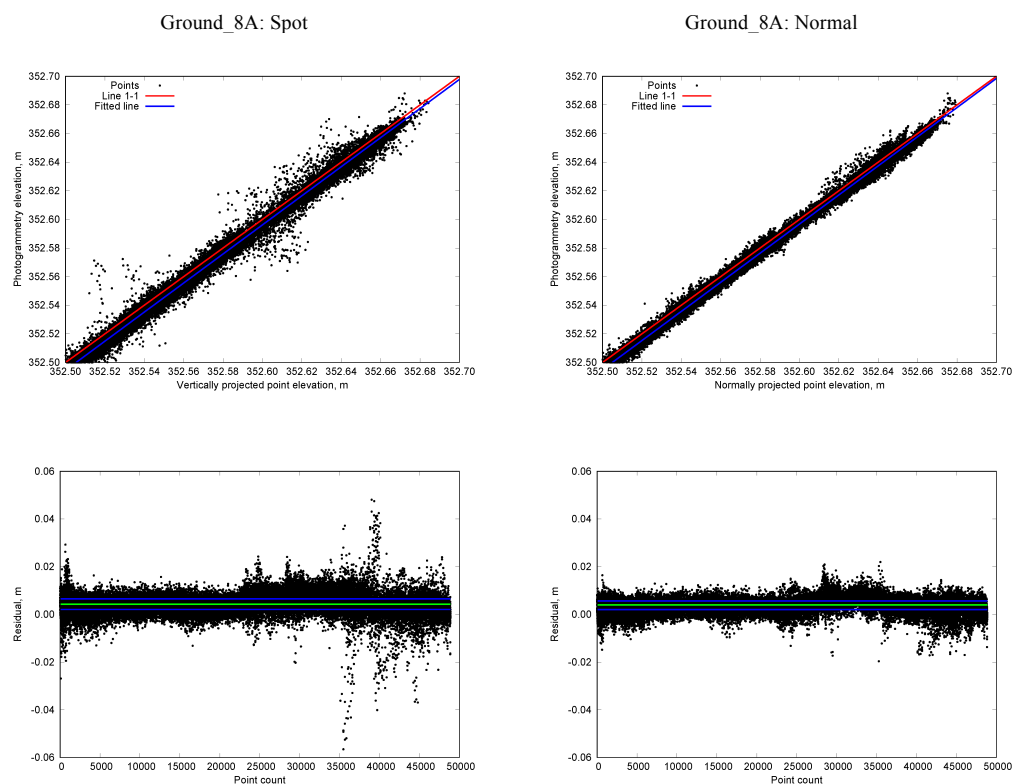
Figure 7: Four photo pair (Ground_8A) spot elevation comparison (left column) and normal to plane comparison (right column) of photogrammetry and LiDAR point cloud data with fitted line through the elevations (blue; Spot) and 1:1 line (red; Spot), along with the residuals (green is mean residual and blue lines are 25 and 75 percentiles).
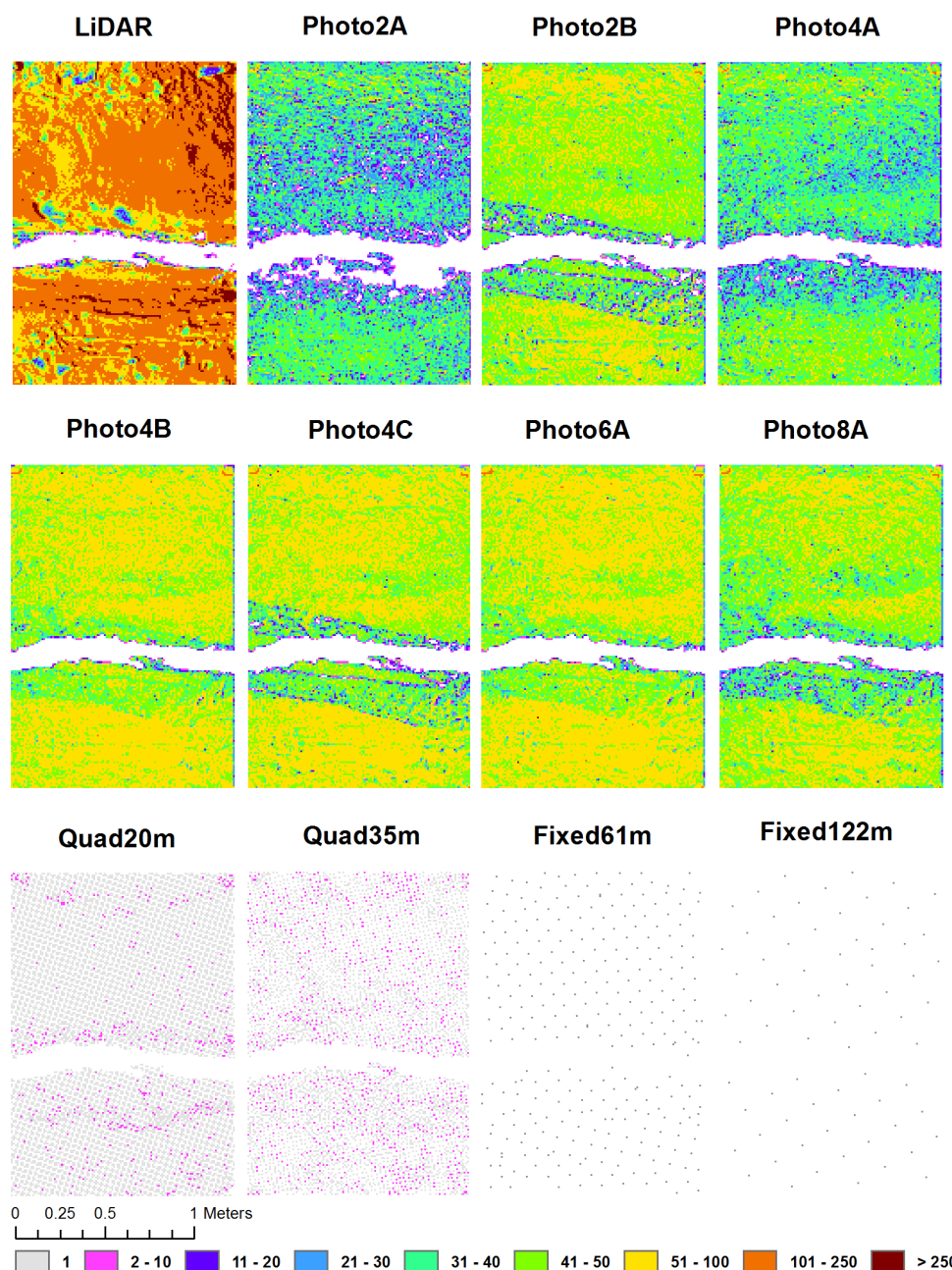
Figure 8: Sampling intensity using the quadrat method for each dataset considered. Individual color represents point sampling count intervals within a 1x1cm virtual grid. Points located in the channel were removed to match the area covered by the LiDAR dataset (herein considered as reference).
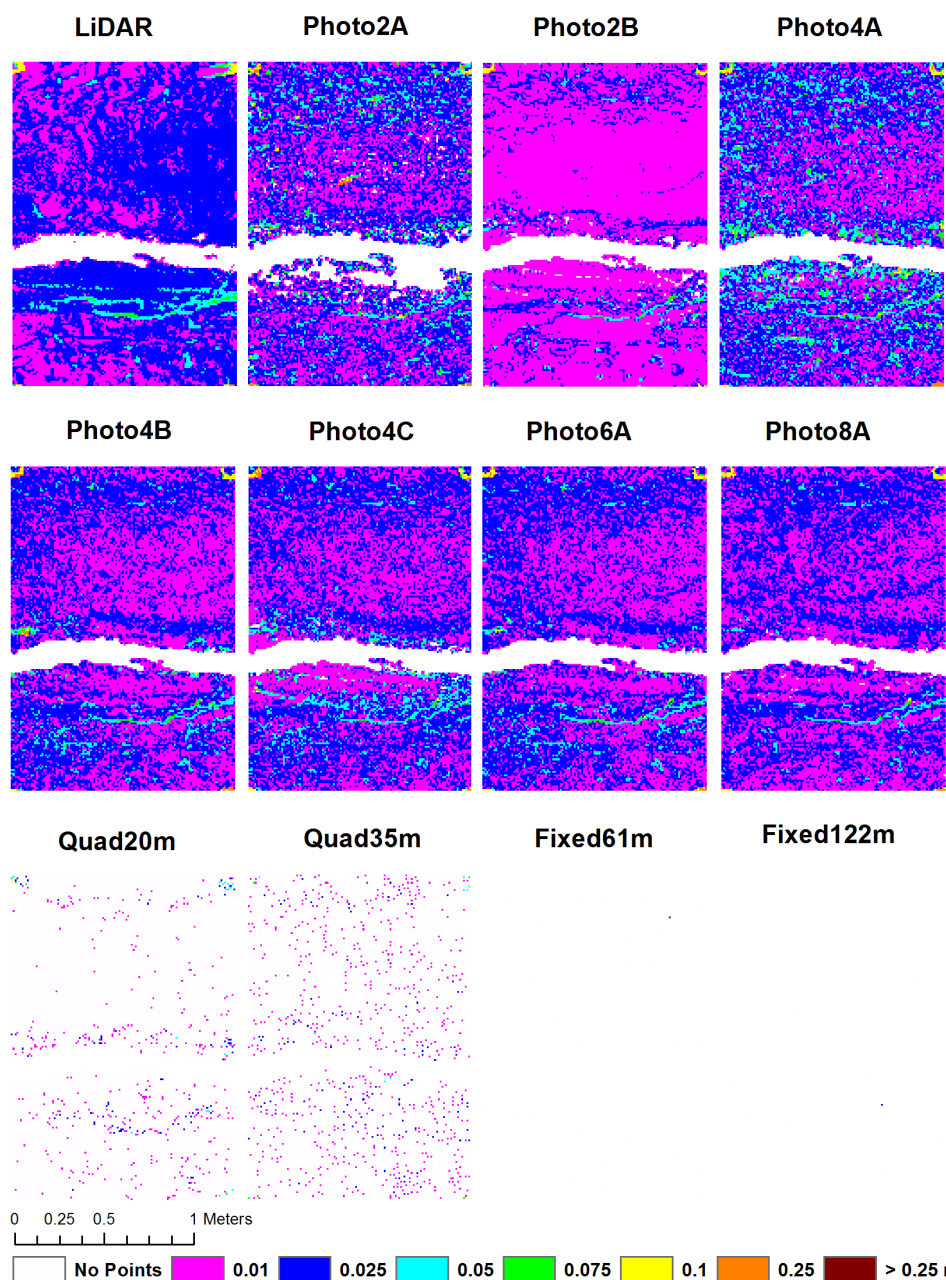
Figure 9: Elevation range (difference between minimum and maximum elevation) represented as individual colors within a 1x1cm virtual grid. Points located in the channel were removed to match the area covered by the LiDAR dataset (herein considered as reference).
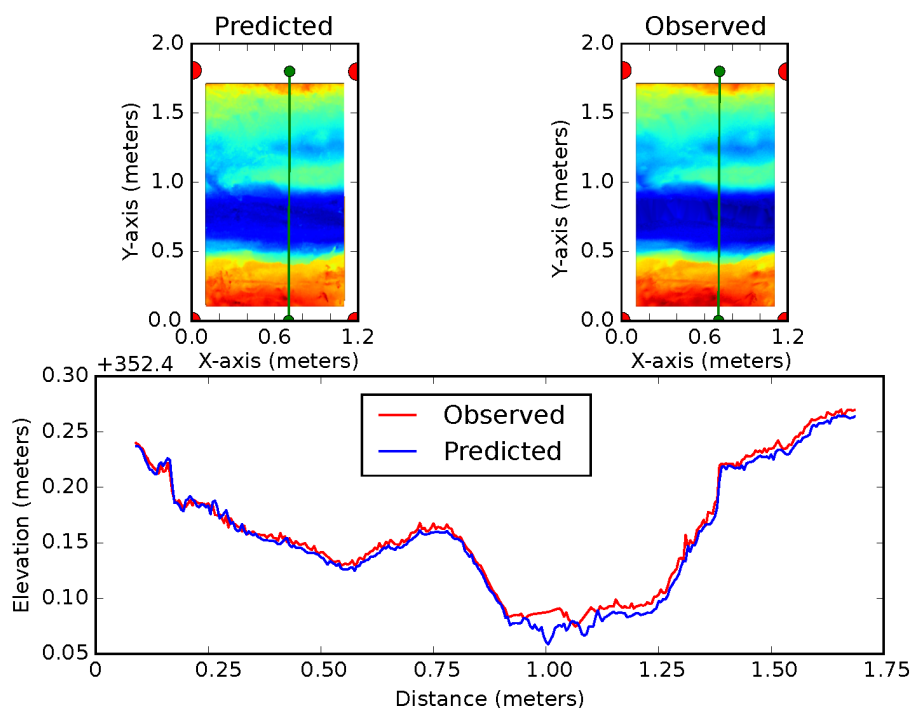
Figure 10: Illustration of interpolating three-dimensional point clouds into raster grid for volume and cross-section analysis of gully monitoring and geomorphologic quantification.
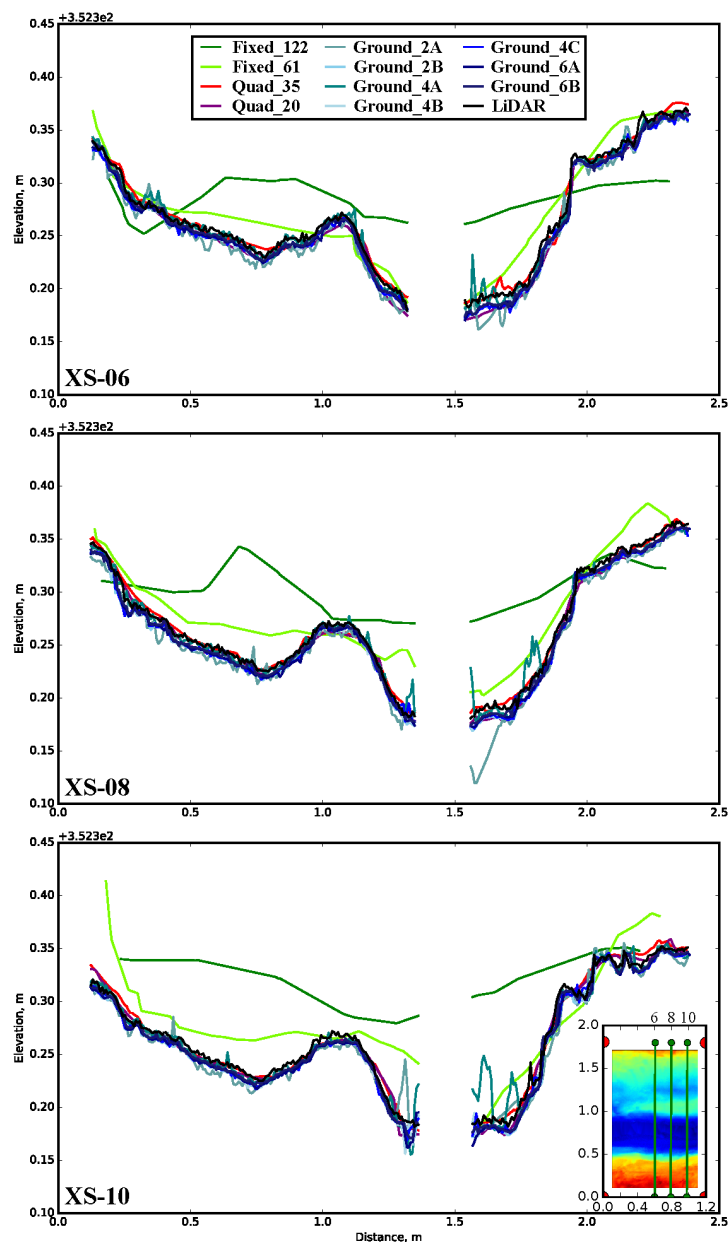
Figure 11: Selected cross-sections generated from interpolating point-clouds into 5x5 mm raster grid file.