Response to Reviewer Comments: *Reviewer #2*

Journal: ESurf
Title: Quantifying uncertainty of remotely sensed topographic surveys for ephemeral gully channel monitoring
Author(s): Robert R. Wells et al.
MS No.: esurf-2017-3
MS Type: Research article
Special Issue: 4-D reconstruction of earth surface processes: multi-temporal and multi-spatial high resolution topography

***Interactive comment on* "Quantifying uncertainty of remotely sensed topographic surveys for ephemeral gully channel monitoring" *by* Robert R. Wells et al.**
A. G6mez-Gutlerrez (Referee): alvgo@unex.es
Received and published: 21 February 2017

REVIEW OF THE PAPER: "Quantifying uncertainty of remotely sensed topographic surveys for ephemeral gully cannel monitoring" by Wells et al.
GENERAL COMMENTS: This paper presents an interesting comparison of techniques- methods used to produce high-resolution topographic surfaces from pictures and laser. These techniques are used to reproduce the surface of an ephemeral gully. I think the results could be of interest for a broad audience, as they will increase the available datasets testing several methods to produce high-resolution topography. Before that, I suggest some modifications that could be used to improve the paper before considering its publication. I would say that these are minor revisions:

1) I miss some important references in the Introduction and Discussion sections, are there other papers comparing terrestrial or aerial SfM to LIDAR-LTS? This, in my opinion, may be addressed in the introduction section. Additionally, the result of these works could be used to enrich the discussion.
We agree and have added several key references to enhance the Introduction and Discussion sections of the manuscript.
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(2): 128-138, 2006.
Casalí, J., Giménez, R., and Campo-Bescos, M.A.: Gully geometry: What are we measuring?, The Soil, 1: 509-513, 2015.
Di Stefano, C., Ferro, V., Palmeri, V., and Pampalone, V.: Testing the use of an image-based technique to measure gully erosion at Sparacia experimental area, Hydrol. Proc., 31: 573-585, doi: 10.1002/hyp.11048, 2017.
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.
Gómez-Gutiérrez, A., Schnabel, S. Berenguer-Sempere, F., Lavado-Contador, F., Rubio-Delgado, J. Using 3D photo-reconstruction methods to estimate gully headcut erosion, Catena, 120, 91-101, 2014.
Micheletti, N., Chandler, J.H., Lane, S.N.: Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smart phone, Earth Surf. Process. Landforms, 40(4): 473–486, 2015.
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, 40: 1656-1671, 2015.
Vinci, A., Brgante, R., Todisco, F., Mannocchi, F., and Radicioni, F.: Measuring rill erosion by laser scanning, Catena, 124: 97-108. doi: 10.1016/j.catena.2014.09.003, 2015.
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. Landforms, 35: 136-156, 2010.

2) Other important point that should be considered is DGPS. GCPs measured by a DGPS are ensuring, probably, a centimeter- level accuracy, so I suggest considering this point in the interpretation of the results.
We agree, however, the methods of GCP positioning and accuracy were not a primary focus of this manuscript. We report the error associated with the aligned datasets, assuming there are no ancillary errors in positioning. Here, we go to extreme lengths to ensure that the comparative datasets are in alignment and all comparisons stem from the locked channel GCP positions. We do agree that the error in GCP position should be reported and have added the error associated with static and kinematic collections for the unit we used in this study. (P4L12-16)

"All GCPs were surveyed using TopCon GR-3 dGPS survey equipment (TopCon Corporation, Tokyo, Japan; 10 mm horizontal and 15 mm vertical kinematic accuracy) to obtain relative position in reference to the state monument point. A static occupation (6 hrs; 3 mm horizontal and 5 mm vertical accuracy) was initiated with the base station, then all GCPs (field, channel and state monument) were surveyed with the rover (20 sec collection interval)."

3) In general, I consider that the text of the manuscript is well-organized but in some parts of the work there are too many sections and sub-sections, I suggest integrating some of them and this would increase the readability of the paper.
We agree and have integrated material to reduce the sub-sections. (P5L15-36)

4) I also miss some methodological details, please see my comments below. The order of figures and tables should be reviewed. References should be reviewed (see my comments below).
We have reordered the figures slightly and added flight details to address this issue. We reviewed and made corrections to the references.

SPECIFIC COMMENTS
Title: I understand that ephemeral gullies are, probably, not visible in satellite images but the terms "remotely sensed" could let to think in the use of this kind of info, so I recommend modifying the tittle to clarify this point, I suggest using "High-resolution remotely sensed".
We agree and have made the change.
"**Quantifying uncertainty of high-resolution remotely sensed topographic surveys for ephemeral gully channel monitoring**"

Abstract:
Lack of standards in landform surveying is due, at least in part, to the variability, complexity, etc. of relief.
We agree.

Line 17, UAVs are not a technique more a platform, the technique is SfM or classical photogrammetry, so I suggest modifying this part of the abstract.
We agree that UAV's are a platform for the technique of photogrammetry and do not see the confusion within these lines of text. "This work evaluated 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 and un-maned aerial vehicle-based photogrammetry platforms were compared to terrestrial LiDAR, defined herein as the reference dataset."

In general for the text, when an acronyms is written, please, the first time when you use the extended form, use capital letters.
We agree and corrected throughout the manuscript.

INTRODUCTION
L5: I guess you refer to Casali instead Casalli, by the way this reference is not in the reference list, please check.
Yes. We agree, corrected the text and added the reference.
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(2): 128-138, 2006.

L11-14: I do not completely agree, LIDAR devices are now available to be the payload of a UAV so you can get great spatial-resolutions, on the other hand you can have the desired temporal resolution for LIDAR data, you just need money to pay for that, I suggest modifying the paragraph.
We agree and provided additional information concerning UAV LiDAR. However, that element was not within the purview of this research. Money will always be an issue at the edge of science. Even with this system, you would still need the information presented herein concerning the GCPs. (P2L18-24)
"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; although, recent developments in UAV LiDAR systems provide 10mm survey-grade accuracy, one million

measurements per second, and 360° Field of View (FoV), all in <1.6 kg payloads. These UAV LiDAR systems can range from $100k to $400k (US dollar), dependent upon level of accuracy and data collection rate (see http://www.rieglusa.com as an example)."

L18: using multiple scan stations is not just for that, but more for avoiding shadows and normalizing the spatial resolution.

We agree and have added information to clarify the use of multiple scanning positions for LiDAR scene capture. (See P2L28)

"Overlapping the same area by multiple scans increases the overall sampling density and assists in occlusion and shadow avoidance, while normalizing spatial resolution."

L30: this paragraph does not flow with the rest of the introduction, you start talking about point clouds but the reader could not know why you start discussing file formats, I suggest you start talking about the typical file format recorded by TLS, LIDAR and produced by SfM and later talking about the classical use of 2.5D file formats like DEM to represent surfaces.

We agree and have added another paragraph to increase readability and help the flow. (P3L1-8)

"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 topographic information is seemingly limitless. At what point can we agree that the results describe "real" change over noise? The alignment of temporal-topographic 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 too is 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."

Equation 1: Does the ICP algorithm include the scale factor? I miss this in the explanation and formulation.

No. The ICP was used to rotate and translate the point clouds only. We added this information to the sentence. (P6L16-20)

"Three-dimensional locational differences between the reference plane generated using the GPS survey (black squares in **Error! Reference source not found.**4) and the planes of each surveyed dataset (LiDAR and photogrammetry) were calculated using the Iterative Closest Point (ICP) algorithm (Besl and Mckey, 1994; James and Robson, 2012; Micheletti et al., 2015) implemented in MATLAB (MathWorks Inc., Natick, Massachusetts) and, since no scale issues were observed, no scaling factor was implemented in the ICP."

Section 2.2.1 I miss error estimations for the DGPS, I guess 2-3 cm, would be nice to tell the readers about that.

We agree and included the information as suggested. (P4L12-16)

"All GCPs were surveyed using TopCon GR-3 dGPS survey equipment (TopCon Corporation, Tokyo, Japan; 10 mm horizontal and 15 mm vertical kinematic accuracy) to obtain relative position in reference to the state monument point. A static occupation (6 hrs; 3 mm horizontal and 5 mm vertical accuracy) was initiated with the base station, then all GCPs (field, channel and state monument) were surveyed with the rover (20 sec collection interval)."

Section 2.2.4. I suggest including some points to support the selection of different software packages to run terrestrial and aerial photogrammetry.

We do not agree. There are far too many processing packages and there is no literature that places one product over another, so there is no need to support a processing software selection. Both are well cited within the literature and both have faults similar to other existing software packages.

Section 2.2.4.1 I miss many details here: Overlaps, number of photos, UAV model, etc., etc.

Of course. All flight parameters are now included in Section 2.2.4. (P5L16-26)

"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 (Figure 2A) by throwing the craft in the air, then the craft flies, captures images and lands itself. Deployment software parameters were: Altitude (117 m, 58 m); Resolution (4.1 cm, 2 cm); 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 an 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. 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. Deployment software parameters were: Altitude (35 m, 20 m); Resolution (0.7 cm, 0.5 cm), 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 and skies were clear."

Section 2.4.1 Sampling intensity I have doubts here about the strategy used, did the authors sampled the number of points using a planimetric (XY) grid?
Yes.

In this case, they need to support this strategy. I think that estimations of volumetric point densities and 3D representations of this variable would work well in this case as. Problems using a grid in this case are important when representing vertical walls or headcut walls where points are distributed in the Z-coordinate direction, in this case, if you use a grid you will have a high value for the sampling intensity; however, if you have a look to the point cloud, in some cases, you will realize that sampling intensity can be low.
We understand the point of contingency and agree with your assessment of 3D point distributions in the Z-coordinate direction; however, this inadequacy does not preclude the point sampling strategy using a planimetric grid. We submit that overhanging walls do not exist within the study channel examined here and further submit that all point counts were carried out prior to data gridding. In this study, the water film within the channel posed a problem, so we removed this area from consideration.

Figure 10: better explanations with A)..., B)..., C)..., what kind of information was used?
We regret the confusion. In an earlier version of the manuscript, we attempted to discuss the material as predicted (photogrammetry) and observed (LiDAR), and this figure made it through without correction. In the revision, we corrected the figure by stipulating the data within the panel (i.e. (a), (b) and (c)) as suggested. The figure represents a bit of scripting we use as a visualization tool. In the revision, this Figure 10 is Figure 11.

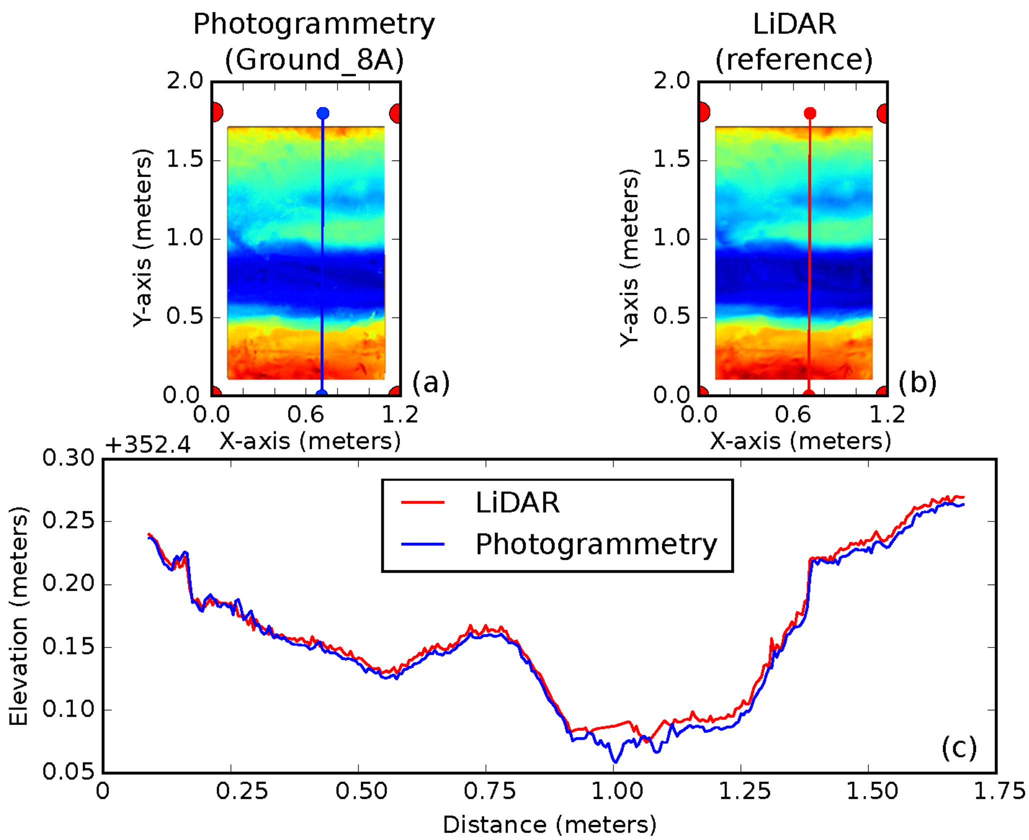

Figure 11: Illustration of three-dimensional point cloud interpolation into raster grids for volume and cross-section analysis of gully monitoring and geomorphologic quantification. Direct comparison of Ground_8A photogrammetry (a) and LiDAR (b) raster grid data with a highlight of one specific cross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

Figure 11: very difficult to understand and quite difficult to get conclusions from it. More than Fixed 122 is not working well for this Cross-section.

In the revision, this figure was moved to Figure 12. We regret the difficulty in understanding and have attempted to enrich the discussion to alleviate the confusion. First, we present these data to show that some of the platforms performed much better than others. And, we believe this point is obvious. Second, we agree with the reviewer that more than the Fixed_122 did not work well; however, we must determine the data usage before we consider whether the data is "good" or "bad". For instance, the Ground_4A data appears to have roughness elements in each of the cross-sections that place the data in the "bad" category for process investigation, while a fair amount of the metrics rank this dataset as being very close to the LiDAR. Overall, it ranks 8th. And, the Fixed_61 seems to be the other "bad" dataset; although the field scale survey and supplementary cross-section data may be appropriate for lumped model assessment (e.g. RUSLE2).

For example: (P10L33-P11L5)

"In the gridded elevation evaluations (Min, Max, Mean; Table 5; Figure 12), 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, does shoe that absolute differences from LiDAR were 131% (Fixed_122), 23% (Fixed_61), and 9% (Grround_4A). Comparing elevation information (Table 7; Figure 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 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 17 mm and the two fixed wings had a standard error of 25 mm. The average area percent difference for all cross-sections were within 1.5%, while the two fixed wings had 3% (Fixed_61) and 8% (Fixed_122). It is important to mention that the range of 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.”