Response to Reviewer Anonymous for manuscript submission to Earth Surface Dynamics,

Validation of digital elevation models (DEMs) and comparison of geomorphic metrics on the southern Central Andean Plateau (esurf-2017-4)

We thank the reviewer for the close consideration of our manuscript, and overall positive comments. Highlighted below in **bold are the reviewer comments**. Below each is our response, which will be changed in the final manuscript submission following completion of the interactive review period.

Abstract

I think the abstract is quite complex and its complexity prevents the reader from really gathering the purpose of the paper. I suggest the authors clarify better the aim of the study and organise the results presented by, for example, DEM resolution, rather than specifying the analysis for each DEM source. I believe the authors could skip the exact measurements of the errors in the abstract, in favor of a more general overview of their results. This should help improving the readibility of the abstract.

The first reviewer noted a similar issue and we have revised the abstract. We copy the revised abstract here:

“

In this study, we validate and compare elevation accuracy and geomorphic metrics of satellite-derived digital elevation models (DEMs) on the southern Central Andean Plateau. The plateau has an average elevation of 3.7 km, and is characterized by diverse topography and relief, lack of vegetation, and clear skies that create ideal conditions for remote sensing. At 30 m resolution, the SRTM-C, ASTER GDEM2, stacked ASTER L1A stereopair DEM, ALOS World 3D, and TanDEM-X have been analyzed. The higher resolution datasets include 12 m TanDEM-X, 10 m single-CoSSC TerraSAR-X / TanDEM-X DEMs, and 5 m ALOS World 3D. These DEMs represent the state-of-the-art for optical (ASTER and ALOS) and radar (SRTM-C and TanDEM-X) spaceborne sensors.

We assessed vertical accuracy by comparing standard deviations of the DEM elevation versus 307,509 differential GPS measurements across 4,000 m of elevation. For the 30 m DEMs, the ASTER datasets had the highest vertical standard deviation at > 6.5 m, whereas the SRTM-C, ALOS World 3D, and TanDEM-X were all < 3.5 m. Higher resolution DEMs had generally lower uncertainty, with both the 12 m TanDEM-X and 5 m ALOS World 3D having < 2 m vertical standard deviation. Analysis of vertical uncertainty with respect to terrain elevation, slope, and aspect revealed the low uncertainty across these attributes for the SRTM-C (30 m), TanDEM-X (12-30 m), and ALOS World 3D (5-30 m). Single-CoSSC TerraSAR-X / TanDEM-X 10 m DEMs and the 30 m ASTER GDEM2 displayed slight aspect biases, which were removed in their stacked counterparts (TanDEM-X and ASTER Stack).

Based on low vertical standard deviations and visual inspection alongside optical satellite data, we selected the 30 m SRTM-C, 12-30 m TanDEM-X, 10 m single-CoSSC TerraSAR-X / TanDEM-X, and 5 m ALOS World 3D for geomorphic metric comparison in a 66 km^2 catchment with a distinct river knickpoint. Consistent *m/n* values were found using chi plot channel profile analysis, regardless of DEM type and spatial resolution. Slope, curvature, and drainage area were calculated and plotting schemes were used to assess basin-wide differences in the hillslope-to-valley transition related to the knickpoint. While slope and hillslope length measurements vary little between datasets, curvature displays higher magnitude measurements with fining resolution. This is especially true for the optical 5 m ALOS World 3D DEM, which demonstrated high-frequency noise in 2-8 pixel steps through a Fourier frequency analysis. The improvements in accurate space-radar DEMs (e.g., TanDEM-X) for geomorphometry are promising, but airborne or terrestrial data is still necessary for meter-scale analysis.

“

Introduction

I suggest some further scientific literature to consider, that is in my opinion important to provide a complete framework for this study, and might also help in improving the discussion

when comparing this work to others. Recent challenges in geomorphometry have been shown in (Sofia et al., 2016). As well, aside from transient landscapes (Andreani and Gloaguen, 2016) and channel network analysis, new challenges in geomorphometry includes modelling anthropogenic landscapes (Tarolli, 2014; Passalacqua et al., 2015; Tarolli and Sofia, 2016). Concerning DEM errors, numerous studies provide interesting analysis, both on DEMs themselves and on the derived topographic parameters such as slope, curvature or other attributes (Albani and Klinkenberg, 2003; Albani et al., 2004; Raaflaub and Collins, 2006; Temme et al., 2006; Xuejun and Lu, 2008; Heritage et al., 2009; Fisher et al., 2013; Sofia et al., 2013)

We thank the reviewer for pointing us to these additional references. We do note that some of these references already appear in the manuscript. The studies of Tarolli (2014) (referenced on Page 3, Line 28 and 32) and Passalacqua et al. (2015) (referenced on Page 3, Line 25 and 32) are reviews that primarily deal with lidar data, which is not the focus of this study. We reference Fisher et al. (2013) when discussing the issues surrounding the noisy ASTER GDEM2 dataset (Page 32, Line 22; Page 30, Line 34). We have selected a few (but not all) of the most relevant studies listed above to add to the revised manuscript. These references, followed by our reason for adding them, are listed below:

Albani et al. (2004) discusses the effect of different window sizes on the calculation of slope, aspect, and curvature from gridded DEMs. It is shown that errors propagate less as the window size is increased, although there is also a loss of fine-scale topographic information when increasing the window size. The first reviewer pointed out that the effects of different window sizes should be briefly discussed, and we will add this reference to that discussion.

Raaflaub and Collins (2006) is another important reference to be included in the methods section (Sect. 3.4.2.), where we discuss the calculation of slope and filtering. These authors demonstrate that errors propagate into slope calculations particularly strongly when using a steepest neighbor algorithm. In our study we calculate slope via the D-infinity algorithm (http://hydrology.usu.edu/taudem/taudem5/index.html), which is not as susceptible as nearest neighbor calculations to these impacts, because slopes are divided between two cells. In our study, we also experimented with other routing and slope calculations, but did not find large differences to the D-infinity method.

Sofia et al. (2016) (from the present journal, Earth Surface Dynamics) provides a nice review and references for the current state of geomorphometry. We therefore include this reference early in the introduction.

Sofia et al. (2013) provides an analysis of curvature outliers in lidar datasets caused primarily by outliers. These errors in curvature measurement are reduced with increasing window size, following the results of Albani et al. (2004). Although this study is concerning high-resolution lidar data, we consider this reference relevant to our 5 m ALOS World 3D data, which shows curvature outliers that we find are related to error, and are reduced when resampling the data to coarser resolutions (approximately equivalent to increasing the window size of slope and curvature calculation). The reference will therefore be added accordingly in the discussion (Sect. 5.2.2.).

Methods

I wonder why the authors only considered a simple analysis based on the SD of residual, and do not consider a complete analysis of errors such as that presented for example in (Höhle and Höhle, 2009). I am also curious to see the differences in errors before filtering the outliers. A further commenting on what DEM presented the highest changes in accuracy after filtering should be done, to provide the reader with an idea of the general quality of the datasets as well.

Our study is not purely about DEM validation, but rather the geomorphometric potential of the state-of-the-art of satellite DEMs as well. Therefore, we chose to simplify our uncertainty analysis. We favor the use of a straightforward metric like the mean and standard deviation to provide a clear metric for comparing the elevation accuracy of the DEMs, particularly since the errors mostly follow normal distributions (cf. Fig. 3 and 4). This allows us to move into an analysis of the derivatives of elevation, without dwelling too long on the various ways that uncertainty can be represented. However, we have compared the full elevation distribution, but note that the mean and standard deviation capture the essence. We note that the key information are the spatial correlation or consistency of the DEM data – the comparison on a pixel-by-pixel basis is not always relevant for geomorphometric studies. This is because geomorphometric studies use the spatial content of DEMs and higher-order derivatives and not absolute elevation values. Nevertheless, we reference several studies in the discussion (Sect. 5.1.) that make fuller assessments of elevation accuracy. With this in mind, we appreciate that the inclusion of the pre-filtering results may be helpful for the interested reader, and have therefore modified Table 2, shown below. We have also modified the methods section (Sect. 3.3.) to include on Page 8, Lines 21-27:

“

Differences of ±30 m were filtered out as outliers caused by bad data and processing errors, and the percentage reduction in number of measurements from this filtering is reported along with the pre-filtering mean and SD. While many other studies suggest additional statistical tests (e.g., Höhle and Höhle, 2009), our simplified method allows us to move into further analysis of the derivatives of elevation. We have compared the full error distribution, but note that the mean and standard deviation capture the essence. The key information is the spatial correlation or consistency of the DEM data because geomorphometric studies use the spatial content of DEMs and higher-order derivatives and not absolute elevation values.

“

Table 2. Results of pixel-by-pixel DEM vertical accuracy (DEM minus dGPS). Mean and standard deviation before filtering denoted in parentheses, with value of n/a if there were no outliers filtered.

Dataset	Mean (m)	Standard Deviation (m)	Number of post-filtered rasterized measurements[a]	Reduction after ±30 m filtering (%)
30 m SRTM-C	2.18 (2.33)	3.33 (13.74)	64,782	0.02
30 m AW3D30	1.59 (1.66)	2.81 (16.19)	63,413	0.03
30 m ASTER GDEM2	-0.15 (0.02)	9.48 (17.65)	63,308	2.30
30 m ASTER Stack[b]	4.56 (4.58)	6.93[c] (7.00)	15,506	0.12
30 m TanDEM-X	-1.29 (-1.12)	2.42 (14.57)	55,791	0.02
12 m TanDEM-X	-1.41 (-1.31)	1.97 (11.16)	108,029	0.02
10 m CoSSC TDX (7 February 2011)	1.99 (2.36)	2.02 (21.26)	28,982	0.03
10 m CoSSC TDX (6 November 2012)[d]	1.32 (n/a)	3.83 (n/a)	22,182	0.00
10 m CoSSC TDX (25 August 2013)	2.94 (n/a)	3.22 (n/a)	22,175	0.00
5 m AW3D5	2.40 (n/a)	1.64 (n/a)	14,306	0.00

a, After ±30 m outlier filtering
b, Generated for Pocitos Basin by weighted stacking of eight manually generated ASTER L1A DEMs
c, Compare with 11.42 m and 10.06 m SD for single L1A DEM and ASTER GDEM2, respectively, clipped to same area
d, CoSSC TDX DEM selected for geomorphometric analysis

Results

In some instances, I found a bit of confusion between grid resolution increasing/ grid size decreasing, please double check on this.

The first reviewer also noted this, and we have made sure to change the terminology accordingly.

Sincerely,

For both authors,

Ben Purinton

Universitaet Potsdam, Germany

purinton@uni-potsdam.de

References

Albani M, Klinkenberg B. 2003. A Spatial Filter for the Removal of Striping Artifacts in Digital Elevation Models. Photogrammetric Engineering Remote Sensing 69: 755–765

Albani M, Klinkenberg B, Andison DW, Kimmins JP. 2004. The choice of window size in approximating topographic surfaces from Digital Elevation Models. International Journal of Geographical Information Science 18 (6): 577–593 DOI: 10.1080/13658810410001701987

Andreani L, Gloaguen R. 2016. Geomorphic analysis of transient landscapes in the Sierra Madre de Chiapas and Maya Mountains (northern Central America): implications for the North American–Caribbean–Cocos plate boundary. Earth Surface Dynamics 4 (1): 71–102 DOI:10.5194/esurf-4-71-2016

Fisher GB, Bookhagen B, Amos CB. 2013. Channel planform geometry and slopes from freely available high-spatial resolution imagery and DEM fusion: Implications for channel width scalings, erosion proxies, and fluvial signatures in tectonically active landscapes. Geomorphology 194: 46–56 DOI:10.1016/j.geomorph.2013.04.011

Heritage GL, Milan DJ, Large ARG, Fuller IC. 2009. Influence of survey strategy and interpolation model on DEM quality. Geomorphology 112 (3–4): 334–344 DOI: 10.1016/j.geomorph.2009.06.024

Höhle J, Höhle M. 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS Journal of Photogrammetry and Remote Sensing 64 (4): 398–406 DOI: 10.1016/j.isprsjprs.2009.02.003

Passalacqua P, Belmont P, Staley DM, Simley JD, Arrowsmith JR, Bode CA, Crosby C, DeLong SB, Glenn NF, Kelly SA, et al. 2015. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. Earth-Science Reviews 148: 174–193 DOI: 10.1016/j.earscirev.2015.05.012

Raaflaub LD, Collins MJ. 2006. The effect of error in gridded digital elevation models on the estimation of topographic parameters. Environmental Modelling Software 21 (5): 710–732 DOI: http://dx.doi.org/10.1016/j.envsoft.2005.02.003

Sofia G, Hillier JK, Conway SJ. 2016. Frontiers in Geomorphometry and Earth Surface Dynamics: Possibilities, Limitations and Perspectives. Earth Surface Dynamics 4: 721–725 DOI: 10.5194/esurf-4-721-2016

Sofia G, Pirotti F, Tarolli P. 2013. Variations in multiscale curvature distribution and signatures of LiDAR DTM errors. Earth Surface Processes and Landforms 38(10): 1116–1134 DOI: 10.1002/esp.3363

Tarolli P. 2014. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. Geomorphology 216: 295–312

Tarolli P, Sofia G. 2016. Human topographic signatures and derived geomorphic processes across landscapes. Geomorphology 255: 140–161 DOI:10.1016/j.geomorph.2015.12.007

Temme AJAM, Schoorl JM, Veldkamp A. 2006. Algorithm for dealing with depressions in dynamic landscape evolution models. Computers Geosciences 32 (4): 452–461 DOI: http://dx.doi.org/10.1016/j.cageo.2005.08.001

Xuejun L, Lu B. 2008. Accuracy Assessment of DEM Slope Algorithms Related to Spatial Autocorrelation of DEM Errors. In Advances in Digital Terrain Analysis SE - 16, Zhou Q, , Lees B, , Tang G (eds).Springer Berlin Heidelberg; 307–322. DOI: 10.1007/978-3-540-77800-416