DOI: 10.24850/j-tyca-2024-04-05

Articles

Evaluation of the Digital Elevation Model from the Shuttle Radar Topography Mission (SRTM) on the Papaloapan Macro-Basin, Mexico, using LiDAR as benchmark

Evaluación del modelo de elevación digital de la misión topográfica de radar en transbordador (SRTM) en la macrocuenca del Papaloapan, Mexico, usando LiDAR como referencia

Edgar Misael Uribe[1], ORCID: https://orcid.org/0000-0001-5993-4244

José Cruz Escamilla[2], ORCID: https://orcid.org/0000-0001-7059-7752

Abigail Juárez[3], ORCID: https://orcid.org/0000-0002-0772-5667

[1]Earth and Materials Sciences Academic Area, Universidad Autónoma del Estado de Hidalgo, Hidalgo, México, edgar_uribe8920@uaeh.edu.mx

[2]Earth and Materials Sciences Academic Area, Universidad Autónoma del Estado de Hidalgo, Hidalgo, México, jocesca@uaeh.edu.mx

[3]Environmental Geology, Academic Area of Earth and Materials Sciences, Universidad Autónoma del Estado de Hidalgo, Hidalgo, México, ju354755@uaeh.edu.mx

Corresponding author: Edgar Misael Uribe-Alcántara, edgar_uribe8920@uaeh.edu.mx

Abstract

The Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) is evaluated using a LiDAR DEM from INEGI as benchmark in the Papaloapan Basin (~58 000 km^2) in Mexico. Three representative regions are selected: 1) a hilly region with strong slopes and elevations over 3 000 m; 2) a transitional region with relatively smoother slopes and elevations around 1 000 m, and 3) a floodplain with flat terrain and elevations below 100 m. The straight comparison of both datasets shows very similar elevation values at the hilly and transitional regions. However, in the floodplain, the relationship has a parabolic shape, and errors are relatively higher, in terms of the elevation range. This is probably due to systematic errors in SRTM being very close to the actual low elevations. Maps of errors suggest strong association with water bodies and the aspect. For example, in the transitional region, most negative errors are found on slopes facing east, while positive errors are found on slopes facing west. Three-dimensional histograms of errors *vs*. topographic features (elevation, slope, and aspect) are estimated. The histograms suggest a systematic error, which means SRTM could be improved with a simple calibration at least in these cases. Evaluations of public DEMs from different sources in Mexico are considered necessary for the identification of their strengths and weaknesses. We believe these evaluations might provide the grounds for the creation of improved MEDs in the future either by either a simple calibration or through composite MEDs from multiple sources.

Keywords: SRTM, Digital Elevation Model, DEM, LiDAR, Papaloapan Basin, Evaluation of INEGI products, Composite DEM, multi-source DEM.

Resumen

Se evalúa el modelo de elevación digital (MED) del Shuttle Radar Topography Mission (SRTM) empleando datos del MED LiDAR de INEGI como referencia en la cuenca del Papaloapan (~58 000 km^2) en México. Se seleccionaron tres regiones representativas: 1) una región montañosa con pendientes fuertes y elevaciones superiores a los 3 000 m; 2) una región transicional con pendientes relativamente más suaves y elevaciones alrededor de 1 000 m, y 3) una planicie de inundación con terreno plano y elevaciones menores a los 100 m. La comparación directa entre ambos MED muestra valores de elevación muy similares en las regiones montañosa y transicional. Sin embargo, en la planicie de inundación, la regresión muestra una forma parabólica, y los errores son relativamente más altos, en términos del rango de elevación. Esto probablemente se debe a errores sistemáticos en SRTM muy cercanos a las elevaciones bajas. Los mapas de errores sugieren una fuerte asociación con cuerpos de agua y el aspecto. Por ejemplo, en la región transicional, la mayoría de los errores negativos se encuentran en pendientes orientadas al este, mientras que la mayoría de errores positivos están en pendientes orientadas al oeste. Se estimaron histogramas tridimensionales de errores *vs*. rasgos topográficos (elevación, pendiente y aspecto). Los histogramas sugieren un error sistemático, lo cual implica que el SRTM podría mejorar con una calibración simple al menos en los presentes casos. Las evaluaciones de MED públicos de diferentes fuentes en México se consideran necesarias para identificar sus fortalezas y debilidades. Estas evaluaciones podrían constituir la base para la creación de MED mejorados en el futuro, ya sea mediante simple calibración o mediante MED compuestos provenientes de fuentes múltiples.

Palabras clave: SRTM, modelo elevación digital, LiDAR, MED, cuenca del Papaloapan, evaluación de productos de INEGI, MED compuesto, MED multifuente.

Received: 23/06/2022

Accepted: 22/01/2023

Published Online: 27/01/2023

Introduction

Digital Elevation Models (DEMs) are becoming an increasingly necessary resource for many environmental studies, especially related to hydraulics and hydrology, because topography is a key factor in determining water distribution and circulation. In Mexico, the official national DEM is the Mexican Continuum of Elevation (Continuo de Elevación Mexicano, CEM 3.0) published by the National Institute of Statistics and Geography (INEGI, 2017). Unfortunately, Uribe-Alcantara, Escamilla-Casas y Cruz-Chavez (2018) have showed this DEM has extremely high errors (particularly in floodplains) associated with an artificial treatment of water bodies by INEGI, which unfortunately has not been documented in the official literature of this product. Therefore, users of public DEMs in Mexico are in need of finding a solution to this lack of accuracy issue in the official public DEM for Mexico, *i.e.* the CEM.

Two possible solutions are: 1) using alternative public DEMs, and 2) creating multi-source DEMs. Regarding the first solution, although there are a few public DEMs available in Mexico, they offer different and sometimes complementary advantages and disadvantages. For example, the Shuttle Radar Topography Mission (**SRTM)** has the advantage of being available nationwide but its spatial resolution is relatively low (pixel size of 90 m). On the other hand, INEGI has also published a LiDAR DEM for Mexico (INEGI, 2017). This dataset has an excellent spatial resolution (5 m) but unfortunately it is not available nationwide. However, where available, LiDAR is considered an appropriate benchmark because of its higher resolution and accuracy associated with the LiDAR technology and closeness to the earth, compared to other remote sensing techniques like the one used by SRTM. There are certainly more accurate methods and technologies like drones and topographic surveys, however, their products are not suitable to evaluate nationwide products, such as SRTM, because they are not public, and they are not available for large extensions.

Ideally, it would be very convenient if we could combine the strengths of each DEM to create the best DEM possible. There are a couple of studies on the possible combination of multi-source DEMs to create a single DEM (Baghdadi *et al*., 2005; Gesch & Wilson, 2001). Eventually, the authors of this paper want to explore the creation of multi-source DEMs to create a DEM for Mexico that combines the strengths of each DEM. However, in order to do so, the evaluation of individual strengths and weaknesses of different DEMs is necessary. A formal assessment is expected to eventually provide some guidance on how to best combine public DEMs available in Mexico. An evaluation of the CEM 3.0 has already been performed (Uribe-Alcantara *et al.*, 2018). In this paper, we evaluated SRTM using LiDAR as a benchmark to identify its strengths and weaknesses, as a function of elevation and other topographic features.

Data and methodology

Two DEMs are compared: SRTM and LiDAR. The SRTM's project is a joint mission between the U.S. National Geospatial-Intelligence Agency (NGA) and the U.S. National Aeronautics and Space Administration (NASA). The objective of this project was to create a DEM for the region between parallels 56° N and 56° S. We used SRTM version 4, distributed by the Consultative Group on International Agricultural Research-Consortium for Spatial Information (CGIAR- CSI) (Jarvis, Reuter, Nelson, & Guevara, 2008). This version has processed data voids, and its resolution is 90 m. The biggest advantage of SRTM is perhaps associated with its availability. This dataset is publicly available over most of the world. The biggest disadvantage is perhaps its low resolution.

On the other hand, LiDAR from INEGI is used as a benchmark. This dataset has 5 m pixel size. This resolution is very high so handling such a large dataset may be inconvenient, particularly for meso- and macro

Tecnología y ciencias del agua, ISSN 2007-2422, 15(4), 194-219. DOI: 10.24850/j-tyca-2024-04-05

products. Therefore, the resolution plays both as an advantage, in terms of accuracy and precision, but also as a disadvantage in terms of processing requirements. The biggest LiDAR's disadvantage is the fact that the dataset is not available all over Mexico. The availability is patchy and, unfortunately, there is no practical way to learn the country's coverage because INEGI´s documentation and metadata are extremely poor (INEGI, 2017).

Since LiDAR availability is limited and it has a very high resolution, evaluating SRTM all over Mexico is neither feasible nor practical. Instead, the evaluation was performed in the Papalopan basin. This macro basin (57 716 km^2) was selected because its topographic features are representative of the elevation range in Mexico (sea level to 5 610 m.a.s.l.). This basin is still quite large for a complete analysis. Furthermore, LiDAR is not available throughout the basin. Therefore, three representative regions were selected for the evaluation:

1. Hilly region: Elevations over 3 000 m with strong slopes and intense spatial variability.
2. Transitional region: Elevations around 1 000 m with mild slopes and moderate spatial variability.
3. Floodplain: Elevations below 100 m with flat slopes and smooth spatial variability.

We consider that these regions are representative of large basins in Mexico, which usually start in hilly regions at very high elevations (thousands of meters) with strong slopes and hilly terrain; then the stream network flows into middle elevations (around 1 000 m), where slopes and spatial variability are both relatively smoother; and streams finally reach floodplains with very low elevations, flat terrain, and smooth variability. In addition to the Papaloapan basin, the Grijalva-Usumacinta basin is also a good example of this pattern found in large Mexican basins.

SRTM data was downloaded from CGIAR's website (Jarvis, Reuter, & Nelson, 2014), while LiDAR data was obtained from INEGI (2017). Both datasets are distributed in mosaics. As mentioned earlier, LiDAR data has limited availability so it was necessary to identify LiDAR data for each representative region. We were able to identify 12, 24 and 16 tiles for the hilly, transitional and floodplain regions, correspondingly. These tiles were merged into a single DEM for each region. Figure 1 shows the location of each region within the Papaloapan basin. Figure 2, Figure 3, and Figure 4 show LiDAR for each one of the regions: Hilly, transitional, and floodplain, correspondingly.

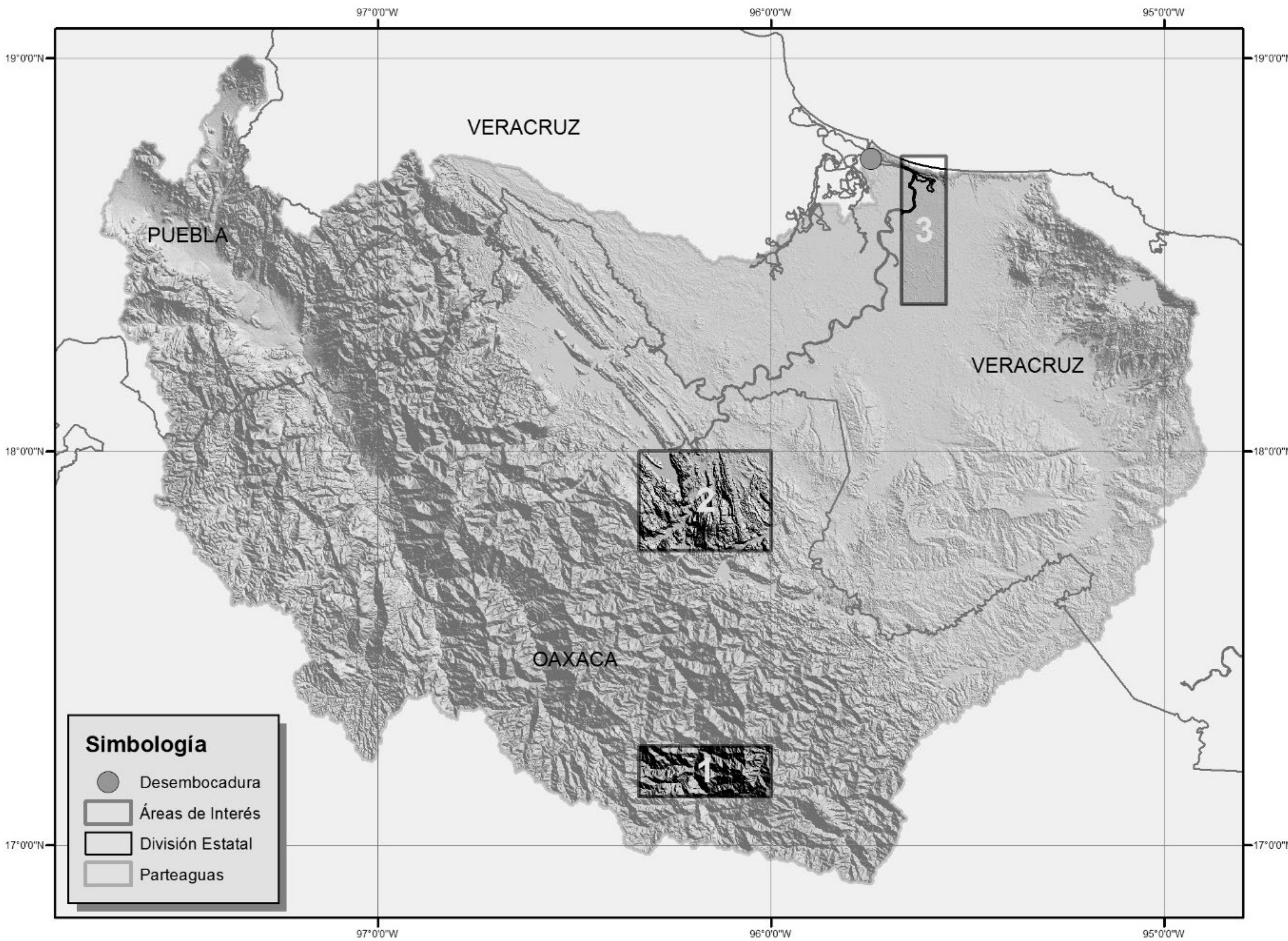

Figure 1. Papaloapan basin and its outlet (circle). Analyzed regions: 1) hilly, 2) transitional, and 3) floodplain (taken from Uribe-Alcantara *et al*. (2018)).

Tecnología y ciencias del agua, ISSN 2007-2422, 15(4), 194-219. DOI: 10.24850/j-tyca-2024-04-05

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

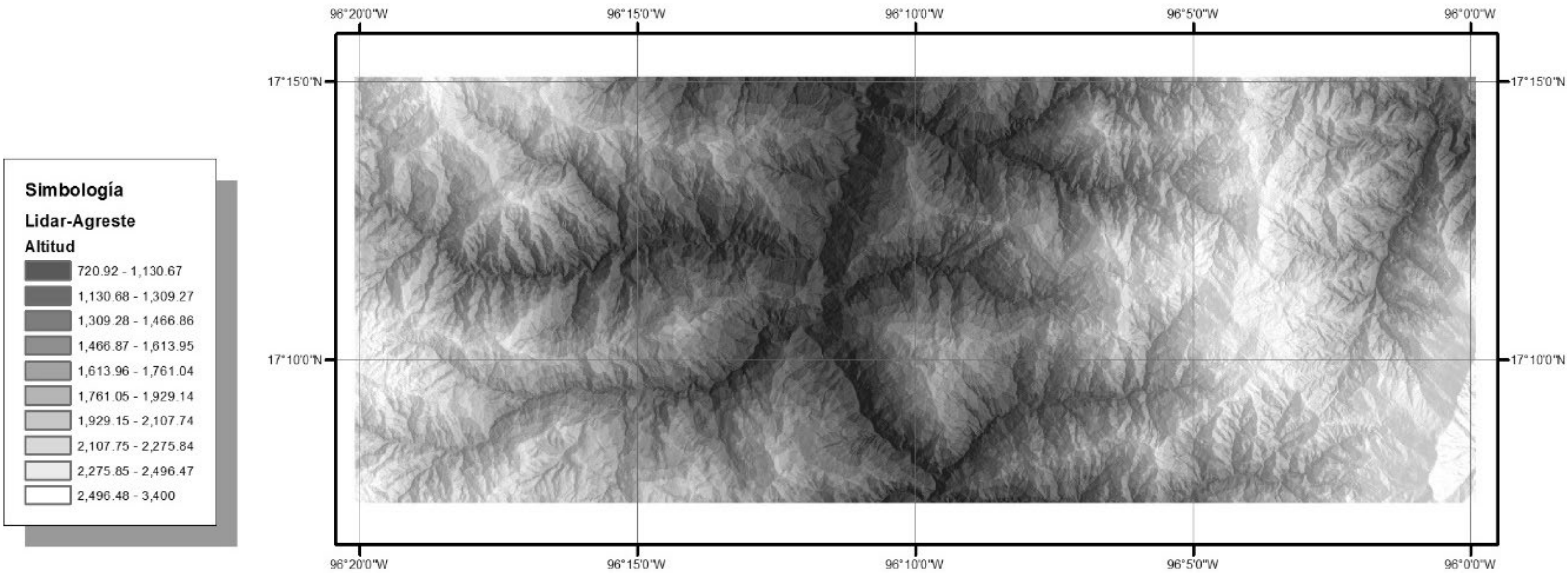

Figure 2. LiDAR's Digital Elevation Model for the hilly region.

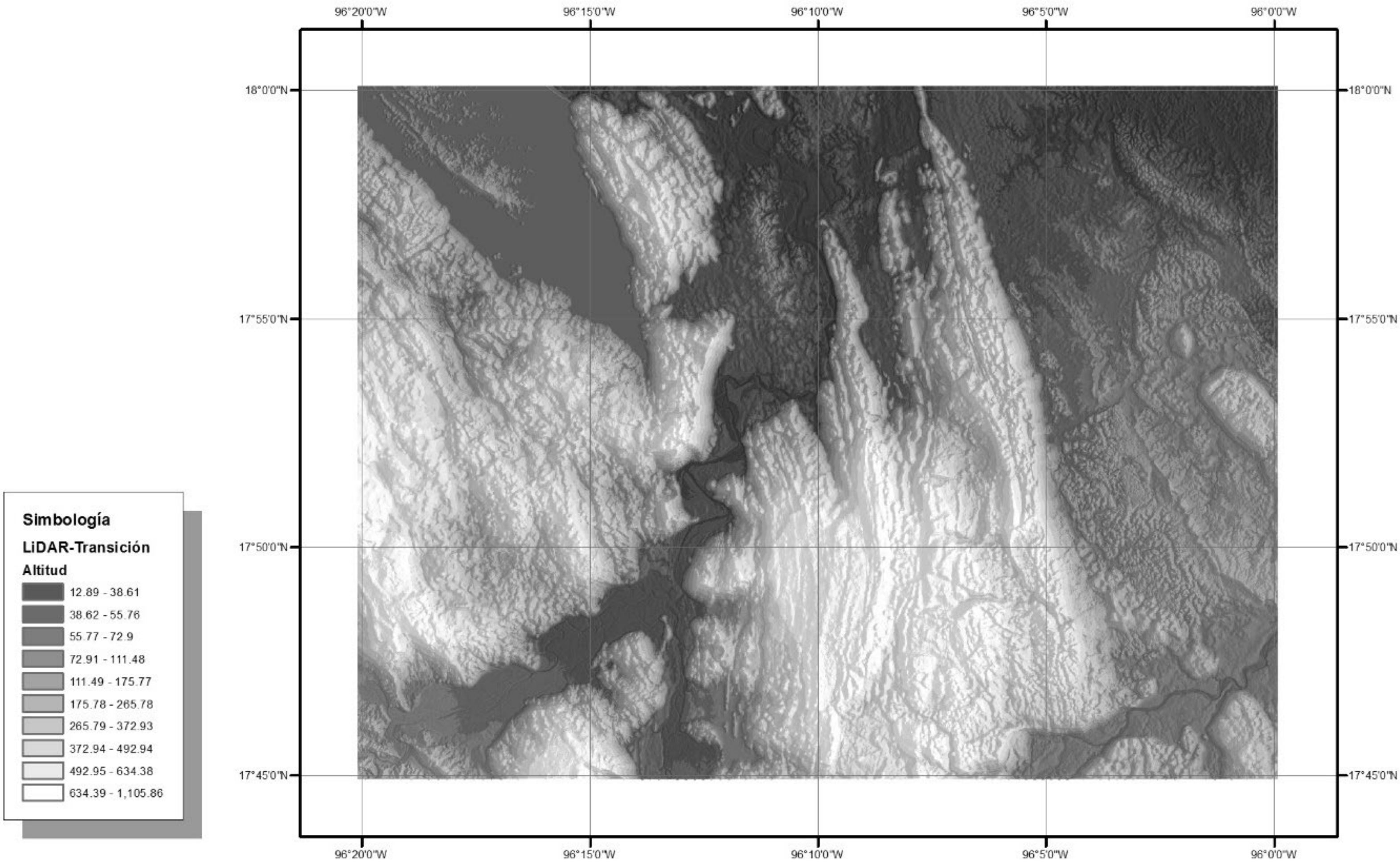

Figure 3. LiDAR's Digital Elevation Model for the transitional region.

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

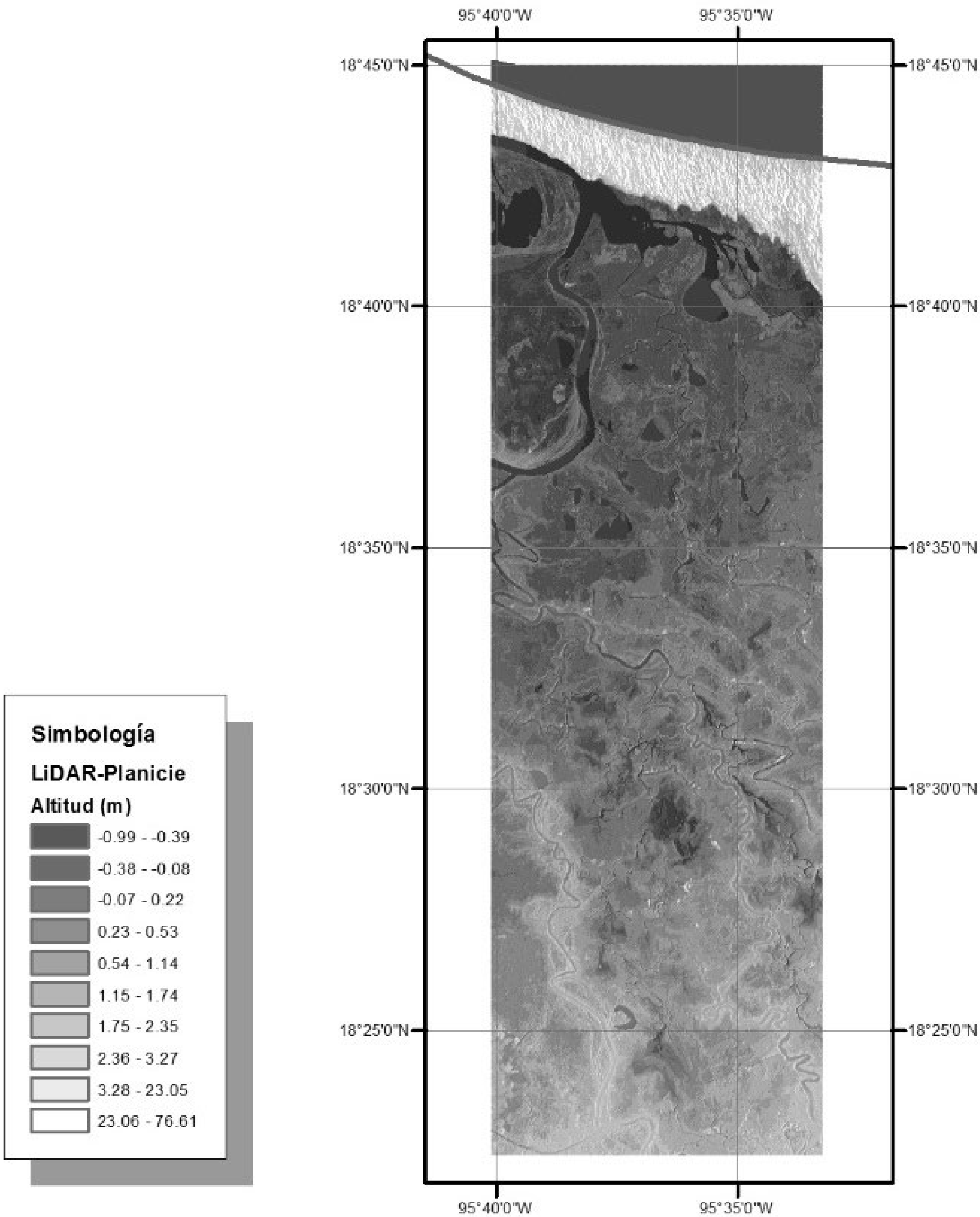

Figure 4. LiDAR’s Digital Elevation Model for the floodplain.

Once the domains for each region were defined based on LiDAR-DEMs, the corresponding SRTM-DEMs were extracted from the original CGIAR data. The evaluation consisted simply on the comparison between SRTM and LiDAR, which was used as benchmark. Thus, errors were calculated using the following equation:

$$Error = Z_{SRTM} - Z_{LiDAR} \quad (1)$$

where:

Z_{SRTM} = elevation in SRTM

Z_{LiDAR} = elevation in LiDAR

Therefore, if the errors are positive, SRTM is overestimating elevation, but if the errors are negative, SRTM is underestimating elevation. However, in order to compare both datasets, the pixels must be consistent, *i.e.* they should have the same grid framework. To achieve consistent grids, LiDAR (5 m pixel size) was upscaled to reach SRTM's pixel-size (90 m pixel size). The procedure is the following: 1) LiDAR was projected from UTM (Zone 15, Datum WGS84) to geographic coordinates to match SRTM's projection, and 2) the resulting grid was aggregated to a 90 m pixel-size, using SRTM's grid as a template for the resulting calculation. This aggregation procedure ensures that all LiDAR pixels falling inside each SRTM's pixel, are averaged and assigned to a grid with the same SRTM's grid framework so a straightforward comparison between pixels from both DEMs is feasible.

Finally, since several papers have pointed out that errors in DEMs from remote sensors may be associated with aspect, slope and elevation (Bater & Coops, 2009; Goulden, Hopkinson, Jamieson, & Sterling, 2016; Uribe-Alcantara *et al*., 2018); these topographic features were calculated, and plotted along with errors, using three-dimensional histograms to explore the relationship between topographic features and errors. Also, as mentioned earlier, there is an evaluation of the CEM 3.0 by Uribe-Alcantara *et al*. (2018), where the authors identified large errors in the same floodplain. Thus, a comparison between the CEM 3.0 and SRTM was feasible, and considered pertinent to evaluate if SRTM presents the same problems than the CEM 3.0, and also to confirm the occurrence of an artificial modification of the elevation values.

Results

Once both DEMs shared the same grid framework, we proceeded to analyze differences between elevations. Figure 5 shows scattergrams of LiDAR *versus* SRTM elevations, as well as the best polynomial fit for each region. Table 1 shows polynomial fit coefficients and norm of residuals. The hilly and transitional regions show very close linear relationships between both DEMs (slopes are very close to one). Both regions also share a similar constant (*i.e.* intercept). Both the linear relationship and the constant intercept suggest a simple linear systematic calibration could conveniently improve SRTM in these regions. On the other hand, the transitional region shows a stronger norm of residuals than the hilly region, which suggests errors are relatively larger. The floodplain shows

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

a different pattern. Errors resemble a parabolic behavior with a much larger relative standard deviation. The best fit for this region was achieved with a cubic regression. The norm of residuals are much lower than in the other two regions because the range of elevations is much smaller.

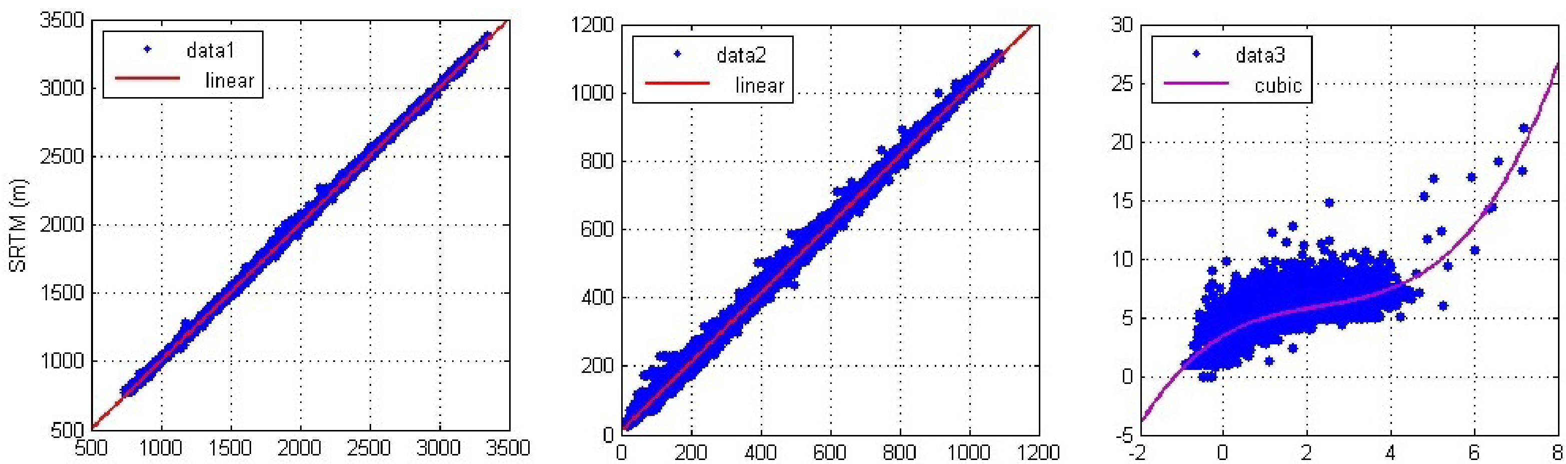

Figure 5. Scattergrams of LiDAR (horizontal axis) *versus* SRTM elevations (vertical axis) for each one of the regions: Hilly region (left), transitional region (center), and floodplain (right).

Table 1. Polynomial fit coefficients and norm of residuals for each region. For the floodplain, two fits were calculated (linear and cubic).

Region	**Polynomial coefficients**				**Norm of residuals**	**Normalized norm of residuals**
	X^3	**X^2**	**X^1**	**X^0**		
Hilly			0.9991	12.8650	1056.00	0.3985
Transitional			1.0055	13.2520	1372.00	1.2472
Floodplain			1.0789	3.6555	86.24	4.1067
Flodplain	0.0931	-0.6423	2.0765	3.4898	82.32	

However, the norms of residuals need to be normalized to allow for comparisons between errors in these three regions with very different ranges of values. Table 1 shows normalized norms of residuals by elevation range. Thus, we can conclude that the floodplain shows the strongest normalized errors, perhaps because elevations at these coastal regions are close to the error associated with remote sensors. The region with the second largest errors is the transitional; and the third largest errors correspond to the hilly region.

Table 2 shows the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and Pearson Correlation Coefficient (PCC) for each region. The largest errors are associated with the transitional region; the second largest, with the hilly region; and the lowest, with the floodplain. However, if we take into account the elevation range, the normalized highest errors are once again associated with the floodplain.

Table 2. Mean Absolute Errors (MAE), Root Mean Square Errors (RMSE), and Pearson Correlation Coefficient (PCC) for each of the three regions.

Region	MAE (m)	RMSE (m)	PCC
Hilly	14.15	17.71	0.9996
Transition	15.53	19.24	0.9987
Floodplain	3.74	3.93	0.6857

As discussed earlier, studies suggest errors in aerial and satellite DEMs are associated with slope, water bodies, or even the angle between the remote sensor and the surface. In order to explore the spatial distribution of errors, maps are calculated for each region. Figure 6 shows errors for the hilly region. The largest positive errors are associated with the stream network, but there is no other obvious spatial pattern. The error distribution seems symmetrical, with the largest negative error at -118 m, and the largest positive error at 132 m. Figure 7 shows errors in the transitional region. In this case, errors seem to be clearly associated with terrain aspect. Slopes facing east have negative errors, while the largest positive errors seem associated with slopes facing west. In this case, water bodies show positive errors (*i.e.* Miguel de la Madrid Dam). Finally, Figure 8 shows errors in the floodplain, as well as in the water bodies reported by INEGI (scale 1:50 000). The figure also shows two cross-sections, one for errors (top plot), and a second one for elevation (bottom plot). These sections correspond to the same sections reported by Uribe-Alcantara *et al*. (2018), where the CEM 3.0 displayed the largest errors (*i.e.* hundreds of meters; top plot), and also artificial terraces associated with decimal truncation in the CEM 3.0 were evident (bottom plot). SRTM does not show the same problems than CEM 3.0. Quite the opposite, the errors have reasonable magnitudes (below 6 m), and they remain around 3 m most of the time. On the other hand, elevation shows gradual changes, not artificial terraces, like CEM 3.0. The errors also show a patchy pattern across the whole region. This is probably associated with SRTM's scanning paths. In this case, the largest water bodies show negative errors.

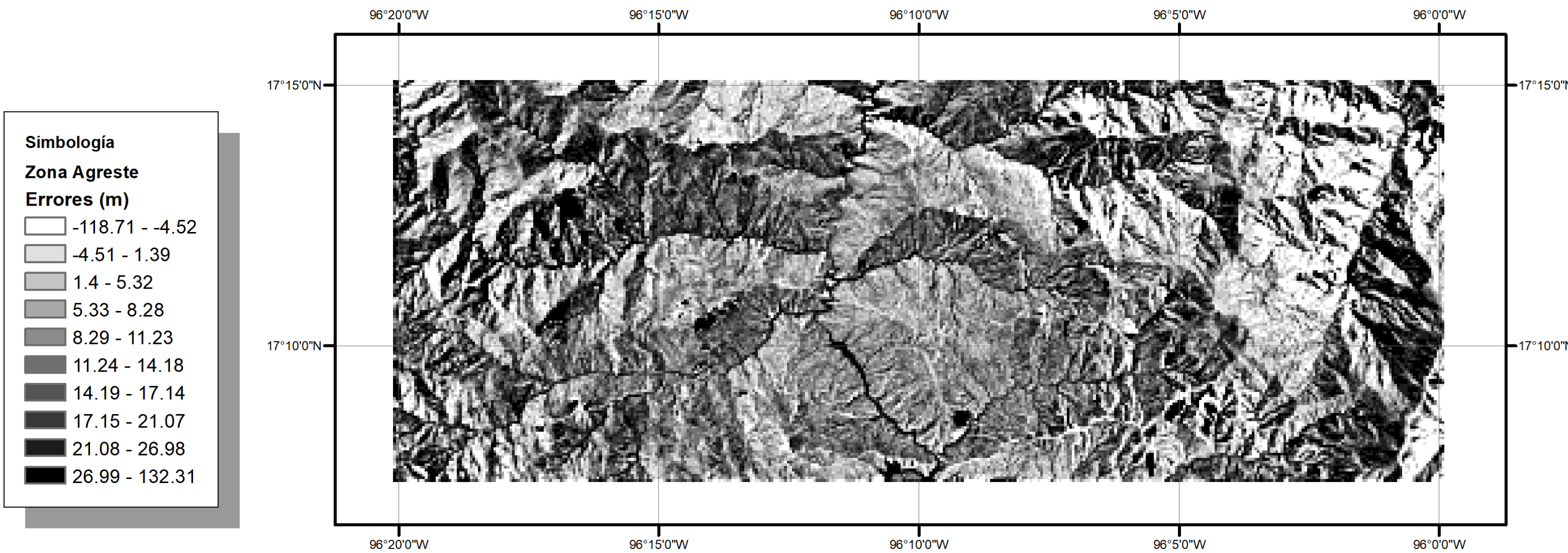

Figure 6. Map of errors for the hilly region, classified in ten quantiles.

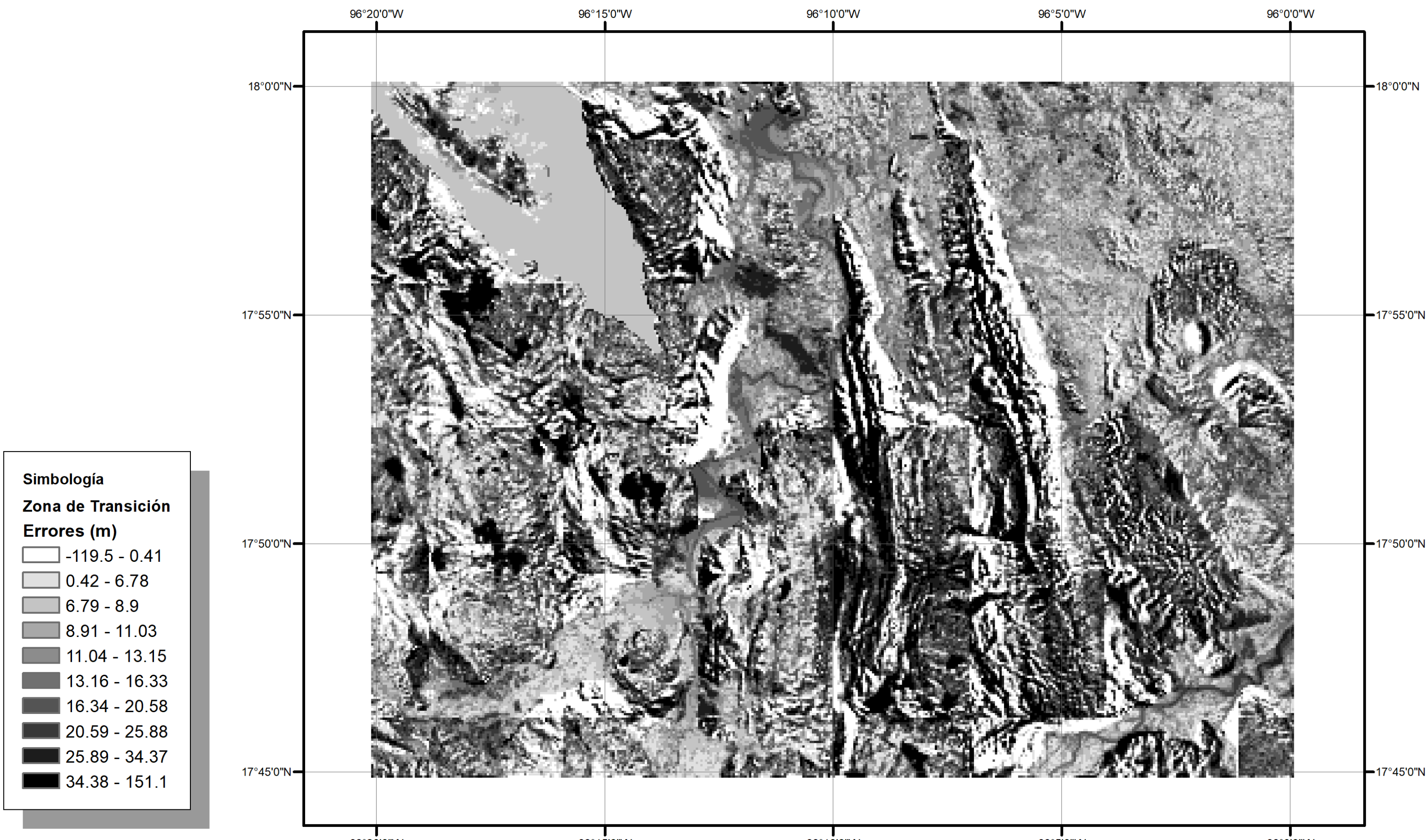

Figure 7. Map of errors for the transitional region, classified in ten quantiles.

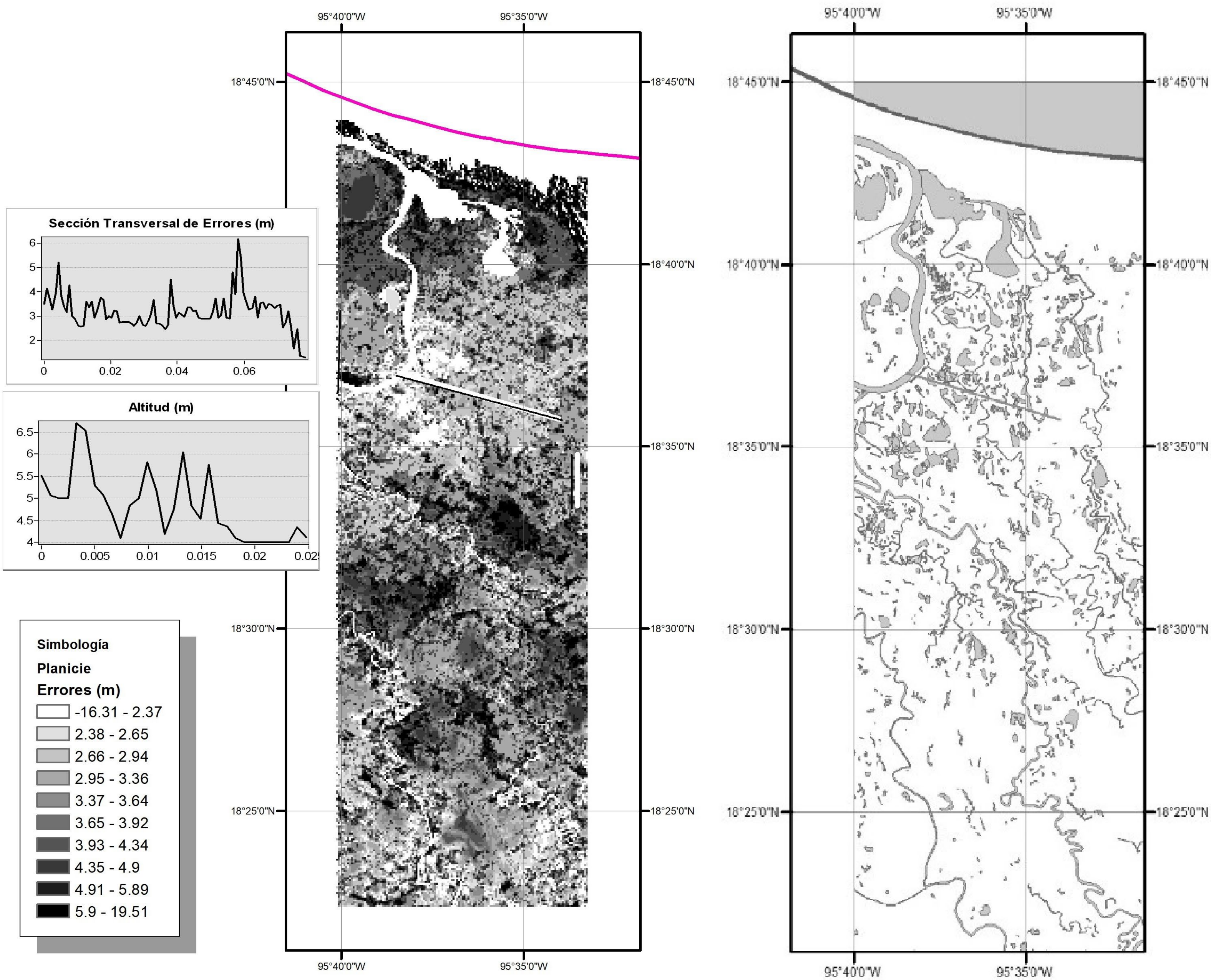

Figure 8. Map of water bodies (right side); errors for the floodplain(left), classified in ten quantiles, and cross-sections of errors (upper plot) and elevations (lower plot) for regions with high and low errors, correspondingly, in the CEM.

Finally, Figure 9 shows three-dimensional errors as function of elevation, aspect and slope for each region. Unlike, errors in CEM 3.0, the only scattergram that shows a clear relationship is the one associated with aspect in the hilly region. In this case, the histogram has a parabolic behavior. The maximum overestimations (~10 m) are observed around 180°, while the largest underestimations (~-20 m) are observed at 0° and 360°. This behavior was also observed in the CEM 3.0. Errors in remote sensors can be associated with the angle between the sensor and the surface (Bater & Coops, 2009; Goulden *et al*., 2016). We can infer that LiDAR is susceptible to this error because the same pattern was apparent during the evaluation of CEM 3.0, which is not derived from any remote sensing technique. However, we cannot know to what degree the pattern in the current comparison is also associated to SRTM, because both remote-sensing errors (in LiDAR and SRTM) would be intertwined.

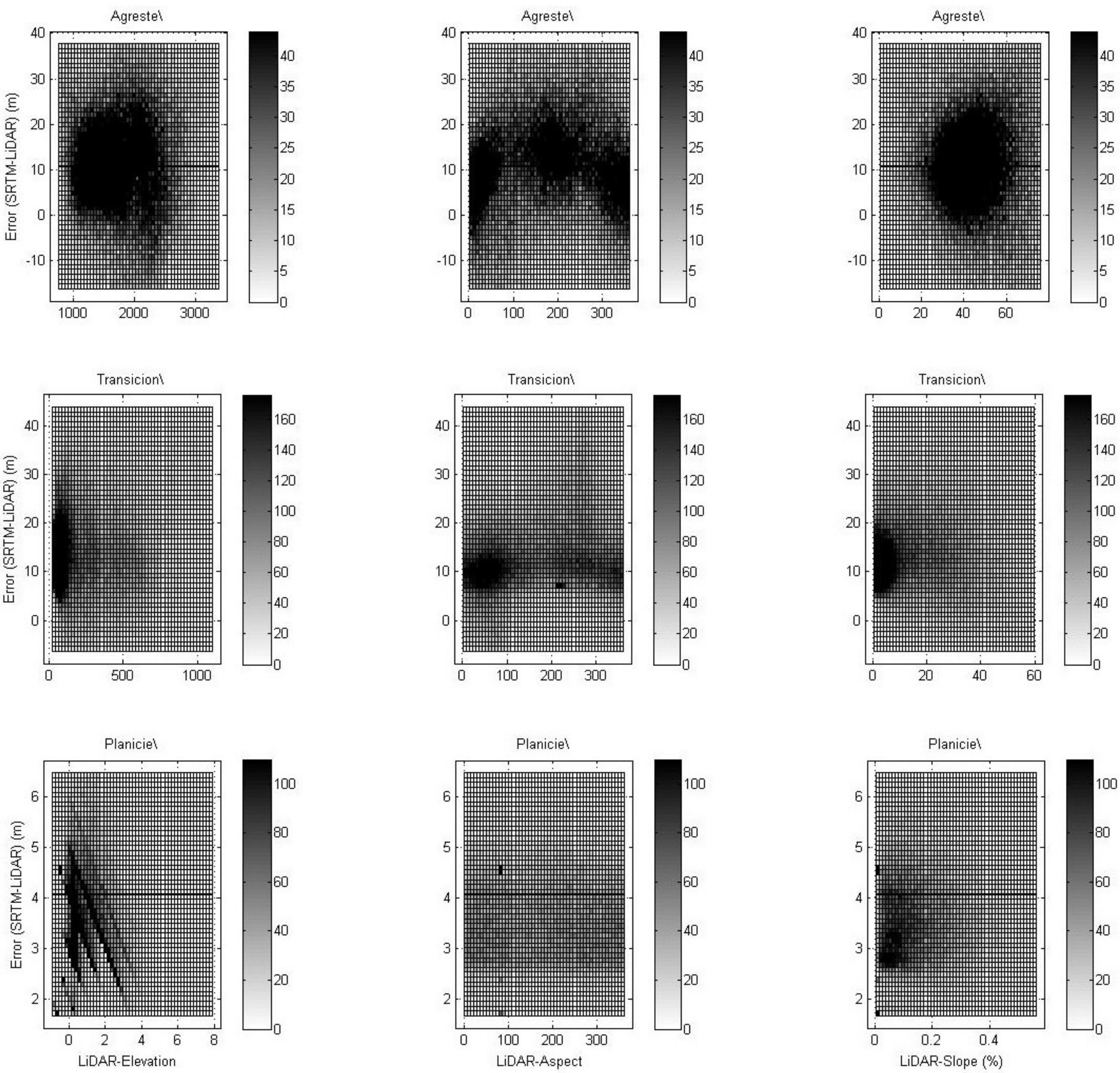

Figure 9. Tri-dimensional histograms of errors *versus* elevation, aspect and slope, from left to right, correspondingly, and for each region: hilly, transition, and floodplain, from top to bottom, correspondingly. The gray scale corresponds to the frequency.

On the other hand, scattergrams for the transitional region show a systematic error, independent of topographic factors. For example, the scattergram of aspect *versus* errors shows a clear horizontal line with the most prevalent errors around 10 m. This is also observed in the other regions although not as clearly. The scattergrams of elevations, for example also show a predominant error around 10 m. Again, these results suggest SRTM may benefit from a simple linear regression calibration.

Conclusions

The evaluation of SRTM DEMs at three representative regions of elevation, using LiDAR as a benchmark shows that both, the hilly and transitional regions have very similar elevation values in LiDAR and SRTM. In fact, the linear regressions have slopes close to one, and intercepts relatively close to zero. The most evident difference between both regions is that errors are relatively larger in the transitional region. The floodplain, on the other hand, does not show a linear relationship between elevations. The scattergram shows a parabolic shape for most of the data. In fact, the best polynomial fit for this behavior was a third degree polynomial. When taking into consideration the range of values, the floodplain is the region with the highest relative errors.

Error maps show that the most important factors for error distribution are water bodies and aspect. In general, errors in water bodies tend to be negative, *i.e.* underestimation. On the other hand, in the transitional region, negative errors are mostly found on slopes facing east, while positive errors are mostly found on slopes facing west. Cross-sections of elevations and errors show that, unlike CEM 3.0 (the official DEM for Mexico published by INEGI), SRTM does not show extreme errors (at hundreds of meters where actual elevations are below 8 m) around water bodies nor artificial terraces due to decimal truncation.

Three-dimensional histograms of errors *versus* topographic features (elevation, aspect and slope) show that errors seem both symmetrical around a constant value in most of the cases, which suggest SRTM could benefit from a simple calibration. The only exception is the histogram of errors *versus* aspect in the hilly region, where we can observe that overestimations are mostly positive, with the only exception being slopes facing north, where we can observe mostly underestimations.

As mentioned earlier, we consider this evaluation a preliminary step for either a simple calibration or the creation of a multi-source DEMs in Mexico. Currently, most DEMs available in Mexico have both advantages and disadvantages. For example, LiDAR has a patchy coverage; CEM has extreme errors in floodplains; and SRTM has a low resolution. Thus, a possible solution could be the creation of multi-source DEMs, which take advantage of all the strengths to mitigate individual weaknesses. However, formal evaluations of DEMs available in Mexico are a necessary preliminary step before creating multi-source DEMs. We expect evaluations like this will be able to provide some guidance during the creation of multi-source DEMs in the near future.

Acknowledgements

We wish to acknowledge the support of the projects: 1) "Mapa de inundación para el Municipio de Tula, Hidalgo", and 2) "Caracterización de intrusiones magmáticas recientes en la Sierra de Pachuca", for the development of this paper. We would also like to thank reviewers and editors for their valuable contribution to the improvement of this article.

References

Baghdadi, N., Cavelier, S., Chiles, J. P., Bourgine, B., Toutin, T., King, C., Daniels, P., Perrin, J., & Truffert, C. (2005). Merging of airborne elevation data and Radarsat data to develop a Digital Elevation Model. *International Journal of Remote Sensing*, 26(1), 141-166. DOI: 10.1080/01431160410001709020

Bater, C. W., & Coops, N. C. (2009). Evaluating error associated with lidar-derived DEM interpolation. *Computers & Geosciences*, 35(2), 289-300. DOI: 10.1016/j.cageo.2008.09.001

Gesch, D., & Wilson, R. (2001). Development of a seamless multisource topographic/bathymetric elevation model of Tampa Bay. *Marine Technology Society Journal*, 35(4), 58-64. DOI: 10.4031/002533201788058062

Goulden, T., Hopkinson, C., Jamieson, R., & Sterling, S. (2016). Sensitivity of DEM, slope, aspect and watershed attributes to LiDAR measurement uncertainty. *Remote Sensing of Environment*, 179 (Suppl. C), 23-35. DOI: 10.1016/j.rse.2016.03.005

INEGI, Instituto Nacional de Estadística y Geografía. (2017). *Modelo digital de elevación de alta resolución LiDAR, Tipo terreno con resolución de 5 m*. Recovered from http://www.inegi.org.mx/geo/contenidos/datosrelieve/continental/presentacion.aspx

Jarvis, A. G. E., Reuter, H. I., & Nelson, A. D. (2008). *Hole-filled SRTM for the globe: Version 4: Data grid.* Recovered from http://srtm.csi.cgiar.org/

Uribe-Alcantara, E. M., Escamilla-Casas, J. C., & Cruz-Chavez, E. (2018). Evaluación del modelo de elevación digital oficial mexicano (continuo de elevación mexicano, CEM 3.0) de INEGI. *Tecnología y ciencias del agua*, 9(6), 153-172. DOI: 10.24850/j-tyca-2018-06-07