

Spatial autocorrelation in the generation of high amplitude DTMs for 3D urban models from LiDAR data

Cristina Catita
Instituto Dom Luiz,
Universidade de Lisboa
Campo Grande 1749-016
Lisboa, Portugal
cmcatita@fc.ul.pt

Paula Redweik
Centre of Geology
Universidade de Lisboa
Campo Grande 1749-016
Lisboa, Portugal
pmredweik@fc.ul.pt

Joana Pereira
Dep. Eng. Geográfica, GE
Universidade de Lisboa
Campo Grande 1749-016
Lisboa, Portugal
fc39452@alunos.fc.ul.pt

Abstract

With this study we investigate a methodology that is capable of reconstructing a DTM from LiDAR data in a city that is often called the city of the seven hills, Lisbon. The strong changeable height that characterizes the terrain defies algorithms that have been successfully applied in flat terrain. The developed methodology is based on the analysis of local indicators of spatial autocorrelation (LISA) using Lisbon's LiDAR datasets. To generate a ground surface the approach went through successive approximations in order to obtain a rough terrain surface representing the local tendencies. The residuals obtained by subtracting this new grid from the DSM were evaluated with the Local Moran Indicator, allowing the classification in ground points and non-ground points based on the spatial association of positive and negative residuals. The methodology was used to produce DTMs that are compatible with 3D building models.

Keywords: LISA, Moran, Spatial analysis, Urban Elevation Models

1 Introduction

Three dimensional urban models intend to describe geometrically and semantically urban objects like buildings, trees, city furniture, streets and other typical urban items. They allow not only a dynamic visualization of a realistic urban environment but also to conduct 3D spatial analysis not feasible in 2D, for example when studying phenomena that have a three dimensional dispersion such as noise and flood prediction models. The level of detail of such urban models can vary according to the objectives they are built for. For a digital detailed geometric description of buildings several modelling techniques can be used such as photogrammetry, laser scanning and CAD 3D modelling tools. Texture mapping operations can provide a realistic look to the models when this is relevant for the actual application.

The surface upon which the building models lie in a 3D urban model should describe the local relief with a level of detail similar of that used for representing the objects.

The automated generation of relief models has been explored by several new technologies where airborne LiDAR scanning assumes an important role and is still a topical area of research.

Deriving a digital terrain model (DTM) from dense LiDAR data, especially for large-scale applications, requires procedures for automatic removal of objects covering the ground surface. Automatic classification of LiDAR clouds into ground and non-ground points has been tested by several different filter algorithms [8]. Methods based on adapted interpolators [2, 5], local slope detectors [10, 11] and mathematical morphological operators [4, 9] are the most popular approaches. Most algorithms have been successfully applied in cities where the ground surface doesn't show sudden height changes.

The object of this study was to investigate a methodology capable of reconstructing the DTM from LiDAR data of a city with high spatial variability on slope gradient of topographic

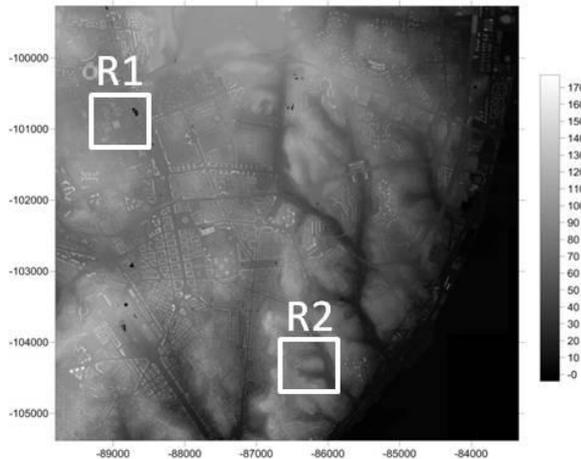
forms. The obtained terrain surface should be compatible with modeled 3D buildings rather than ignoring their existence, replacing buildings footprint locations by horizontal terraces with an adequate height. The developed methodology is based on the analysis of local indicators of spatial association (LISA) [1, 3]. The Local Moran statistic (I) is applied on this study to assess the significant spatial clustering of similar values around an individual location. In order to investigate this, differences between a generated rough surface, following the local terrain tendency, and the original Digital Surface Model (DSM) are evaluated in terms of spatial association. Mapping this measure allows the identification of local clusters of hot and cold spots as well as spatial outliers of the corresponding residuals. Results show that hot spots (similar spatial trends of high values) are mostly related with non-ground objects whereas cold spots (similar spatial trends of low values) are related with ground points. In hilly regions this separation is not so clear, so that additional considerations on the behavior of trend surfaces have to be taken into account as explained in 2.2. The methodology here proposed is tested on several LiDAR data sets from Lisbon. Figure 1 shows two examples of the tested urban areas: region 1 (R1) presents smooth slope terrain, high buildings and large streets (recent town area); region 2 (R2) presents very steep slopes, high building density and narrow streets (old town area).

2 Methodology

The data sets used in the project are extracted from a LiDAR DSM, provided by LOGICA, covering a significant part of the built area of Lisbon (39 km²) (Figure 1). In the origin of the DSM was a LiDAR coverage made 2006, where elevation and intensity of the first and last returns were recorded for each laser pulse with a TopoSys II airborne LiDAR scanner (83 kHz pulse rate). The equipment was flown on a helicopter and the density of the original sample was 20 points per m². The

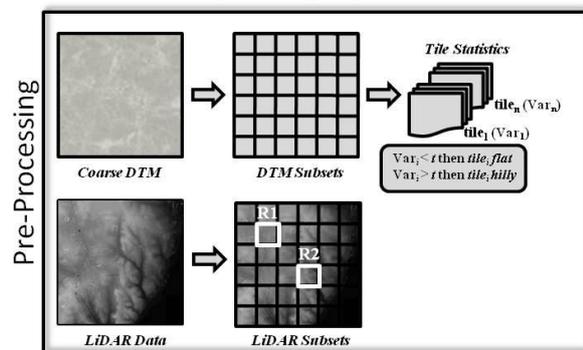
data were resampled to a 1m x 1m raster and present a documented horizontal accuracy of 0.5 m and a vertical accuracy of 0.15 m [7].

Figure 1: LiDAR DSM, Lisbon with test regions 1 and 2.



Because of computing and software limitations, the original LiDAR DSM has been arranged as a collection of tiles of 500 per 500 meters, where sets may overlap (Figure 2). For the entire study region a coarse DTM derived from an independent source was also calculated. The variance of the altitude is then determined in the region covered by each tile in the coarse DTM and the terrain relief is classified in two categories: flat (variance < 5m) and hilly (variance > 5m). This evaluation depends on the dimension of the subsets and on the local terrain characteristics. After this preliminary evaluation, the methodology is differentiated for each kind of terrain.

Figure 2: Flowchart depicting the methodology used on pre-processing step.

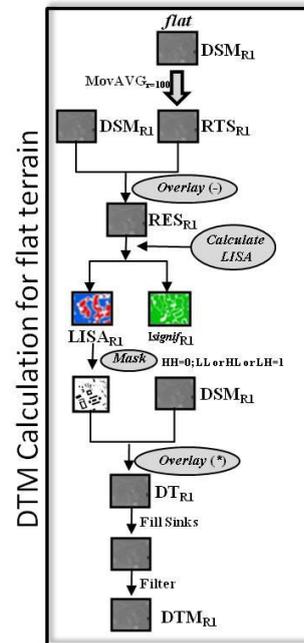


2.1 Flat terrain

The region 1 of our study corresponds to the campus of the Faculty of Sciences (FCUL). This is a typical urban area containing several teaching and research buildings, a museum and part of the tree canopy of Campo Grande and the

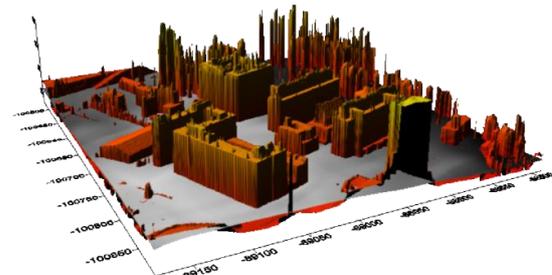
museums garden with some very high trees. Figure 3 shows the applied methodology.

Figure 3: Flowchart depicting the methodology used for Region 1 (R1) corresponding to flat terrain.



The first step on the DTM calculation was to produce a rough terrain surface (RTS) applying a moving average to the data points with a neighborhood of 100 x 100 cells, which represents the local elevation tendencies. This new surface is smoother than the original raster DSM and corresponds locally to an average height between ground and non-ground points. The residuals (RES) obtained by subtracting this new surface from the DSM grid are mostly negative for ground points and positive for non-ground points (Figure 4).

Figure 4: Original DSM (in shades of orange) overlaid with rough terrain surface (in gray) of test Region 1.



On the second step of our procedure, those residuals are evaluated in terms of local spatial association. A local indicator of spatial autocorrelation (LISA) is computed for

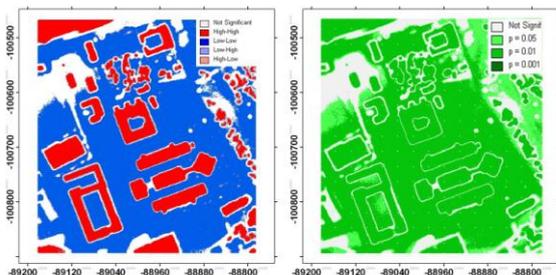
each residual grid point in order to classify how strong the spatial association is between the point and its neighbors.

The extent of significant spatial clustering of similar values (hot or cold spots) around an observation identifies stationary regions while non-stationarity is demonstrated by spatial heterogeneity. For this purpose, the Local Moran statistic I_i is applied [1]:

$$I_i = \frac{Z_i - \bar{Z}}{S^2} \sum W_{ij}(Z_j - \bar{Z}) \quad (1)$$

where \bar{Z} is the mean intensity over all observations, Z_i and Z_j are the intensities of observations i and j , respectively (where $i \neq j$), S^2 is the variance over all observations, and W_{ij} is a distance weight for the interaction between observations i and j . Figure 5 (left side) shows the location of grid points classified according to the type of spatial autocorrelation. The high-high (red) and low-low (blue) suggest clustering of similar values indicating locations of positive autocorrelation, whereas, the high-low (pink) and low-high (light blue) locations indicate spatial heterogeneity or negative autocorrelation. Locations with significant local Moran statistics are represented in Figure 5 (right side) in different shades of green (the corresponding p values of significance are given in the legend), while areas of non-significant Moran I_i are represented by white points.

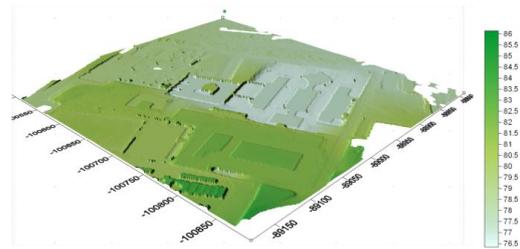
Figure 5: LISA (left) and Moran significance (right) maps of Region 1 (R1).



Since non ground points are well discriminated by the significant high-high clusters, these were used to generate a raster binary mask in the third step of our approach. The grid values of this mask are set to 0 on high-high points and to 1 on any other points. Afterwards, the product of the original LiDAR DSM and the raster produced mask is calculated (DT surface on Figure 3). This way the remaining points were considered ground points on this new surface and kept their LiDAR elevation values. The removed grid points, assigned with zeros, represent mostly non-ground points and were on this step considered artificial sinks. Using this methodology, about 83% of non-ground points were detected correctly comparing to cartography.

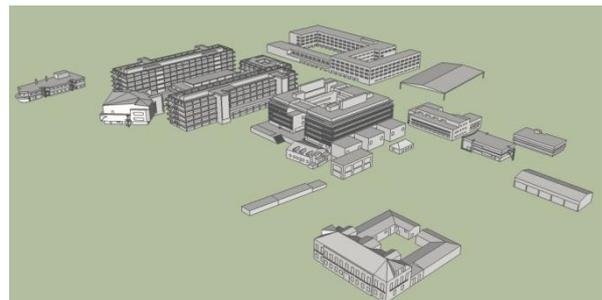
Finally, filling only the artificial sinks with the height of the lowest boundary cell of each sink consisted on the almost last procedure of the methodology here presented. Because the remaining non-ground points were sparse, a minimum spatial filter with an adequate kernel size can produce suitable results. Following this approach, the DTM obtained for our geographic region 1 is represented on Figure 6.

Figure 6: DTM produced for Region 1.



For large scale applications it is often relevant to have quite detailed building models. Structures on the roof could possibly be obtained from high density airborne LiDAR data, but details of vertical facades can't be modeled from such data. Therefore, 3D buildings were interactively generated using the modeling software from Google, Google SketchUp 8 (SU) (Figure 7). The extrusion of the buildings resulted from the combination of the building footprints, obtained by photogrammetry, and the height information measured in stereo models. All the other components, such as doors, windows and balconies, were added based on photographs of the site and images from Bing's "Bird's Eye View" taking advantage of the ability of the SU software to juxtapose photographs as textures to the 3D surfaces and model from there.

Figure 7: Buildings model of the Campus of FCUL



The integration of the 3D building models with the generated DTM (Figure 8) was then a simple procedure, since the georeference was common and the buildings were modeled ground up, meaning that no underground floors were considered in the modeling.

Figure 8: 3D model of FCUL Campus with generated DTM.



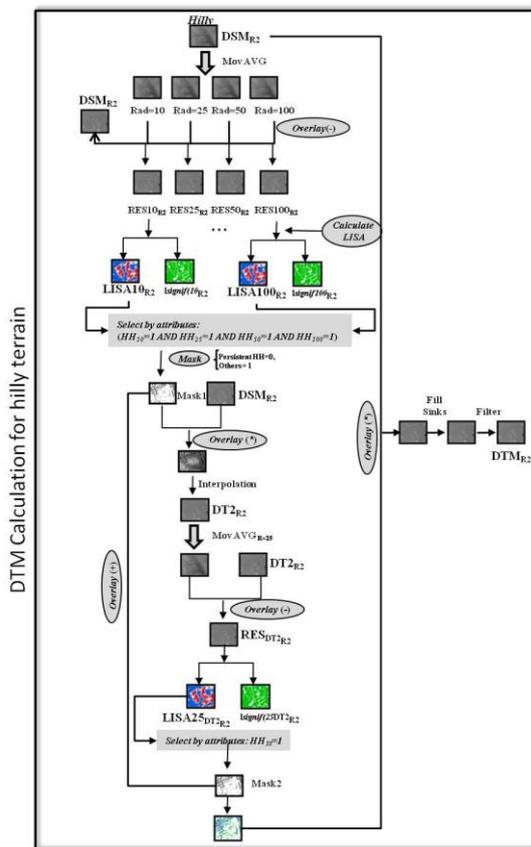
Attending to the purpose of the urban model to be used in solar radiation potential studies, this representation was found suitable.

2.2 Hilly terrain

The above described method can be successfully applied to extract flat terrain from the LiDAR DSM, but it doesn't work well in hilly terrain, since hills and buildings behave similar in terms of spatial autocorrelation of the residuals to a trend surface.

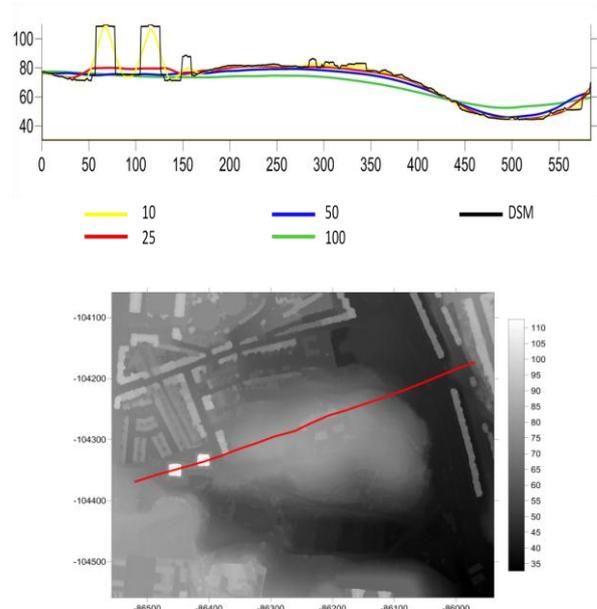
Considering the relief of region 2, one of the critical zones of the DSM with steep slopes and buildings both on the valley and on the top of the hill, another approach was followed. Figure 9 shows the applied methodology.

Figure 9: Flowchart depicting the methodology used for Region 2 (R2) corresponding to hilly terrain.



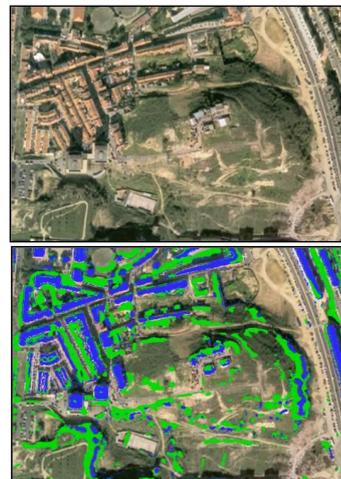
Four trend surfaces were calculated using a moving average with different neighborhood radius. As the radius gets smaller the trend surfaces become less smooth, as exemplified on the profiles in Figure 10: 100 (green), 50 (blue), 25 (red) and 10 cells (yellow). Also the residuals from the trend surfaces to the original DSM become smaller in the hills than in the buildings. For each trend surface, these residuals were evaluated in terms of local spatial autocorrelation generating four different LISA maps similar to those shown in Figure 5.

Figure 10: Trend surfaces profiles (above) and profile location (below).



Significant clusters of high residuals detected by the LISA simultaneously on all trend surfaces should indicate non ground points to be removed from the DSM, but due to the relief heterogeneity and built density, a relevant portion of non ground points can't be detected at this stage. An iteration is needed. Those points are going to be part of a final mask to be completed in further steps. On the other hand they are also removed from the DSM in this step, resulting in a surface with holes which were filled with interpolated heights of their boundary. This way, a second approximate elevation surface DT2 is generated where the higher non ground points are absent but some lower non ground points still exist (Figure 11).

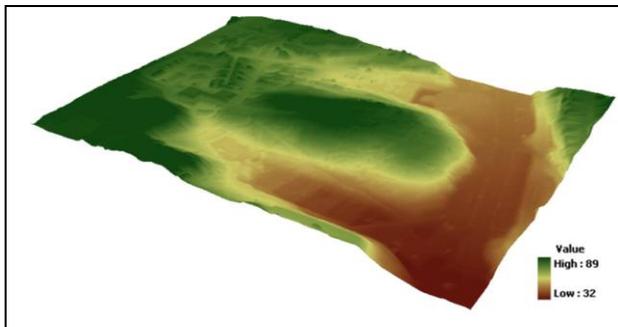
Figure 11: Orthophoto (above) and mask (below) with classified non ground points after two iterations (blue 1st and green 2nd iteration).



A new trend surface is generated from the surface DT2 with a moving average in a 25 cells neighborhood. This neighborhoods dimension adapts the better to the ground and is not so influenced by the buildings as smaller radius (see Figure 10). The residuals to DT2 are once again analyzed with LISA in order to detect new high-high point clusters. The points in the clusters are added to the final mask. The process is iterated until there is no significant clustering of residuals meaning that the residuals are randomly located in space. This can be tested by a global indicator of spatial autocorrelation (e.g. Global Moran's I). Values of Global Moran ranges from -1 (indicating perfect dispersion) to +1 (perfect correlation). Values near zero indicate a random spatial pattern.

Finally, the last calculated mask is then applied to the original LiDAR DSM removing non ground points and, like in 2.1, the generated holes are considered artificial sinks being consequently filled with the height of the lowest boundary cell of each sink (Figure 12).

Figure 12: DTM of region 2.



3 Results

This study presents a new methodology for generating a DTM based on evaluation of the local spatial autocorrelation of terrain surface differences. Starting from a dense DSM of a complex cityscape and steep terrain, this approach removes iteratively non ground objects approximating the data to terrain topography.

Moreover, classified ground points preserve the original LiDAR DSM height information. The areas classified as non-ground can be either filled or interpolated depending on the application requirements. Filling those areas with an appropriate constant value in order to create horizontal terraces, allows a suitable integration with 3D building models.

The resulting DTM is as detailed as required for most urban studies.

Nevertheless, some disadvantages have to be pointed out:

- the high density of the LiDAR DSM limits the process in terms of calculation time and area dimensions;
- the iterative and neighbourhood focused nature of the methodology is very expensive in terms of computing resources;
- the success of the methodology still relies too much on interactive judgement.

The methodology is being tested in other areas in order to be improved in terms of effectiveness and degree of automatism.

References

- [1] Anselin, L. Local indicators of spatial association – LISA". *Geographical Analysis*, 27, 93-115, 1995.
- [2] Chen, Q., Gong, P., Baldocchi, D. and Xin, G. Filtering airborne laser scanning data with morphological methods. *Photogrammetric Engineering and Remote Sensing* 73(2), 175-185, 2007.
- [3] Getis, A and Ord, J.K. Local Spatial Statistics: An Overview. In *Spatial Analysis: Modeling in a GIS Environment*, edited by P. Longley and M. Batty. John Wiley & Sons: New York, 1996.
- [4] Kilian, J., Haala, N. and English, M. Capture and evaluation of airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 31(B3), 383-388, 1996.
- [5] Kraus, K. and Pfeifer, N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 53(4), 193-203, 1998.
- [6] Kraus, K. and Pfeifer, N. Advanced DTM generation from LiDAR data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34(3/W4), 23-30, 2001.
- [7] Leitão, J., Boonya-aroonnet, S., Prodanovic, D. and Maksimovic, C. Influence of DEM resolution on surface flow network for pluvial urban flooding and simulations of integrated system. *Proceedings of the 11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK, 11pp., 2008.
- [8] Liu, X. Airborne LiDAR for DEM generation: some critical issues *Progress in Physical Geography*, 32 (2008), pp. 31–49, 2008.
- [9] Lohmann, P., Koch, A. and Schaeffer, M. Approaches to the filtering of laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 33(B3), 540-547, 2000.
- [10] Vosselman, G. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 33(part B3/2), 935-934, 2000.
- [11] Zhang, K. Q., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J. H. and Zhang, C. C. A progressive morphological filter for removing nonground measurements from airborne LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing* 41(4), 872-882, 2003.