MSEL
MODELLING IN SCIENCE EDUCATION AND LEARNING

Modelling in Science Education and Learning
Volume 6(3), No. 10, 2013.
Instituto Universitario de Matemática Pura y Aplicada

Mathematical modelling applied to LiDAR data

Javier Estornell, Jesus M. Martí-Gavilá, M.Teresa Sebastiá, Juan Andres González
Universitat Politècnica de València
jaescre@cgf.upv.es, jemarga@cgf.upv.es, mtsebastia@hma.upv.es, juagonr1@hma.upv.es

Abstract

El objetivo de este artículo es explicar la aplicación de varios cálculos matemáticos de LiDAR (Light Detection and Ranging) para estimar los parámetros de la vegetación y el modelado del relieve de una zona forestal en la localidad de Chiva (Valencia). Para representar la superficie que describe la topografía de la zona, en primer lugar, se aplicaron filtros morfológicos iterativamente para seleccionar los puntos de tierra Lidar. A partir de estos datos, se aplicó la estructura de red irregular triangulada (TIN) para modelar el relieve de la zona. A partir de datos LiDAR se calculó también el modelo de la altura del dosel (CHM). Este modelo ha permitido la obtención de suelo desnudo, arbustos y cartografía de la vegetación árbol en el área de estudio. Además, la biomasa se estimó a partir de medidas tomadas en el campo en 39 parcelas circulares de radio de 0.5*m y el percentil 95 de los datos de altura LiDAR incluidos en cada parcela. Los resultados indicaron una alta relación entre las dos variables (biomasa medida y percentil 95) con un coeficiente de determinación* (R^2) *de* 0.73*. Estos resultados ponen de manifiesto la importancia del uso de modelos matemáticos para obtener información de la vegetación y el relieve del terreno a partir de datos LiDAR.*

The aim of this article is to explain the application of several mathematic calculations to LiDAR (Light Detection And Ranging) data to estimate vegetation parameters and modelling the relief of a forest area in the town of Chiva (Valencia). To represent the surface that describes the topography of the area, firstly, morphological filters were applied iteratively to select LiDAR ground points. From these data, the Triangulated Irregular Network (TIN) structure was applied to model the relief of the area. From LiDAR data the canopy height model (CHM) was also calculated. This model allowed obtaining bare soil, shrub and tree vegetation mapping in the study area. In addition, biomass was estimated from measurements taken in the field in 39 circular plots of radius 0.5 m and the 95th percentile of the LiDAR height data included in each plot. The results indicated a high relationship between the two variables (measured biomass and 95th percentile) with a coefficient of determination (R^2) *of* 0.73*. These results reveal the importance of using mathematical modelling to obtain information of the vegetation and land relief from LiDAR data.*

Keywords: LiDAR, TIN, filtro morfológico, DEM, Interpolación.
LiDAR, TIN, morphological filter, DEM, Interpolation

1 Introduction

The framework of this study is related to the contents of the elective subjects "Applied Remote Sensing" and "Applied Remote Sensing in the coastal zone". The first subject is taught since the 2004/05 academic year in the fourth year of the Degree in Environmental Sciences. The second subject is taught since the 2011/12 academic year in the Master's Degree in Assessment and Environmental Monitoring of Marine and Coastal Systems, in the Escola Politècnica Superior de Gandia (Universitat Politècnica de València). In these subjects, LiDAR technology is used in several environmental fields, such as viewshed analysis, modelling of dunes and mainly forest and agriculture applications.

The LiDAR (Light Detection And Ranging) technology is an active remote sensing system that is based on measuring the elapsed time between the emission of a pulse of energy and its arrival at an airborne sensor in this case, after being reflected by some element of the terrestrial surface. These data contain information of the coordinates x, y and z of the points where the reflections occur, either on the ground or on any object above the topographic surface such as vegetation and buildings (Figure 1). One of the main handicaps of using LiDAR data is that ground and non-ground points must be identified from the raw data. To do this, filtering algorithms are applied. A review of methods can be found in [1]. An important group of them are based on the use of morphological filters what involves selecting minimum or maximum values for a specific location [2]. These ones will be applied in this article.

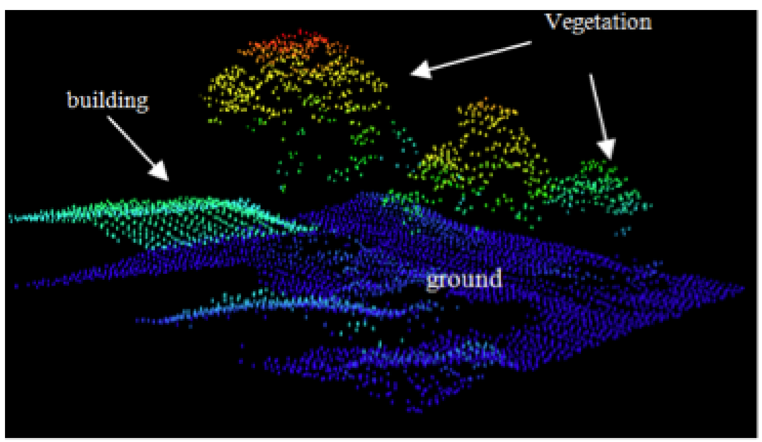


Figure 1: Information generated by an aerial LiDAR system in which points can be distinguished for the ground, vegetation and a building.

The first step is to create a continuous surface of the land relief from discrete data (ground points), that is to obtain a Digital Elevation Model (DEM). A DEM is defined as a numeric structure of data that represents the spatial distribution of the elevation of a topographic surface [3]. To do this transformation, two structures are commonly used: Triangulated Irregular Network (TIN) and raster (image format) (Figure 2). A TIN is a vector data structure that divides a geographic area into non-overlapping and contiguous triangles. The vertices of each triangle are defined in this case by the coordinates (x, y and z) of the LiDAR ground points. The assumption of this structure is that land relief is defined from the plane defined by each triangle. In the second case, raster format, the terrestrial surface is defined by a regular array of cells that contain z values.

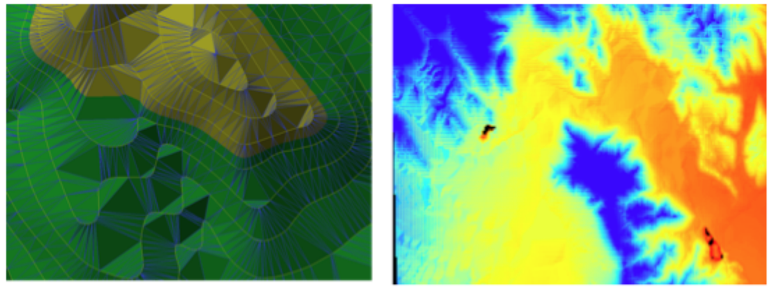


Figure 2: Examples of a TIN model (left) and a raster model (right) to represent a Digital Elevation Model.

The non-ground points allow calculating a digital surface model (DSM) defined as a digital representation of earth's surface in raster format that contains information of the elevation of geographic features such as buildings and vegetation (Figure 3). A DSM contains the same data that a DEM in the bare ground areas. The difference between a DSM and a DEM is defined as the normalized Digital Surface Model (nDSM) that in forest environments coincide with the Canopy Height Model (CHM). These surfaces contain the height of the geographic features (vegetation, buildings).

LiDAR data have been successfully applied in several fields such as: changes in beach sand [4]; building extraction [5]; and mainly in forest applications [6, 7, 8]. The result of combining the non-ground points and the DEM to convert point elevations into heights may be used as entry data too [9].

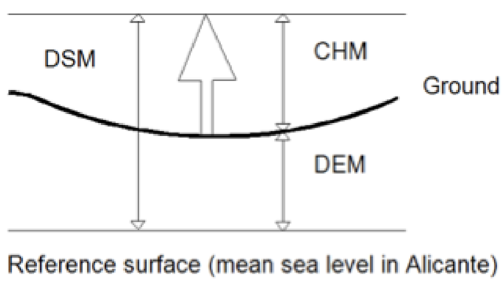


Figure 3: Scheme to represent a DSM, a DEM, and a CHM.

In forest applications some variables of trees (height, biomass, and volume) have been estimated by plots using several statistics derived from LiDAR data such as maximum height, mean height and several percentiles of the height distribution in the sampled plots [10, 11]. These are potential independent variables to estimate some parameters of the vegetation. The foundation of this methodology is based on the existing relations among the statistics derived from LiDAR data and some field measurements of the trees. For example, there is a strong relationship between the volume of a tree and its height that can be derived from LiDAR data. In this article, the estimation of shrub biomass will be done following this methodology.

The objective of this article is to apply mathematical modelling to the LiDAR data to obtain a DEM, a DSM, a CHM for the determination of the presence of shrub vegetation and estimation of biomass.

2 Methodology

The study area is located in the municipality of Chiva (Spain), covering an area of $10km^2$, for which LiDAR data was available. This zone is mountainous and predominantly covered by shrub vegetation, in which the elevation of the terrain varies between 442 and $1,000m$ above sea level with an average slope of 45 %. The most abundant species is *Quercus coccifera*, which is widely spread in the Mediterranean region. The mean height of the shrub vegetation found in this area was $1.27m$ and the standard deviation was $0.29m$.

The filtering process applied to select ground data from raw LiDAR data utilized an algorithm which uses an iterative process to select minimum elevations (Figure 4). These points are considered as ground points. To select them a series of progressively smaller windows are used. First, an initial DEM is calculated using the minimum points in an entry window ($10m$). In a second step, new minimum elevations are selected by using a smaller window ($5m$). These points are compared to the values of the initial DEM. The objective is to remove from all minimum points selected in this step, the non-ground points. In a dense vegetation area the likelihood of selecting points with minimum elevation that belong to the vegetation is high. This fact can be explained by the difficulty of some LiDAR systems to penetrate the dense vegetation cover. Then, to increase the accuracy of a DEM the non-ground points have to be removed introducing a height threshold. In the comparison above mentioned, the points whose differences are higher than this threshold are removed. The height threshold used in this application was $1m$. With the ground points, it was computed a new digital elevation model (DEM2). The algorithm has a new step in which it was used a smaller window ($2.5m$) to select minimum points. A new comparison was done among the minimum points selected in this step and the DEM2. It was used the same threshold ($1m$) to remove the non-ground points. With the ground points it was calculated the final DEM. Further information of this algorithm can be found in [12]. Thus, when this filtering methodology is applied, DEM accuracy is influenced by window size, number of iterations and height thresholds.

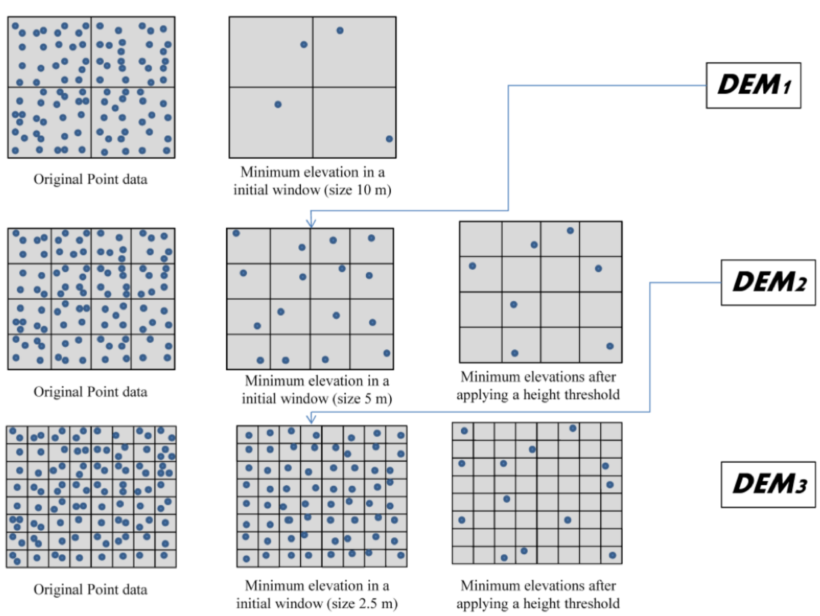


Figure 4: Algorithm used for calculating DEM.

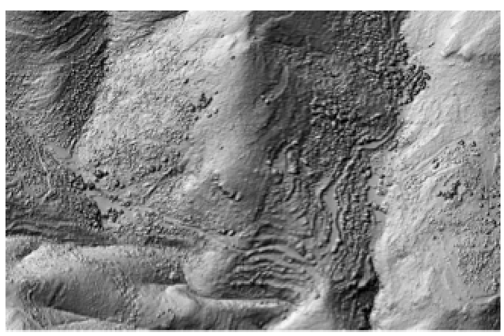
Each DEM of the algorithm was calculated using a Delaunay triangulation[13]. This triangulation accomplishes the following properties: triangles must be as equilateral as possible; the circumference of the three vertices of the triangle contains no other point; the order in which the points are processed does not affect the result of triangulation.

A digital surface model (DSM) was also created selecting the highest elevation points in a neighbourhood of cell size $1 \times 1m^2$. The difference between the image of the DSM and the DEM provided the CHM. The cells in this model contain information about height of vegetation and ground. Considering the CHM, three intervals were created to separate ground, shrubs and tree vegetation. Cells with values in height from $0m$ to $0.3m$ were classified as ground; cells with values in height from $0.3m$ to $2.5m$ were classified as shrubs; cells with values in height larger than 2.5 were classified as tree vegetation.

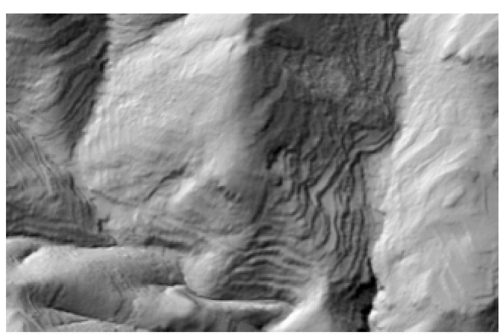
Finally, it was calculated a model to estimate shrub biomass from statistics derived from LiDAR data in 39 plots of $0.5m$ where this variable was measured in field. For this purpose field vegetation was cut and weighed within each plot. The 95th percentile of the LiDAR heights within each plot was used as independent variable.

3 Results and discussion

In Figure 5, the differences between the DSM and the DEM can be observed in an area with vegetation. In these details, it can be visualized that the vegetation has been properly filtered in the process of calculating a DEM (Figure 5(b) with respect to the DSM (Figure 5(a)).



a)



b)

Figure 5: Details of the shaded relief image of the DSM (a) and of the DEM (b).

 ISSN 1988-3145

In Figure 6 the CHM can be observed. It has been visualized creating three classes. For this purpose three ranges of heights have been defined: $0 - 0.3m$; $0.3 - 2.5m$; $h > 2.5m$. It can be seen that the number of cells corresponding to the last interval is less than for the other classes. The heights of this interval correspond to the trees, indicating that the study area has a low presence of this vegetation. Conversely, the class with the highest number of cells is the second $(0.3m - 2.5m)$, which comprises the characteristic heights of shrubs. In the details of Figure 6 a good correspondence between the image of the CHM and a high spatial resolution orthophoto in bare areas and in areas with trees and shrub is shown.

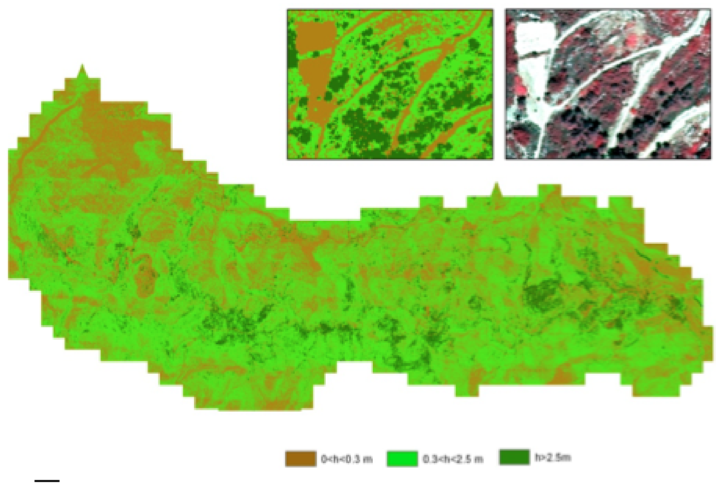


Figure 6: CHM image of the study area. Visualization was performed by grouping the height values in three classes: $0 - 0.3m$, $0.3 - 2.5m$, $h > 2.5m$. This figure also contains a detail of the CHM and a multispectral image. The image is displayed in false colour, IR bands, R and G are assigned to the R, G, B, respectively.

Regarding the estimation of the biomass, a good correlation was obtained using the 95th percentile of the LiDAR height points included in each plot ($R^2 = 0.73$) what indicates that this variable shows a strong explanatory power for predicting shrub biomass (Figure 7).

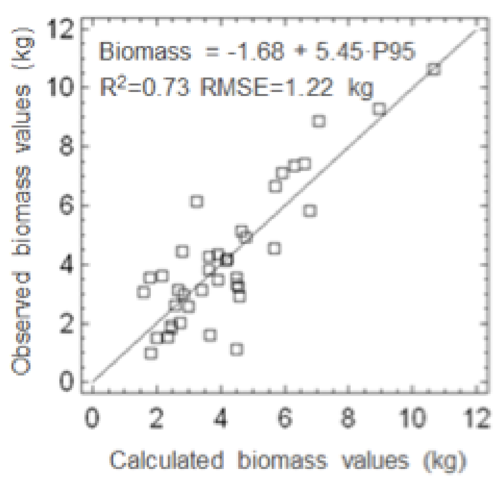


Figure 7: Comparison between observed and calculated biomass values.

4 Conclusions

The results of this study indicate the need to use mathematical modelling with LiDAR data to define the land relief of an area and to obtain variables of shrub vegetation such as the biomass and the cover occupation. This example illustrates the importance of mathematics in such a multidisciplinary task as the environmental analysis.

References

[1] G. Sithole, G. Vosselman. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, **59**(1-2), 85–101 (2004).

[2] K. Zhang, S.-C. Chen, D. Whitman, M-L. Shyu, J. Yan, C. Zhang. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing, **41**(4) 872–882 (2003).

[3] A.M. Felicísimo. Modelos digitales del terreno: principios y aplicaciones en las ciencias ambientales. Oviedo: Ediciones Pentalfa. (1994).

[4] R. Shrestha, W. Carter, K. Slatton, B. Luzum, M. Sartori, M. Airborne Laser Swath Mapping: Quantifying changes in sandy beaches over time scales of weeks to years. ISPRS Journal of Photogrammetry and Remote Sensing, **59**(4) 222–232 (2005).

[5] G. Sohn, I. Dowman. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing, **62**(1), 43–63 (2007).

[6] S.C. Popescu. Estimating biomass of individual pine trees using airborne LiDAR. Biomass & Bioenergy, **31**(9), 646–655 (2007).

[7] M. García, D. Riaño, E. Chuvieco, F.M. Danson. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. Remote Sensing of Environment, **114**(4), 816–830 (2010).

[8] J. Estornell, L. A. Ruiz, B. Velazquez-Marti. Study of shrub cover and height using LiDAR data in a Mediterranean area. Forest Science, **57**(4), 171–179 (2011).

[9] J. Estornell, L. A. Ruiz, B. Velazquez-Marti, A. Fernández-Sarría. A. Estimation of shrub biomass by airborne LiDAR data in small forest stands. Forest Ecology and Management, **262**, 1697–1703 (2011).

[10] E. Næsset. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing, **52**, 49–56 (1997).

[11] Y. Li, H.-E. Andersen, R. McGaughey. A comparison of statistical methods for estimating forest biomass from light detection and ranging data. Western Journal of Applied Forestry, **23**(4), 223–231 (2008).

[12] J. Estornell, L.A. Ruiz, B. Velazquez-Marti, T. Hermosilla, T. Analysis of the factors affecting LiDAR DTM accuracy in a steep shrub area. International Journal of Digital Earth, **4**, 521–538 (2011).

[13] V. J. D. Tsai. Delaunay triangulations in TIN creation: an overview and a linear-time algorithm. International Journal of Geographical Information Systems, **7**(6), 501–524 (1993).