

## Biomass estimation using LiDAR data

Leyre Torre-Tojal\*<sup>1</sup>, Javier Maria Sanchez Espeso<sup>2</sup>, Aitor Bastarrika<sup>1</sup>  
and Jose Manuel Lopez-Guede<sup>3</sup>

<sup>1</sup>Department of Mining and Metallurgical Engineering and Materials Science, Faculty of Engineering, University of the Basque Country UPV/EHU, Nieves Cano 12, 01006 Vitoria-Gasteiz, Spain

<sup>2</sup>Department of Geographic Engineering and Graphic Expression Techniques, School of Civil Engineering, University of Cantabria, Avda. de los Castros, 44, 39005 Santander, Spain

<sup>3</sup>Department of Systems Engineering and Automatic Control, Faculty of Engineering, University of the Basque Country UPV/EHU, Nieves Cano 12, 01006 Vitoria-Gasteiz, Spain

---

### ABSTRACT

Forest ecosystems play a very important role in carbon cycle because they suppose one of the biggest carbon reservoirs and sinks. Estimating the aboveground forest biomass is critical to understand the global carbon storage process. Different models to estimate aboveground biomass in the *Pinus radiata* specie in a specific region of Spain have been developed, using, exclusively, public and accessible data with low point density gathered periodically from Light Detection and Ranging (LiDAR) flights. The point clouds data were processed to obtain metrics considered as predictive variables and afterwards, the multiple regression technique has been applied to generate the biomass estimation models. The best models explain 76% of its variability with a standard error of 0.26 ton/ha in logarithmic units. The methodology can be considered as highly automated and extensible to other territories with similar characteristics. Our results support the use of this approach for more sustainable management of forest areas.

### Keywords:

LiDAR;  
Biomass;  
Multiple linear regression;

URL:  
[dx.doi.org/10.5278/ijsepm.2018.17.7](https://doi.org/10.5278/ijsepm.2018.17.7)

---

### 1. Introduction

The significant increase in concentrations of the greenhouse effect causes additional warming of the surface and atmosphere of the Earth, which can have a negative effect on natural ecosystems and on humanity [1]. The most important greenhouse gas is carbon dioxide (CO<sub>2</sub>), whose atmospheric concentrations have been seriously altered by men in the global carbon cycle [2]. Due to the global efforts to reduce CO<sub>2</sub> emissions, the demand for renewable sources of electricity is fast growing [3]. Concentrations of CO<sub>2</sub> in the atmosphere will continue

rising unless major changes are made in the way fossil fuels are used to provide energy services [4].

Biomass – the fourth largest energy source after coal, oil and natural gas - is the largest and most important renewable energy option at present and can be used to produce different forms of energy. The sustainability potential of global biomass for energy is widely recognized. Energy from biomass currently contributes approximately 10% of global energy supply [5]. About 5% of this energy should cover about 50% of the world's total energy use at present [6].

Another opportunity for emission reductions is the sustainable management of the forests. Deforestation

---

\* Corresponding author e-mail:

produces the biggest effects on the carbon cycle. It amounts to 20% of the total emissions of CO<sub>2</sub> [7], due to the lack of photosynthetic ability of the removed forest vegetation, and the simultaneous liberation of huge amounts of carbon, accumulated in the forest ecosystems for a long time. An important aspect of the carbon cycle is how much CO<sub>2</sub> is retained in the biomass, which will be, later on, exchanged naturally with the atmosphere. Therefore, the interest of estimating forest biomass has increased due to its regulating ability in the cycling of carbon.

Aboveground biomass is defined as the total amount of aboveground living organic matter in trees expressed as oven-dry tons per unit, including leaves, twigs, branches, main bole, and bark [8].

Traditionally, and nowadays, allometric equations have been used to relate the biomass measured in the field with forest variables easily measurable in field (diameter, height) [9, 10]. However, one of the biggest inconveniences of forest inventories is the huge amount of money and time spent in field measurement. To palliate this deficit, Remote Sensing techniques have been developed and successfully applied [11]. As an alternative methodology, airborne LiDAR (Light Detection and Ranging), has become a powerful tool to characterize forest canopy with accuracy and effectiveness.

LiDAR technology constitutes an active remote sensing massive measurement system, based on the emission of a laser that swipes the surface recording the returns from every element impacted. The LiDAR's measurement system is based in the obtained response time for every return. Since the speed of light and the elapsed time are known, the distance to the object that has generated the return can be immediately calculated. Therefore, the X, Y, Z coordinates of every reflected object can be obtained. When the laser is mounted in an aircraft or helicopter, the system is called Airborne LiDAR System (ALS). The positioning and orientation system is a key component of the equipment because it provides the orientation of the laser beam, composed by a differential Global Positioning System (GPS) and an Inertial Movement Unit (IMU). This technology allows gathering huge quantity of measurements with good altimetric accuracies. The accuracy depends on the sensor features, the flight parameters and the absolute orientation system, but typically, it is, approximately, 15 cm [12]. As a result, huge geo-referenced points clouds are recorded, which allows for the characterization of the forest canopy. These sensors provide forest variables related with the canopy, such as mean height, dominant height or mean diameter [13, 14]. After these previous

studies, the prediction of the biomass using LiDAR derived variables has been the logical evolution [15].

New technologies has improved sensor's features, such as the Frequency Repetition Rate (FRR), hence, a higher density point cloud could be managed and other application scopes would be opened. Single tree studies were carried out with good results, the developed models could explain the 87% of the variability of the biomass ( $R^2 = 0.87$ ), using LiDAR dataset with 2–5 pulse/m<sup>2</sup> [16, 17]. Kaartinen & Hyppä [18], applied different methods to single-tree extraction with various points densities. They concluded that no significant increase was appreciated in the accuracy of tree crown delineation while point density was not lower than 2 points/m<sup>2</sup>.

Other authors has used full waveform LiDAR (10 m footprint size) to forest characterization [19] with very good results ( $R^2 = 0.87–0.91$ ), but they pointed that the footprint size was a critical parameter, because precision of the height measures decreases with sizes bigger than 10 m for forest applications. Biomass estimation using large footprint size LiDAR has been applied in different stages such as tropical forest where determination values for the model of 0.94 with and Root Mean Square Error (RMSE) of 16 ton/ha was obtained using multiple linear regression techniques [20]. Poorer results were obtained in temperate forests where the developed model could explain only the 0.66 of the variability of the model with an RMSE of 41.989 ton/ha. The discrepancies were attributed to the high amount of biomass concentrated in the study area (New England), that could create saturation problems in height-biomass relationship when biomass was bigger than 300 ton/ha [21].

Newer methodologies based on machine learning use Random Forest or Supported Vector Machines methods to estimate forest structural variables, including biomass [22, 23]. In the first study the determination coefficient was low ( $R^2 = 0.5$ ), even though better results were obtained for other variables like Lorey's height, defined as the weighted mean height whereby individual trees are weighted in proportion to their cross-sectional area at 1.5 m off the ground ( $R^2 = 0.89$ ). In the second study, a comparison between machine learning methods and linear regression technique was made and they concluded that Support Vector Regression technique achieved more accurate adjustments, while linear regression got the worst results.

At the global scale, Geoscience Laser Altimetry System (GLAS) flown onboard the IceSat platform (which operated from 2003 through 2008), is the only lidar sensor onboard a satellite platform available for terrestrial ecology applications. As Zolkos et al.

concluded comparing different lidar types, biomass models based on GLAS metrics had the lowest absolute errors of all lidar types (RMSE = 40 Mg ha<sup>-1</sup>). This can be attributed to several factors: (i) most GLAS studies evaluated boreal forests, which had considerably lower biomass than other forests types, and (ii) the larger size of GLAS footprints reduces variance in forest structural measurements via spatial averaging. Moreover, GLAS was not designed for vegetation studies and has relatively large canopy height errors that partly account for its reduced utility for biomass estimation [24].

The main aim of this research is to develop a highly automatized methodology to estimate the aerial biomass of the *Pinus Radiata* specie in the Arratia-Nervi3n using low density LiDAR data (0.5 pulse/m<sup>2</sup>) and totally free software and tools. The LiDAR data are obtained from the National Plan of Aerial Orthophotography (PNOA), these data are fully public and they are supposed to be updated every 3 or 4 years, representing a stable, free and updated data source for the entire national extension. The applied methodology allows the actualization of the results of the estimated biomass when new data may be available for the entire territory, with a value of R = 0.76 and an RMSE = 0.26 in logarithmic units. The accurate estimation of the forest biomass is essential to conduct suitable management policies regarding this renewable energy source.

The structure of the paper is as follows: Firstly, the used materials are described. The study area and the both main data sources: Forest National Inventory 4 (FNI4) and the LiDAR data, are analyzed. Secondly, the applied methodologies are explained. The forest and LiDAR data processing is detailed. In the third section the obtained results are shown, while in the fourth section these results are compared with other similar studies and the results are analysed in depth. Finally, the main conclusions of the study are exposed.

## 2. Materials

### 2.1. Study area

The research has been developed in the Arratia-Nervi3n region, placed in the Historical Territory of Biscay (Figure 1). This region is integrated by 14 municipalities, with an area of 40.000 ha. The altitude varies from 55 to 1480 m, with a mean slope of 19°. The main specie is the *Pinus Radiata* D. Don., which represents 60% of the total species in the study area.

### 2.2. National forest inventory

The dendometric data for the *Pinus Radiata* species were collected from the Fourth Forest Inventory in the Basque Autonomous Community during the summer of 2011. The sample plots were distributed along the territory coinciding spatially with the kilometeric grid of the

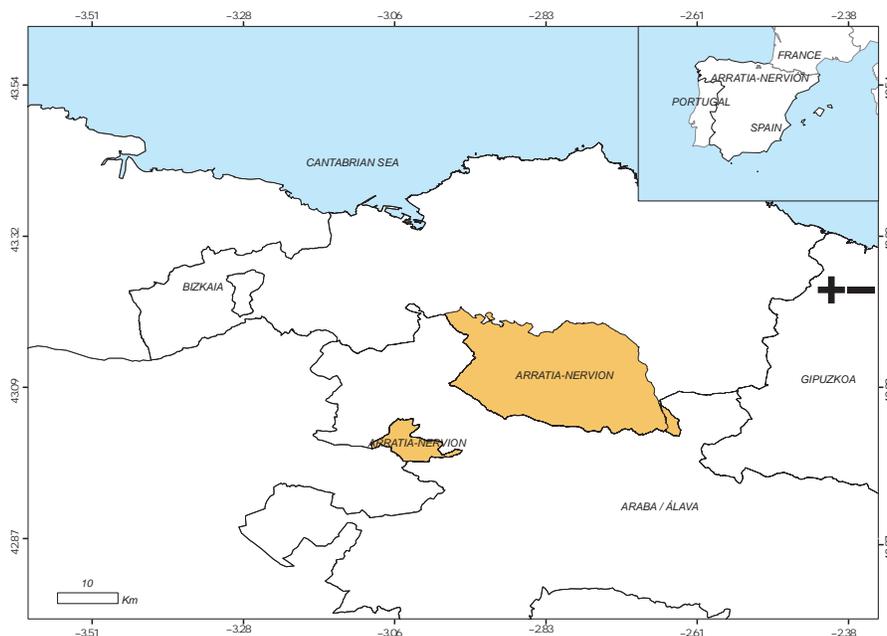


Figure 1: Location of the study area, Arratia-Nervi3n

UTM (Universal Transversal Mercator) for the 30N time zone cartography in tree-covered areas, with a mean density of a plot per square kilometre, in the ED50 (European Datum 1950) reference system. The circular plots of maximum radius 25 m had an area of 0.2 ha.

### 2.3. LiDAR data

The LiDAR data were obtained from the LiDAR flight of the Autonomous Community of the Basque Country, carried out in the summer of 2012 by the Basque Government. The LiDAR data were projected using the ETRS89 (European Terrestrial Reference System 1989) reference system and the applied cartographic projection was the UTM. The ellipsoidal heights were referred to the ETRS89 system, using GRS80 ellipsoid as reference surface. The LiDAR data can be found at [ftp://ftp.geo.euskadi.eus/lidar/LIDAR\\_2012\\_ETRS89/LAS/](ftp://ftp.geo.euskadi.eus/lidar/LIDAR_2012_ETRS89/LAS/).

The beam divergence ( $\Delta$ ) is the angle of the laser beam from its emission until the impacts against the surface. The recommended values for the beam divergence in forest scope are 0.1–0.5 mrad. The beam divergence has direct effect in the footprint size. With a flight height of 1100 m and a beam divergence of 0.5 mrad [12], the obtained footprint size is 0.55 m, considered as small footprint. The ability of the small footprint LiDAR to obtain single tree metrics and accurate DTMs, has become established as the most suitable one for forest applications [25].

The scan angle is transversal to the flight direction, it determines the size of the area swiped by the laser. An appropriate value for the scan angle in forest applications would be  $12^\circ$  [26]. Even if this parameter has not influence in DEM or CHM creation, in very dense forests, other trees often shade the side area of a tree crown that could have been exposed at a high scanning angle causing a shift of the laser height percentiles. However, the more affected parameter by scanning angle would be the proportion of canopy [27].

**Table 1: LiDAR flight technical parameters**

LiDAR flight technical parameters	
Scan angle	60°
Pulse Repetition Rate (PRR)	100 kHz
Laser beam divergence	< 0.5 mrad
Flying mean speed	185 km/h
Point density	0.5 points/ m <sup>2</sup>
Flying Mean Height	1100 m
Overlap	15%

The Pulse Repetition Rate (PRR) is the number of pulses emitted, normally, per second. The most used value for PRR fluctuates between 100 and 200 KHz. Csanyi [28] showed that the mean error in the determination of the height increases when the frequency increases, because the energy emitted by the pulse is less and this circumstance could cause certain degradation in the distance measurement accuracy.

Point density has not been a determinant factor in the estimation of dendrometric variables such as tree height, diameter at breast height or total basal area until it went down to 1 pulse/m<sup>2</sup> in coniferous forest [29]. Treitz et al. [30] did not find relevant differences when they reduced the density from 3.5 to 0.5. In contrast, other authors pointed that both the standard deviation and the bias of the LiDAR data, decreased when point density increased, but not following a linear relationship [31, 32].

The flight altitude is influential in the underestimation of the tree height, being bigger when the height increases. Yu et al. and Nilsson [33, 34] pointed that with flight heights higher than 1500 m, point density decreased significantly. The random errors of the DTM also increase when the flight height grows, mainly due to the decrease of the point density and the increase of the planimetric error, especially in flat areas [35].

To cover all the study area, one hundred and twenty three. LAS files were processed, with an area per file of  $2 \times 2$  km<sup>2</sup>, and approximately 2 million points per file. The LiDAR data has been analysed previously to the processing in order to contextualize them correctly. Table 2 shows the spatial domain of the LiDAR data, Table 3 represents the distribution of the cloud point by the return number, and finally, Table 4 analyses the classification of the LiDAR returns by the standard given by the American Society for Photogrammetry and Remote Sensing (ASPRS).

The mean point density of LiDAR cloud was 0.78 points/m<sup>2</sup>, as can be seen in Figure 2 the density was irregular. Is important to remark that other LiDAR flights in similar studies has considered bigger point densities (> 0.5 points/m<sup>2</sup>) for the stand approach.

**Table 2: Spatial domain of the LiDAR data**

Spatial Domain	Minimum	Maximum
X Coordinate (m)	492,000	530,000
Y Coordinate (m)	4,758,000	4,784,000
Altitude (m)	45.42	1482.66

If we map point density along all the study area considering all the returns (Figure 3), disseminate areas with less than 0.4 point/m<sup>2</sup> appeared without a valid justification. However, in the rest of the area densities were bigger than the nominal ones.

**Table 3: Spatial domain of the LiDAR data**

Return number	Point number	Percentage
Return 1	381,962,402	91.7%
Return 2	32,079,525	7.7%
Return 3	2,123,779	0.5%
Return 4	198,003	0.0%
Return 5	23,454	0.0%
Return 6	0	0.0%
Return 7+	0	0.0%
Total Returns	416,387,163	

**Table 4. LiDAR point classification according to ASPRS standard**

Category	Points number	Interpretation
0	0	Created, never classified
1	26,348	Unclassified
2	144,894	Ground
3	48,201,484	Low Vegetation
4	17,611,152	Medium Vegetation
5	176,795,016	High Vegetation
6	1,897,663	Building
7	143,143	Low Point (noise)
10	495,649	Reserved

**2.4. Orthophotos**

Orthophotos from the flight made by the Basque Government during the summer of 2012, with a 25 m pixel size, were used to localize possible error and discrepancies. The reference system is the European Terrestrial Reference System 89 (ETRS89) and the coordinate system is UTM for the 30N time zone. The orthophotos are available at [ftp://ftp.geo.euskadi.eus/cartografia/ Cartografia\\_Basica/Ortofotos/ORTO\\_2012/ Hojas\\_JPG/5000/](ftp://ftp.geo.euskadi.eus/cartografia/Cartografia_Basica/Ortofotos/ORTO_2012/Hojas_JPG/5000/).

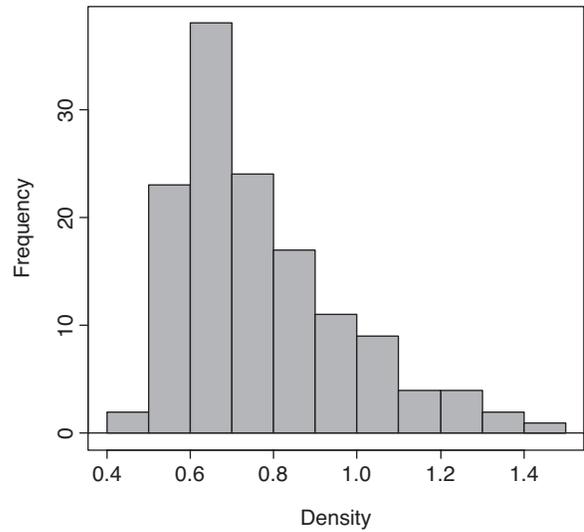


Figure 2: Distribution of the point density per m<sup>2</sup>

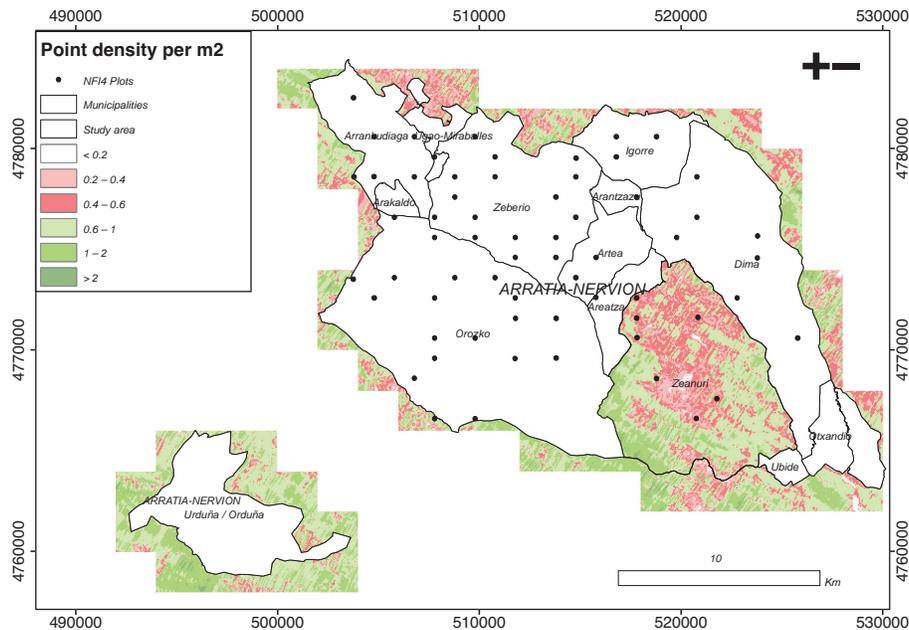


Figure 3: Point density per 50 m diameter plot taking into account all the returns

### 3. Methodology

#### 3.1. Biomass estimation

From the 118 plots placed in the area of interest, 63 were chosen for the sample, having more than 80% of occupation of *Pinus Radiata* (Figure 4). Allometric equations were applied to estimate the biomass for every tree [36], as nested plot methodology [37] had been applied to every plot, final per hectare values were interpolated using diameter dependent expansion factors.

In order to check if the obtained control values could be clustered or if they follow a random spatial pattern, spatial autocorrelation Moran's I index [38, 39] was applied to the following variables: biomass, tree height and number of trees. This index varies between -1 and 1, corresponding to values near to 1 or -1, bigger spatial autocorrelation.

If we compare the figures, neither biomass (Figure 5) or number of trees (Figure 6) seem to follow a clustering pattern, being it completely random, but certain

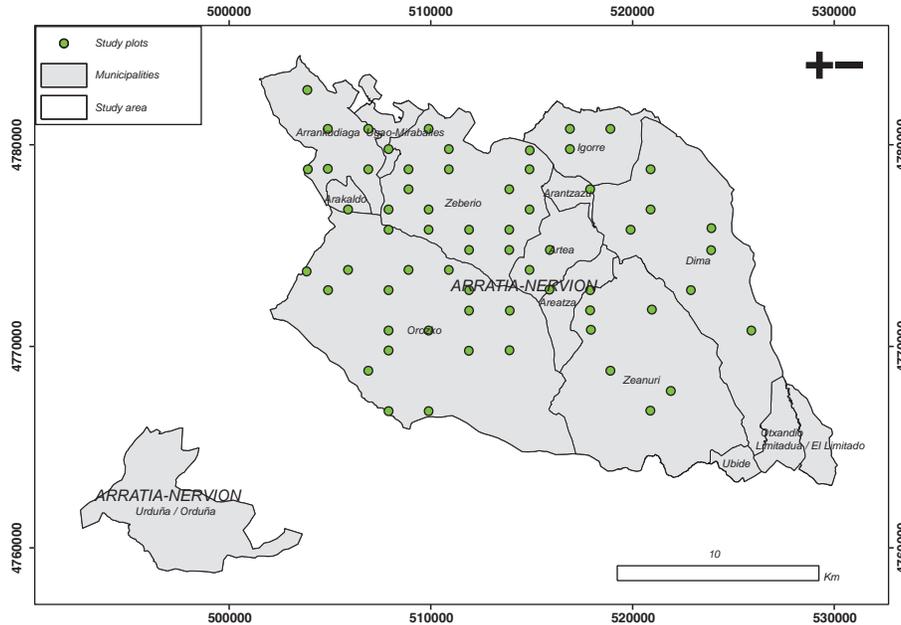


Figure 4: Distribution of the sample plots in ETRS89 reference system.

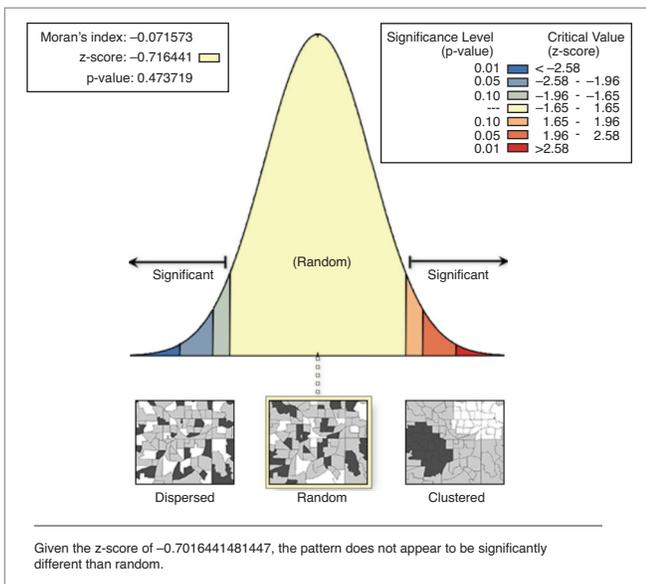


Figure 5: Moran's I index applied to biomass

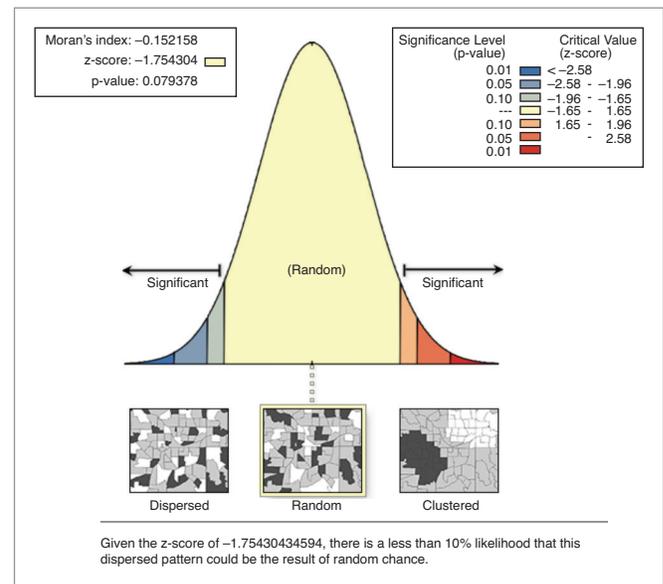


Figure 6: Moran's I index applied to tree height

clustering can be appreciated in the tree height (Figure 7). This pattern is disperse for a probability of 0.9.

### 3.2. LiDAR data processing

Once the reference values of the biomass were calculated, LiDAR data were processed to obtain various products. Digital Terrain Models (DTM) and Digital Surface Models (DSM) for the area were produced. The DTM is a continuous function which represents the elevation of every planimetric position,  $z = f(x,y)$ , while the DSM includes only the points that are in the highest part of the area. The Canopy Height

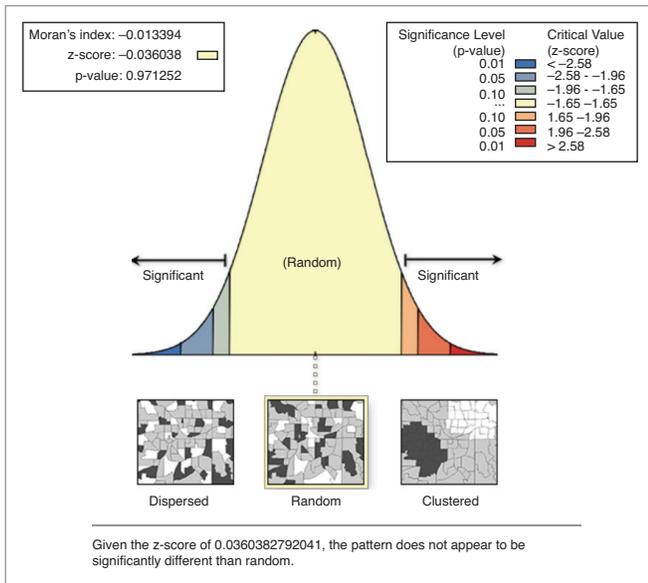


Figure 7: Moran's I index applied to the number of trees

Model (CHM) was later computed by the subtraction of the previous two models (Figure 8).

On one hand, the conventional metrics given by commercial software were calculated. On the other hand, an algorithm was designed to obtain specific density metrics from the cloud. This algorithm divides the entire cloud into equal height slides, in order to compute metrics related with the point density of each slide [40]. For this purpose the LiDAR dataset was integrated in a database, so the point clouds and the MDT generated before, using routines developed with PostGis spatial database extender for the PostgreSQL Database Management System. Once different low limits and upper limits were tested, the algorithm developed to calculate the canopy density metrics divided the point clouds into 10 vertical layers of equal height, setting the low limit on 2 m, to avoid shrubs and the upper limit as the percentile 95 of the height distributions. Then, the routine calculates the proportion of points of the total number of points contained above each layer, finally, 10 canopy densities were computed (tr\_1...tr\_10) (Table 5).

### 3.3. Models generation

Firstly, the correlation matrix between calculated field biomass and obtained LiDAR metrics was obtained to check which variables were more correlated [41]. It was decided to use the logarithm because of possible homocedasticity problems in the model [42]. Secondly, these linear relationships were verified graphically, creating dispersion diagrams between the measured field biomass and its logarithm and each obtained LiDAR metric.

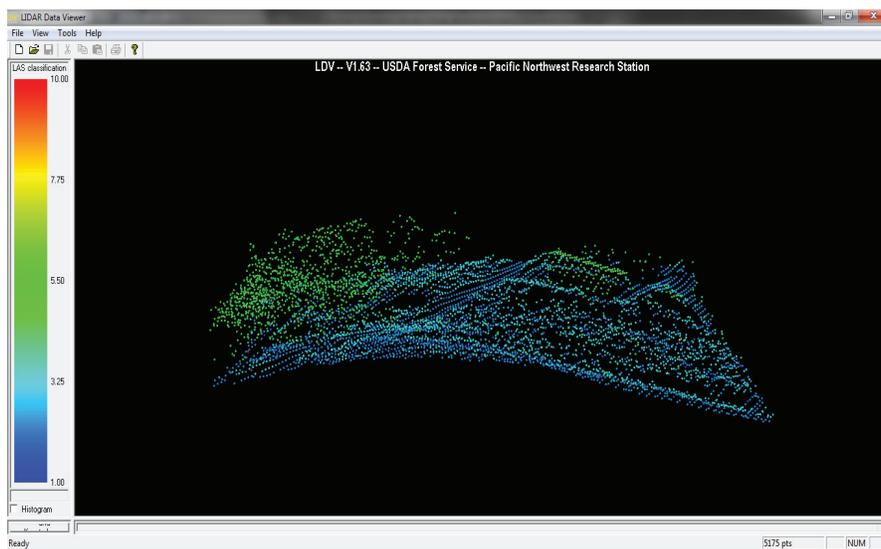


Figure 8: Normalized LiDAR data

Table 5: Obtained LiDAR explanatory variables

Variable	Description	Variable	Description
count	Number of returns above the minimum height	ccr	canopy relief ratio:((mean - min) / (max - min))
densitytotal	total returns used for calculating cover	eqm	Elev quadratic mean
densityabove	returnsaboveheightbreak	ecm	Elev cubic mean
densitycell	Density of returns used for calculating cover	r1count,...,r9count	Count of return 1,...,9 points above the minimum height
min	minimumvalueforcell	rothercount	Count of other returns above the minimum height
max	maximumvalueforcell	allcover	(all returns above cover ht) / (total returns)
mean	mean valueforcell	afcover	(all returns above cover ht) / (total first returns)
mode	modal valueforcell	allcount	number of returns above cover ht
stddev	standard deviation of cell values	allabovemean	(all returns above mean ht) / (total returns)
variance	variance of cellvalues	allabovemode	(all returns above ht mode) / (total returns)
cv	coefficient of variation for cell	afabovemean	(all returns above mean ht) / (total first returns)
cover	cover estimate for cell	afabovemode	(all returns above ht mode) / (total first returns)
abovemean	proportion of first (or all) returns above the mean	fcountmean	number of first returns above mean ht
abovemode	proportion of first (or all) returns above the mode	fcountmode	number of first returns above ht mode
skewness	skewness computed for cell	allcountmean	number of returns above mean ht
kurtosis	kurtosis computed for cell	allcountmode	number of returns above ht mode
AAD	average absolute deviation from mean for the cell	totalfirst	total number of 1st returns
p01,...,p99	1st,...,99th percentile value for cell	totalall	total number of returns
iq	75th percentile minus 25th percentile for cell	tr1,tr2,...,tr10	point density metrics

Then, the models were generated using combinations of two and three variables. The main criteria to choose the appropriate model was the determination coefficient ( $R^2$ ), but several statistical tests were carried out to evaluate the level of fulfilment of the hypothesis of the linear regression technique:

- a. Shapiro-Wilkinson normality test: This test measures the disagreement between the residuals and the normality probabilistic line. The null hypothesis assumes that the data are normally distributed [43].
- b. Breusch-Pagan homoscedasticity test: In order to assure the homoscedasticity of the residuals, this test analyzes if the variance of the residuals is constant. The null hypothesis assumes that the data are homoscedastic.
- c. Durbin-Watson autocorrelation test: This test checks the autocorrelation level of the regression to cast aside dependencies in the residuals. The null hypothesis assumes that the autocorrelation is null.
- d. Variance Inflation Factor (VIF) multi-collinearity test: When a variable included in a model is linear combination of other variables of the model, multi-collinearity problems can occur. The ideal value of the VIF is one, representing total absence of multi-collinearity, when the value of the factor is bigger than 10, the multi-collinearity is high [44].
- e. Regression Specification Error Test (RESET): This test is designed to detect if there are any neglected nonlinearities in the model. The null hypothesis states that the functional form is correctly specified.
- f. Bonferroni outlier detector test: This test verifies the absence of outliers. The null hypothesis assumes the absence of atypical values in the plot.

All the tests were performance with a significance level of 0.95, except for the VIF test.

### 4. Results

As can be seen in Table 6, for the obtained models the best results involve high percentiles of the LiDAR heights and point density metrics. The values for the applied tests show that almost all the models satisfy the hypothesis of the multiple linear regression technique. The relation between the logarithm of the biomass and the percentiles of the height is clearly linear, although the dispersion increases when the biomass values increase (Figure 9). The density metric shows higher dispersion values, as it was verified because of their signification in the models, where the height percentile could explain most of the variability.

The positioning of the sample plots was made with a GPS navigator, obtaining autonomous observations without differential correction. To evaluate the incidence of this possible positioning error, each plot was moved 10 m in the directions of the wind rose, obtaining eight new parcels around the original one. The similarity between the nine samples was measured by the Cohen-Kappa concordance test. The test revealed that the minimum value for Kappa coefficient was 0.91.

### 5. Discussion

All the models in Table 3 obtain similar values for the determination coefficient ( $R^2 = 0.75$ ), and identical Root Mean Square Error value ( $RMSE = 0.26$  in logarithmic units). The main differences can be observed in the results of the applied statistical tests. In most cases, the models satisfy the hypothesis of the linear regression technique with a significance level of 0.95. No homoscedasticity problems have been detected in the models and no other tendencies apart from the linear relationship between the variables can be considered. In addition, in Figure 10 the linear relationship can be graphically seen, even if it is much stronger between the biomass and the height percentile, than between the density metric and the biomass. In all the cases, the most correlated combinations involved a height metric and a density one. Combinations of three variables did not improve the results so much ( $R^2 = 0.78$ ) and the third variable was not statistically significant in the models, being the models less harmonious.

The results obtained here are comparable to those from similar studies using LiDAR technology to

**Table 6: Values obtained for the different models, p-values are shown for all the test except VIF**

Variables	R <sup>2</sup>	R <sup>2</sup> adjus	SE	Shapiro-Wilk	Breustch-Pagan	Durbin-Watson	VIF	RESET	Bonferroni
p99-allabovemean	0.77	0.76	0.26	0.10	0.12	0.54	1.28	0.62	0.04
p95-allabovemean	0.77	0.76	0.26	0.07	0.23	0.55	1.27	0.64	0.06
p95-tr_3	0.76	0.75	0.26	0.30	0.24	0.54	1.11	0.25	0.05
count-p95	0.76	0.75	0.26	0.13	0.19	0.69	1.10	0.96	0.05
p95-allcover	0.76	0.75	0.26	0.18	0.20	0.61	1.12	0.49	0.05
p95-tr_2	0.76	0.75	0.26	0.10	0.23	0.58	1.10	0.28	0.06
p95-tr_1	0.76	0.75	0.26	0.18	0.20	0.60	1.11	0.51	0.05
p95-tr_4	0.76	0.75	0.26	0.08	0.23	0.52	1.14	0.14	0.05
p90-allabovemean	0.76	0.75	0.26	0.03	0.29	0.57	1.27	0.56	0.05

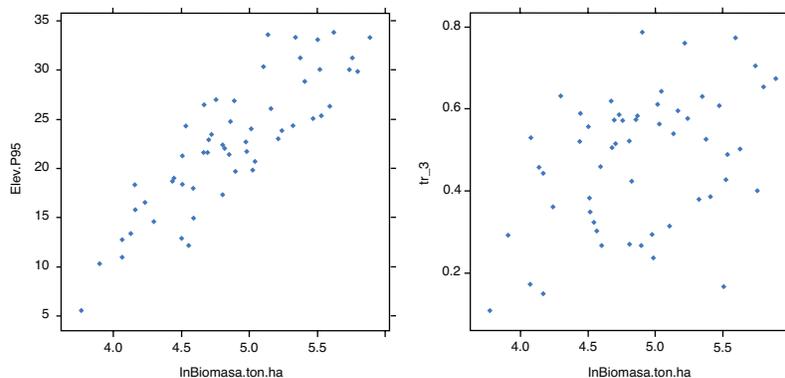


Figure 9: Dispersion diagrams between the biomass and the predictive variables for one of the obtained models

characterise aboveground biomass in coniferous forest with a mean point density of 1.2 points/m<sup>2</sup> [45, 46]. Naeset and Gobakken [47] obtained models where the predictive variables were too a height percentile and a density metric but the highest determination coefficient was 0.86 and a RMSE of 0.25 ton/ha, using four different stratifications.

The positioning error of the plots was not influential (minimum value for Kappa coefficient = 0.91), therefore, the forest canopy can be assumed as homogenous in the study sites.

On one hand, the rough surface of the study area and the big slopes can be influential in the DTM creation, increasing the error when the slope increases. On the other hand, the scan angle applied in the flight used in this study (up to 25°) is much bigger than the recommended one (15°) [48], assuming its influence in the LiDAR metrics and, therefore, in the obtained results.

Thus, the average estimated stem volume growth calculated by means of stand volume growth equations [49] is about 3%. This method involves the selection of one tree whose volume is equal to the mean tree volume for the stand. The volume of the selected tree is then determined and the stand volume estimated by inflating the sample tree volume by the number of trees in the stand. It is therefore likely that a time lag of maximum one growth season has had a very limited impact on the results.

## 6. Conclusions

The developed methodology has been able to obtain models that explain approximately 75% of the variance of the field biomass, with a standard error of 0.26 logarithmic units. This fact reveals that the biomass can be modelled using two independent variables from LiDAR data: a height percentile and a canopy density metric. The results obtained here are comparable to those from similar studies using LiDAR technology to characterise aboveground biomass in coniferous forest, even using highest pulse density [50].

LiDAR technology has been revealed as a powerful tool to characterize the forest canopy in vast areas. Biomass values have been obtained due to the strong relations between this variable and the height of the trees collected by the sensor. The spatial and temporal continuity of the LiDAR data included in the PNOA, which are completely public, enables their application

to national scale and guarantees the durability of the data. National Forest Inventory field plots are already available in many countries and consequently reliable carbon monitoring at national level using LiDAR data is plausible.

The uncertainty in the biomass estimation is very high due to the different error sources implied, such as the forest inventory, the applied allometric equations and finally the ones derived from the LiDAR data itself, but the technique has shown good adjustments between the incoming variable and the different LiDAR metrics.

Further research should be carried out to confirm the validity of the results in different areas and with different species. The relationships between ALS data and biomass will change with site properties, latitude, and altitude and these factors may influence stem form, being this part of the tree the largest biomass component [51].

Numerous studies, debates, and public figures have highlighted the need for a radical change in the very near future, to reduce the environmental damage and risks associated with the existing energy system [52]. Biomass estimation plays a very important role in the future possible policies applied to mitigate the current environmental problems derived from the fossil fuels used to provide energy services.

## References

- [1] Sands P, United Nations The United Nations Framework Convention on Climate Change, Review of European Community & International Environmental Law 1 (3) (1992) pages 270–277. <http://dx.doi.org/10.1111/j.1467-9388.1992.tb00046.x>
- [2] Intergovernmental Panel of Climate Change. Climate change 2001, Synthesis Report. Robert T. Watson, Intergovernmental Panel on Climate Change, 2002.
- [3] Korfiati A, Gkonos C, Veronesi F, Gaki A, Grassi S, Schenkel R, Volkwein S, Raubal M, Lorenz Hurni. Estimation of the Global Solar Energy Potential and Photovoltaic Cost with the use of Open Data. International Journal of Sustainable Energy Planning and Management 9 (2016) pages 17–30. <http://dx.doi.org/10.5278/ijsepm.2016.9.3>
- [4] Berndesa G, Hoogwijkb M, van den Broek R. The contribution of biomass in the future global energy supply: a review of 17 studies. Biomass and Bioenergy 25 (2003) pages 1–28. [http://doi:10.1016/S0961-9534\(02\)00185-X](http://doi:10.1016/S0961-9534(02)00185-X)
- [5] Welfle A. Balancing growing global bioenergy resource demands – Brazil's biomass potential and the availability of

- resource for trade. *Biomass and Bioenergy* 105 (2017) pages 83–95. <http://dx.doi.org/10.1016/j.biombioe.2017.06.011>
- [6] Ladanaï S, Vinterback J. *Global Potential of Sustainable Biomass for Energy*, Uppsala, 2009.
- [7] Montero G, Ruiz-Peinado R, Muñoz M. *Producción de biomasa y fijación de CO<sub>2</sub> por los bosques españoles*. Monografías INIA: Serie Forestal 13, 2005.
- [8] Brown S, *Estimating biomass and biomass change of tropical forests: A primer*, FAO Forestry Paper 134. Rome, FAO, United Nation, 1997.
- [9] Eamus D, Burrows W, McGuinness K, *Review of allometric relationships for estimating woody biomass for Queensland, the northern territory and western Australia*. Australia: Australian Greenhouse Office, 2000.
- [10] Zianis D and Seura SM, *Biomass and stem volume equations for tree species in Europe*, Finnish Society of Forest Science, Finnish Forest Research Institute, 2005.
- [11] Johansen K, Phinn S, Witte C, *Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs*, *Remote Sensing of Environment* 114 (11) (2010) pages 2679–2691. <https://doi.org/10.1016/j.rse.2010.06.004>
- [12] Baltsavias EP, *Airborne laser scanning: Basic relations and formulas*, *ISPRS Journal of Photogrammetry & Remote Sensing* 54 (1999) pages 199–214. [https://doi.org/10.1016/S0924-2716\(99\)00015-5](https://doi.org/10.1016/S0924-2716(99)00015-5)
- [13] Nelson R, Krabill W, MacLean G, *Determining forest canopy characteristics using airborne laser data*, *Remote Sensing of Environment* 15 (3) (1984) pages 201–12. [https://doi.org/10.1016/0034-4257\(84\)90031-2](https://doi.org/10.1016/0034-4257(84)90031-2)
- [14] Nilsson M, *Estimation of tree heights and stand volume using an airborne lidar system*, *Remote Sensing of Environment* 56 (1) (1996) pages 1–7. [https://doi.org/10.1016/0034-4257\(95\)00224-3](https://doi.org/10.1016/0034-4257(95)00224-3)
- [15] Condes S and Riaño D, *El uso del escáner laser aerotransportado para la estimación de la biomasa foliar del pinus sylvestris L. en Canencia (Madrid)*, *Cuadernos De La Sociedad Española De Ciencias Forestales* 19 (2005) pages 63–70.
- [16] Hyppä J and Inkinen M, *Detecting and estimating attributes of single tree using LiDAR*, *The Photogrammetric Journal of Finland*, 16 (2) (1999) pages 27–42.
- [17] Popescu SC, *Estimating biomass of individual pine trees using airborne lidar*, *Biomass Bioenergy* 31 (9) (2007) pages 646–55. <https://doi.org/10.1016/j.biombioe.2007.06.022>
- [18] Kaartinen H and Hyppä J, 2008. EuroSDR/ISPRS project, commission II "tree extraction" final report, euro SDR official publication. Dublin, Ireland: European Spatial Data Research; 2008.
- [19] Lefsky MA, Cohen WB, Acker SA, Parker GG, Spies TA, Harding D, *Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests*, *Remote Sensing of Environment* 70 (3) (1999) pages 339–61. [https://doi.org/10.1016/S0034-4257\(99\)00052-8](https://doi.org/10.1016/S0034-4257(99)00052-8)
- [20] Drake JB, Dubayah RO, Knox RG, Clark DB, Blair JB, *Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest*, *Remote Sensing of Environment* 81 (2–3) (2002) pages 378–92. [https://doi.org/10.1016/S0034-4257\(02\)00013-5](https://doi.org/10.1016/S0034-4257(02)00013-5)
- [21] Ahmed R, Siqueira P, Hensley S, *A study of forest biomass estimates from lidar in the northern temperate forests of New England*, *Remote Sensing of Environment* 130 (2013) pages 121–35. <https://doi.org/10.1016/j.rse.2012.11.015>
- [22] Zhao K, Popescu S, Meng X, Pang Y, Agca M, *Characterizing forest canopy structure with lidar composite metrics and machine learning*, *Remote Sensing of Environment* 115 (8) (2011) pages 1978–96. <https://doi.org/10.1016/j.rse.2011.04.001>
- [23] Gleason CJ and Im J. *Forest biomass estimation from airborne LiDAR data using machine learning approaches*. *Remote Sens Environ* 125 (2012) pages 80–91. <https://doi.org/10.1016/j.rse.2012.07.006>
- [24] Zolkos SG, Goetz SJ, Dubayah R. *A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing*. *Remote Sensing of Environment* 128 (2013) pages 289–98. <http://dx.doi.org/10.1016/j.rse.2012.10.017>
- [25] Evans DL, Roberts SD, Parker RC, *LiDAR A new tool for forest measurements*, *The Forestry Chronicle* 82 (2) (2012) pages 211–218. <https://doi.org/10.5558/tfc82211-2>
- [26] White JC, Wulder MA, Varhola A, Vastaranta M, Coops NC, Cook BD, Pitt D, Woods M, *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach (version 2.0)*, 2013)
- [27] Holmgren J, Nilsson M, Olsson H, *Simulating the effects of lidar scanning angle for estimation of mean tree height and canopy closure*, *Canadian Journal of Remote Sensing* 29 (5) (2003) pages 623–632. <http://dx.doi.org/10.5589/m03-030>
- [28] Csanyi N and Toth CK, *LiDAR data accuracy: The impact of pulse repetition rate*. MAPPs/ASPRS 2006 Fall Conference, 2006).
- [29] Jakubowski MK, Guo Q, Kelly M, *Tradeoffs between lidar pulse density and forest measurement accuracy*, *Remote Sensing of Environment* 130 (2013) pages 245–53. <http://dx.doi.org/10.1016/j.rse.2012.11.024>
- [30] Treitz P, Lim K. W, M., Pitt D, Nesbitt D, Etheridge D, *LiDAR sampling density for forest resource inventories in Ontario, Canada* *Remote Sensing* 4 (2012) pages 830–848. [10.3390/rs4040830](https://doi.org/10.3390/rs4040830)

- [31] Anderson ES, Thompson JA, Crouse DA, Austin RE, Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM, *Geoderma* 132 (3-4) (2006) pages 406–415. <https://doi.org/10.1016/j.geoderma.2005.06.004>
- [32] Olsen RC, Puetz AM, Anderson B, Effects of lidar point density on bare earth extraction and DEM creation, *American Society for Photogrammetry and Remote Sensing Baltimore, Maryland* (2009).
- [33] Yu X, Hyypä J, Hyypä H, Maltamo M, 2004. Effects of flight altitude on tree height estimation using Airborne laser scanning, *International Archives of the International Society for Photogrammetry and Remote Sensing* 36 (8/W2) (2004) pages 96–101.
- [34] Nilsson M, Estimation of tree heights and stand volume using an airborne lidar system, *Remote Sensing of Environment* 56 (1) (1996) pages 1–7.
- [35] Hyypä H, Yu X, Hyypä J, Kaartinen H, Kaasalainen S, Honkavaara E, Ronnholm P, Factors affecting the quality of DTM generation in forested areas, *ISPRS WG III/3, III/4, V/3 Workshop "Laser Scanning 2005"*, Enschede, the Netherlands (2005) pages 85–90.
- [36] Canga E, Dieguez-Aranda I, Afif-Khoury E, Camara-Obregon A, Above-ground biomass equations for pinus radiata D. don in Asturias, *Forest Systems (INIA)* 22 (3) (2013) pages 408–415. <http://dx.doi.org/10.5424/fs/2013223-04143>
- [37] Segundo inventario forestal nacional. Explicaciones y métodos. 1986–1995. I.C.O.N.A.; 1990.
- [38] Goodchild MF. Spatial autocorrelation. *Catmog* 47, Geo Books, 1986.
- [39] Griffith D. Spatial autocorrelation: A primer. *Resource Publications in Geography*, Association of American Geographers, 1987.
- [40] Nasset E and Gobakken T, Estimating forest growth using canopy metrics derived from airborne laser scanner data, *Remote Sensing of Environment* 96 (3–4) (2005) pages 53–65. doi:10.1016/j.rse.2005.04.001
- [41] Garcia M, Riaño D, Chuvieco E, Danson FM, Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data, *Remote Sensing of Environment* 114 (4) (2010) pages 816–30. <https://doi.org/10.1016/j.rse.2009.11.021>
- [42] Valbuena MA, Determinación de variables forestales de masa y de árboles individuales mediante delineación de copas a partir de datos LiDAR aerotransportado, 2014
- [43] Shapiro SS and Wilk MB. 1965. An analysis of variance test for normality (complete samples) *Biometrika* 52 (3/4) pages 591–611.
- [44] Kleinbaum D, Kupper L, Nizam A, Rosenberg E. *Applied regression analysis and other multivariable methods*. 3rd ed. Cengage Learning, 2014.
- [45] Nasset E and Gobakken T, Estimating forest growth using canopy metrics derived from airborne laser scanner data, *Remote Sensing of Environment* 96 (3–4) (2005) pages 53–65. doi:10.1016/j.rse.2005.04.001
- [46] Hall SA, Burke IC, Box DO, Kaufmann MR, Stoker JM. Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests, *Forest Ecology and Management* 208 (1–3) (2005) pages 189–209. doi:10.1016/j.foreco.2004.12.001
- [47] Nasset E and Gobakken T, Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser, *Remote Sensing of Environment* 112 (6) (2008) pages 3079–90. <https://doi.org/10.1016/j.rse.2008.03.004>
- [48] Su J and Bork E, Influence of vegetation, slope and lidar sampling angle on DEM accuracy, *Photogrammetric Engineering & Remote Sensing* 72 (2006) pages 1265–1274.
- [49] Delbeck K. Methods for calculation of increment for open forests, *Tidsskrift for Skogbruk*, 73 (1965) pages 5–45.
- [50] Gonzalez-Ferreiro E, Aranda U, Miranda D. Estimation of stand variables in pinus radiata D. don plantations using different LiDAR pulse densities, *Forestry* 85 (2) (2012) pages 281–292. <https://doi.org/10.1093/forestry/cps002>
- [51] Marklund LG, Biomass functions for pine, spruce and birch in Sweden, *Univ. of Agric. Sciences, Dep. of For. Surv. Report*, 45, page 73 (1988).
- [52] Connolly D, Vad Mathiesen B. A technical and economic analysis of one potential pathway to a 100% renewable energy system. *International Journal of Sustainable Energy Planning and Management*. 1 (2014) pages 7–28. <http://dx.doi.org/10.5278/ijsepm.2014.1.2>