MASTER ON GEOMATICS, REMOTE SENSING AND SPATIAL MODELS APPLIED TO FOREST MANAGEMENT

Site index estimation using airborne laser data in *Eucalyptus dunnii* Maide stands in Uruguay

Author:

Iván Gabriel Rizzo Martín

Directors:

Andrés Eduardo Hirigoyen Domínguez

Rafael María Navarro Cerrillo

2021

General index **Page**

Index of figures **Page**

Index of table **Page**

Summary

The estimation of forest variables to support a forest inventory can be approached through the use of different technologies. Although field sampling is the most widely implemented technique, the development of remote sensing techniques increases the possibilities of action in this field. One of these technologies is the airborne LiDAR scanner (ALS). In this study, linear models and non-parametric models with Random Forest imputation were generated to estimate the total height (HT) and site index (SI) of *Eucalyptus dunnii* Maide, based on LiDAR metrics. High spatial resolution continuous rasters for HT and SI were created with these models. The use of a semi-automatic object-oriented segmentation algorithm for stand delimitation based on the SI raster was then carried out. To evaluate the performance of these models, the One Leave One cross-validation technique was implemented, determining for each model the ratio between the RMSE of the model and the RMSE of the cross-validation (RMSEcv). Linear models for HT estimation presented a better fit (R^2=0.84, RMSE=0.84 m, MAPE=0.039, Bias=0.002) than the Random Forest (R^2=0.85, RMSE=1.26 m, MAPE=7.19, Bias=-0.173) model including only independent variable the 99th percentile. The RMSE/RMSEcv ratio presented a higher value for the linear model (0.93) than Random Forest (0.75). For the estimation of the site index (SI), the Random Forest model was applied, which included the LiDAR metrics corresponding to the 99th percentile and the 80th bicentile. This model had an R^2 value of 0.65 and an RMSE value of 1.62 m. Then, on the SI raster generated by the Random Forest model, automatic segmentation was applied, generating segments with high internal homogeneity and low homogeneity between segments. The methodology developed in this work provides accurate estimates and mapping of HT and SI at stand scale based on LiDAR data. In addition, an automatic segmentation method was applied, generating stands based on the SI. This segmentation is very useful for the sector as it is a tool that will improve forest management in terms of harvesting and future plantations.

Keywords: LiDAR; intensive silviculture; *Eucaliptus* spp.; linear model; Random Forest; total height; site index; stand segmentation.

1. Introduction

The surface of Uruguay is 176,251 km^2, of which currently 66% are grasslands (natural, fertilized, improved, or implanted) and around 4.77% is occupied by native forest and 5.91% occupied by planted forest (MGAP, 2019). In Uruguay, commercial forest plantations have been carried out mainly with species of the genus *Eucalyptus* spp. and *Pinus* spp. (MGAP, 2018; FAO, 2020). Within the genus *Eucalyptus* the predominant species are *Eucalyptus grandis* Hill ex Maide and *Eucalyptus dunnii* Maide, and within the genus *Pinus* sp. mainly *Pinus taeda* L. and *Pinus elliottii* Engelm (MGAP, 2019). The wood of *E. grandis* is used for cellulose pulp and for sawmilling, and in the case of *E. globulus* and *E. dunnii* for pulp purposes. In contrast, the *Pinus* sp. is intended for sawing purposes only.

The quality of the forest site determines the potential productivity of an area, referring to the volume of wood generated by a forest in the final harvesting. The term site quality integrates climatic, topographic, and edaphic elements of the site. In Uruguay, the rotation of *Eucalyptus* spp. is around 8 to 10 years for pulp purposes and 20 years for sawmill. In the case of *Pinus* spp., the rotation corresponds to approximately 25 years for sawmills (MGAP, 2019). This potential productivity is quantified with the mean dominant height, being the site index (SI) an indication of the productive capacity of the site. The forest site index refers to the graphic representation that describes the relationship between the dominant height and the age of a stand (Prodan et al., 1997). In Uruguay, SI estimating equations are available for some forest species, as *E. dunnii*, *E. grandis, E. globulus* and *Pinus spp.* (Methol, 2008). These equations consider the current age of the trees, the reference age of the species and the current dominant height. (Methol, 2008).

Traditionally, obtaining information on forests has been carried out through field inventory, presenting important limitations when it is necessary to study large areas (Bergen and Dronova, 2007; Sánchez et al., 2018). Based on the information obtained in the field inventories, forest variables such as volume, total biomass, basal area, and density are estimated, presenting a level of uncertainty and precision (Cruz, 2010). However, there are currently innovative technologies such as remote sensing that includes the use of sensors, the analysis of aerial orthophotographs and other intensive data collection methods that can complement field work (Logroño et al., 2020). This scientific discipline generates greater efficiency when estimating the dasometric variables through lower economic costs, less time invested and less estimation error (Palop et al., 2016).

Light Detection and Ranging (LiDAR) is an active remote sensing system based on the emission of laser pulses that make contact with the surface and bounce off, generating a return signal. This return signal allows calculating the distance from the surface to the sensor since the speed of light is taken as a constant (Palop et al., 2016). LiDAR technology can represent the three-dimensional structure of the forest, allowing the improvement of the estimation of variables such as biomass, volume or basal area, compared to other two-dimensional measurement sensors, such as photographic systems or radiometers (García, 2010; Alberola et al., 2018; Moe et al., 2020; Arumäe, 2020). Precision forestry requires the use of LiDAR technology supplemented with information from field plots. In Australia, LiDAR data have been analyzed to estimate tree height, canopy height, and crown diameter, which were then used to infer canopy volume of trees in the genus *Eucalyptus spp*. (Verma et al., 2019). A study in Estonia, forest height and volume were estimated using LiDAR metrics as an independent variable. In this work the

height was estimated using the 80th percentile of the point cloud and the volume was estimated using the 80th percentile and the cover density. A study conducted in Canada investigated the potential of Random Forest, a machine learning technique, to estimate canopy structure as measured by LiDAR (Ahmed et al., 2015). Another study, conducted in Canada, combined information derived from LiDAR plots with Landsat pixel-based composites to produce annual estimates of forest structure from 1984 to 2016 in more than 650 million hectares of Canada's forest ecosystems (Matasci et al., 2018). In a study in Uruguay, LiDAR metrics have been used to improve inventories of forest stands of *Eucalyptus* spp. (Hirigoyen et al., 2020).

In terms of forest management, stand delineation is of utmost importance for efficient forest planning. Forest stands are uniform in composition, size, age or species, and are managed as a single unit (García et al., 2014). The main objective of such delineation is to generate economic and/or ecological benefits (O'Hara et al., 2013). Traditionally, stand delineation was carried out manually using information from the field and high-resolution photographic images (Dechesne et al., 2016; Arumäe, 2020). This method is not efficient because it requires a lot of analysis time by human operators. In addition, the degree of subjectivity of this method means that the results tend to vary between different operators (García et al., 2014). At present, automatic segmentation is available based on dasometric variables derived from LiDAR metrics and is useful for precision forestry, as it allows the delineation of stands based on these variables. Automatic segmentation is capable of generating more homogeneous stands than those defined manually by traditional methods. Also, another advantage of automatic segmentation is the lower time invested and lower cost when delineating stands compared to traditional methods (Gutiérrez et al., 2013; Hirigoyen et al., 2020; Pukkala, 2020). It should be noted that optimal segmentation should decrease within-segment variability and increase between-segment variability (Sánchez et al., 2018; Pukkala, 2020). In a study, carried out in Spain, accurate segmentations of *Pinus sylvestris* stands were generated, where mean shift or multiresolution segmentation methods were used from LiDAR data (Varo et al., 2017; Varo and Navarro, 2021). In Uruguay, a semi-automatic object-oriented segmentation algorithm was used for stand delimitation based on Airborne Laser Scanning (ALS) data in *Eucalyptus grandis* and *E. dunnii* plantations. In this study, above ground biomass and total volume maps were generated for harvesting tasks based on rasters derived from LiDAR metrics (Hirigoyen et al., 2020).

However, less experience exists in the use of LiDAR data to conduct site quality analysis (Cheng and Zhu, 2012). By using height projection equations, the site index at the reference age of the species can be estimated, and LiDAR can be used in these equations. In the United States, site index was estimated with dynamic site index equations based on a long-term loblolly pine plantation applied on the height raster generated with the dominant mean height model based on LiDAR metrics as independent variables (Gopalakrishnan et al., 2019). In Poland, a study demonstrated how airborne laser scanning bitemporal data collected within an 8-year period can be used for the development of site index models for *Pinus sylvestris* L. based on tree height (Socha et al., 2020). In a study in Norway, the site index (SI) of the forest was estimated using LiDAR metrics as an independent variable. In this work a regression was performed between the SI of each plot against LiDAR metrics for *Picea abies* (L.) H. Karst and *Pinus sylvestris* L. As a result, SI maps were generated for each tree species and for the study area. (Noordermeer et al., 2020). The hypothesis of this work was that the segmentation of *Eucalyptus dunnii* stands according to dominant height allows inferring

site index and mapping site quality for use in forest management. The general aid is to generate a stand-scale mapping from LiDAR data to improve and optimize of forest management based on the estimation of the site index in commercial plantations of *Eucalyptus dunnii* Maide in Uruguay. The specific objectives were i) estimate the dominant height of *E. dunnii* plantations from LiDAR data; ii) estimate the site index based on LiDAR metrics and delimit forest productivity zones based on the site index. The methodology developed in this work provides accurate estimates and mapping of the variable total height (HT) and site index (SI) at stand scale based on LiDAR data. In addition, an automatic segmentation method was applied, generating stands based on the SI. This segmentation is very useful for the sector as it is a tool that will improve forest management in terms of harvesting and future plantations.

2. Materials and methods

2.1 Study area

This work was carried out in commercial plantations of *Eucalyptus dunnii* of the company Forestal Oriental S.A. located in the departments of Río Negro and Paysandú, with coordinates of 32°33′04.51″S - 57°14′39.61″W (Figure 1). Climate of the area is Cfa according to the Koppen-Geiger classification, characterized by hot summers and rainfall distributed throughout the year, with a mean annual temperature of 19.2 °C and annual accumulated precipitation of 1262.5 mm (Castaño et al., 2011). Dominant soil is phaeozem with a mollic horizon and no secondary calcium carbonate in the upper, accumulation of organic matter and saturated in bases in its first meter of depth (FAO, 2015). According to the classification of the National Commission for Agroeconomic Studies of the Land (CONEAT) of Uruguay, predominates soils in the area are characterized by presenting a texture from sandy loam to sandy clay loam, average fertility to low, moderately deep and generally well drained. This soil group presents an average productivity compared to the forest priority soil groups (MGAP, 2018).

2.2 Field data

In the months of May and June of the year 2017, 43 plots with a radius of 10 m (314.16 m^2) were established in the field. A systematic sampling design was carried out according to traditional inventory procedures for monospecific plantations of *Eucalyptus* sp. Each field plot contains data corresponding to diameter at breast height (dbh at 1.3 m, cm), density (number of trees per hectare), basal area (G, m^2 ha^{-1}), total height (H, m) and volume per hectare (m^3 ha^{-1}). Table 1 shows a summary of the measurements made on the sample plots.

Table 1. Silvicultural variables of Eucalyptus *dunnii*: number of plot(n), density (N); diameter at breast height (dbh); basal area (G), total height (TH), total volume (TV), and age. stdev: standard deviation, min: minimum, max: maximum.

	Field Attributes	**Stdev**	**Min**	**max**	**mean**
Eucalyptus dunnii (n=43)	Age (years)	1.71	5.00	10.00	7.44
	TH (m)	2.43	16.20	25.30	20.33
	Dbh (cm)	2.39	12.65	22.99	17.29
	G (m^2 ha^{-1})	5.44	14.11	36.84	149.19
	TV (m^3 ha^{-1})	37.87	80.49	235.94	149.19
	N (tree ha^{-1})	231.90	445.63	1401.00	992.97

2.3 LiDAR data acquisition and processing

The LiDAR sensor data was obtained in April 2017, for an area of 1995 ha (Figure 1), using a Riegl VUX-1 laser scanner installed on an autogyro helicopter, at a flight altitude of 110 m above ground level, a pulse repetition rate of 550 kHz, a wide angular pitch of 0.0687 ° and a field of view of 55 °. The point density is 12 points per m^2. For the georeferencing of the plots, the WGS84 UTM 21 (EPSG: 32721) coordinate system was used. The LiDAR point cloud was processed with the LAStools software through the Windows cmd console, generating the normalized point cloud, the digital surface model (MDS) and the digital vegetation model (MDV). To carry out the normalization of the point cloud, the "lasheight" function was used, taking as an input file the tiles in LAZ format provided by the company that carried out the LiDAR flight. To obtain the MDS and MDV, the "blast2dem" function was implemented, obtaining as products several rasters in ASCII format. To evaluate the distribution of the pulses on the surface, the "lasgrid" function was used, generating raster files, where the digital value of each pixel (1m) corresponds to the density of pulses per m^2. Then, to obtain the metadata information of the total normalized point cloud, the "lasinfo" function was implemented, highlighting the average number of points per m2 as a result. To obtain the rasters of the LiDAR metrics, the "lascanopy" function was used, using a pixel size of 17.8 m, corresponding to the square root of the surface area of the field plot (314.16 m^2).

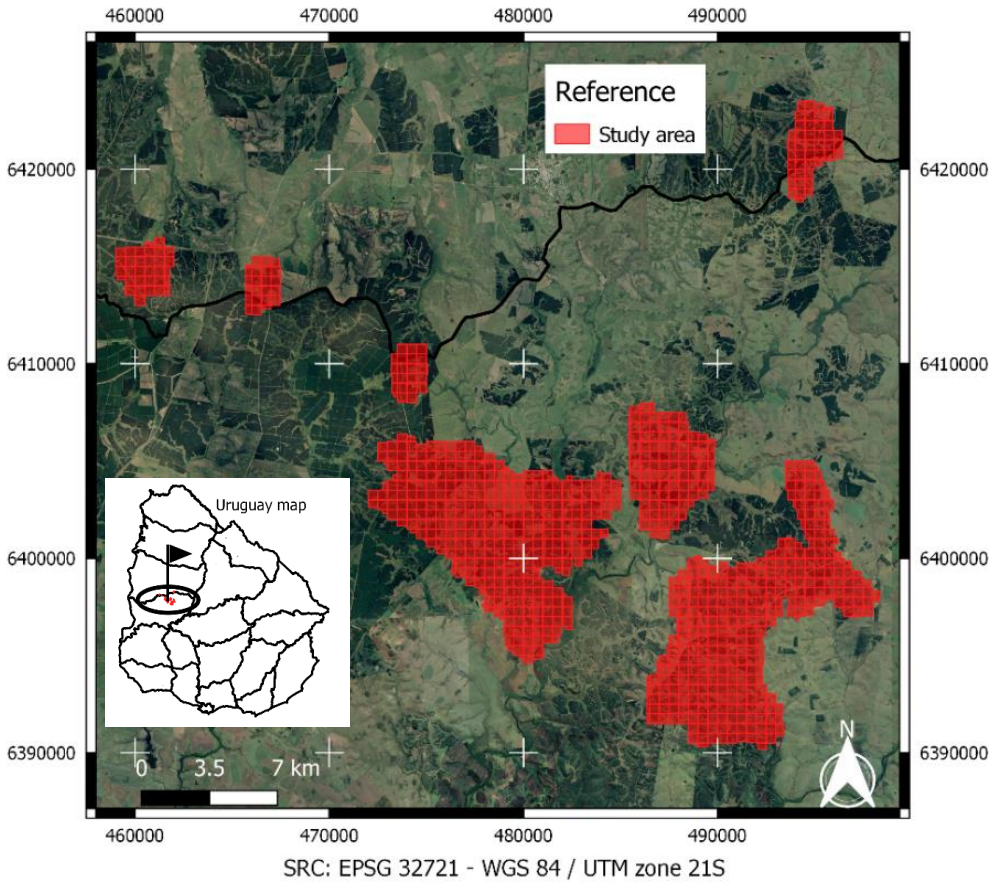

Figure 1.: Study area of *Eucalyptus dunnii* commercial plantations

To carry out the extraction of the LiDAR metrics from the area occupied by the field plots, the “lascanopy” function of the LAStools software was used through the Windows cmd console. The input files used were the normalized point cloud in LAZ format and a vector file in shapefile format that corresponds to the area occupied by the field plots. The parameters used in the code were those corresponding to the metrics to be extracted, "cover_cutoff 2.0" and "height_cutoff 0". The parameter "cover_cutoff 2.0" indicates the height from which to consider the points in terms of the "canopy cover" (similar to the fraction of the covered space). It is indicated that it is considered from two meters in height, to avoid points from non-relevant vegetation, such as scrub. Therefore, it only affects the “cover” metric. The parameter "height_cutoff 0" is similar to the previous one but refers to the cutting height to be considered by all statistics. In this case, it is set to 0 to be able to obtain the percentiles and bicentiles of the entire range of heights. Then, to obtain the database, the LiDAR metrics file and the file with the field data are merged.

2.4 Variable selection and statistical analysis

Before obtaining the estimation models of the total height based on the LiDAR metrics, the study of the correlation matrix between them was carried out. For this, the "cor" function available in the R package was used. Then, the study of the normality of the uncorrelated variables was carried out by means of the "shapiro.test" function available in the R package. Linear predictive models were fitted for the estimation of total height as a function of the set of metrics selected above. The construction of these models was carried out with the "lm" function of the R package.

2.5 Variable selection and k-NN models with method Random Forest

The analysis of the correlation between variables was carried out to reduce data collinearity, and then the k-NN Random Forest algorithm was applied. The Variance Inflation Factor (VIF) was implemented. The correlated variables were eliminated, and the varSelection function of yaImpute (Crookston et al., 2008) in the R package was used to select the variables. The criterion used to select the variables was a critical threshold of VIF>10 (Quinn et al., 2002). This function calculates the generalized root mean square distance (grmsd) at each moment of the aggregation of variables to an imputation algorithm, preserving the variables that strengthen the imputation process (Crookston et al., 2008).

The database was divided into two groups, corresponding to training and validation, containing 70% and 30% of the total data, respectively. Then, the model fit was calculated with the training data, using the "yai" function available in the "yaimute" package of R. Next, the "trainControl" function of the "caret" package was implemented to generate parameters that control the model creation process, allowing the calculation of the k value for the Random Forest imputation method. The value of k was chosen for the imputation that presented the lowest root mean square error (RMSE) value. Then, the "randomForest" package was used to obtain the nonparametric model. In addition, the "q" value was calculated to detect possible overfitting of the model (Weisberg, 1985).

2.6 Model assessment and validation

The coefficient of determination (R^2) was used to determine the goodness of fit of the models. This coefficient represents the quality of the model for estimating the results, and the proportion of variation in the results that the model can generate. The performance of the models generated was evaluated by studying the estimation errors. These estimation errors correspond to the root mean square error (RMSE), the mean absolute percentage error (MAPE) and the model bias (Bias). Simple linear regressions relating observed and predicted values were evaluated using the coefficient of determination (R^2) to assess the selected models (Zald et al., 2016).

After selecting the best model, it was validated. Validation is of utmost importance because it will determine whether the model can be extrapolated to the entire study area (Altman et al., 2000). Therefore, validation will determine the credibility of the estimation of the dependent variable in different cases (Fernández, 2018). To validate the linear models, the one leave one cross validation technique was used. This technique consists of removing one observation from the sample in each interaction and fitting the selected model using the remaining observations and estimating its value (Montealegre et al. 2016, Lekuona et al. 2018, Hirigoyen et al., 2020). Each fit is going to have its R^2, RMSE and Bias values. These values were averaged obtaining R^2, RMSE and Bias from the cross-validation, which were compared with the values of the selected model (Knapp et al.,

2017). The model that presents the best quality of fit corresponds to the one that presents close values of RMSE of the cross-validation and the RMSE of the model to be evaluated. As these RMSE values get closer, the prediction of the model improves, and the overfitting of the data becomes less (Anderson et al., 2005).

When validating the nonparametric Random Forest model, the internal (training) and external (evaluation) accuracy was studied. To carry out this study, the database was divided into training (70% of the data) and evaluation (30% of the data) (Fernández, 2018). These two databases were taken independently, and the estimation errors were calculated. In this case, the quality of the validation estimates is higher when the ratio between RMSE of training data and RMSE of evaluation data is closer to the value of 1. This determines whether the model was applicable to the entire area to be studied. In addition, the cross-validation technique was used. For the comparison of the best linear model (parametric model) and the K-nn Random Forest model (non-parametric model), the R^2 and RMSE were used, since these coefficients summarize the performance of the models and the existing difference between predictions and observations (Zald et al., 2016).

2.7 Obtaining height raster and site index raster

To generate the height raster, the best selected model was applied on the raster of the LiDAR metric corresponding to the independent variable. The raster corresponding to the LiDAR metric in question was obtained using LAStools software. The generated height raster has a pixel size of 17.8 cm, corresponding to the surface of the plots used in this work (314.16 m^2). In addition, this raster presents in each pixel a digital level that refers to the average total height in meters.

To obtain the site index (SI) raster, first, a new attribute corresponding to the SI value was generated in the database. To estimate the SI of the plot, the equation proposed by Methol (2008) for the species *E. dunnii* was applied. This equation is composed of the current total height of the plot (AMD_1), the current age of the trees in the plot (t_1) and the reference age of the species (t_2). In the case of *E. dunnii*, a t_2 equal to 8 years is considered (Methol 2008). Linear and non-parametric models were then generated to estimate SI as a function of LiDAR metrics. The best of these models was applied on the LiDAR metrics raster to obtain the Site Index raster. Eq. (1)

$$AMD_2 = AMD_1 \frac{\left[1 - e^{-0.15 * t_2}\right]^{1.0915}}{\left[1 - e^{-0.15 * t_1}\right]^{1.0915}} \qquad eq(1)$$

2.8 Segmentation Method

The algorithm used to carry out the stand segmentation was based on the non-parametric estimator called Mean Shift (MS). This algorithm was implemented with the Orpheo ToolBox (OTB) software available in QGIS tools (QGIS, 2009). This automatic segmentation method was applied on the site index raster, generating homogeneous segments in terms of the IS value of the pixels. In this work, the function "otbcli_LargeScaleMeanShift" of the OTB tool was used from the Windows CMD console, defining different parameters in the code. The parameters corresponded to spatial

radius, range radius and minimum region size. The value of spatial radius will determine the neighborhood to average, the value of range radius refers to the digital value threshold that the algorithm will consider delimiting segments (this parameter defines the interval in the spectral space), and the value of minimum region size defines the smallest size of the segment (Comaniciu et al., 2002). These parameters define the homogeneity within each segment and the heterogeneity between segments. In segmentation, the aim is for segments to be as homogeneous as possible and for heterogeneity between segments to be high. In this work, different combinations of parameters that generated different segmentations were tested. The set of parameters that resulted in the lowest intra-segment variation and the highest inter-segment variation was chosen. This variation was obtained by means of the statistical zonal tool of the QGIS software, taking as input the shapefile resulting from the segmentation.

2.8.1 Unsupervised evaluation of the segmentation method

To evaluate the segmentation method applied to the site index raster of this work, unsupervised evaluation (NSE) was chosen. This technique is based on expert knowledge and generates better results than the supervised evaluation. This is because the NSE generates results of higher efficiency and lower subjectivity than the ES technique (Wu et al., 2013; Varo et al., 2017). In the unsupervised evaluation each segment should present as much homogeneity as possible within it and in addition, the segments should be differentiated from each other (Varo et al., 2017). In this work to select the best segmentation, the evaluation of the internal homogeneity of the segments and the evaluation of the heterogeneity between segments were carried out. The evaluation of the internal homogeneity of the segments was carried out by means of the internal variance of the segments, selecting as the best segmentation the one with lower variance values and smaller difference between variances. The evaluation of the heterogeneity between segments was determined by means of the variance of the Site Index values of each segment, that is, the one with the highest variance value was selected as the best segmentation.

3. Results

3.1 Linear models

In the case of the linear models for estimating total height for *Eucalyptus dunni*, metrics corresponding to the 99th, 95th and 90th percentiles were used. These LiDAR metrics presented high Pearson´s correlation coefficients with total height (Table 2). The linear models presented similar values for the coefficient of determination (R^2) and a slight variation in the root mean square error (RMSE). Therefore, the linear model with the 99th percentile (model 1, Table 3) was selected since it had a higher ratio between the RMSE of the model and the RMSE of the cross-validation (RMSE/RMSEcv= 0.939). This defines a higher quality of fit by that model compared to the other two models using the 95th and 90th percentile metrics (Table 3).

Table 2. Correlation coefficient of Pearson for each independent variable and the dependent variable of total height: TH (m): total height; p99, p95 and p90: percentile 99, 95 and 90.

	p99	**p95**	**p90**
TH (m)	0.92	0.92	0.92

Table 3. Linear models with their fit values and cross-validation: TH (m): total height; R^2: coefficient of determination; RMSE: root mean square error (m); MAPE: absolute percentage error; Bias: model bias; RMSEcv, MAPEcv y BIAScv: errors of the cross validation.

	Models	R^2	RMSE	MAPE	Bias	RMSE cv	MAPE cv	BIAS cv	RMSE/ RMSEcv
	5.96 + 0.659*p99 model 1	0.84	0.94	0.04	0.00	1.00	0.83	0.00	**0.94**
TH (m)	5.442 + 0.721*p95 model 2	0.85	0.90	0.04	-0.02	1.16	0.97	0.01	**0.77**
	4.444 + 0.794*p90 model 3	0.85	0.99	0.41	0.00	2.04	1.22	0.01	**0.49**

Figure 2 shows the relationship between the observed and predicted total height values for the selected model (model 1). After performing a linear model of estimated total height (HT predicted) as a function of observed total height (HT observed), it was found that the determination coefficient (R^2) corresponded to 0.82. This determined that there is almost perfect agreement between predictions and observations.

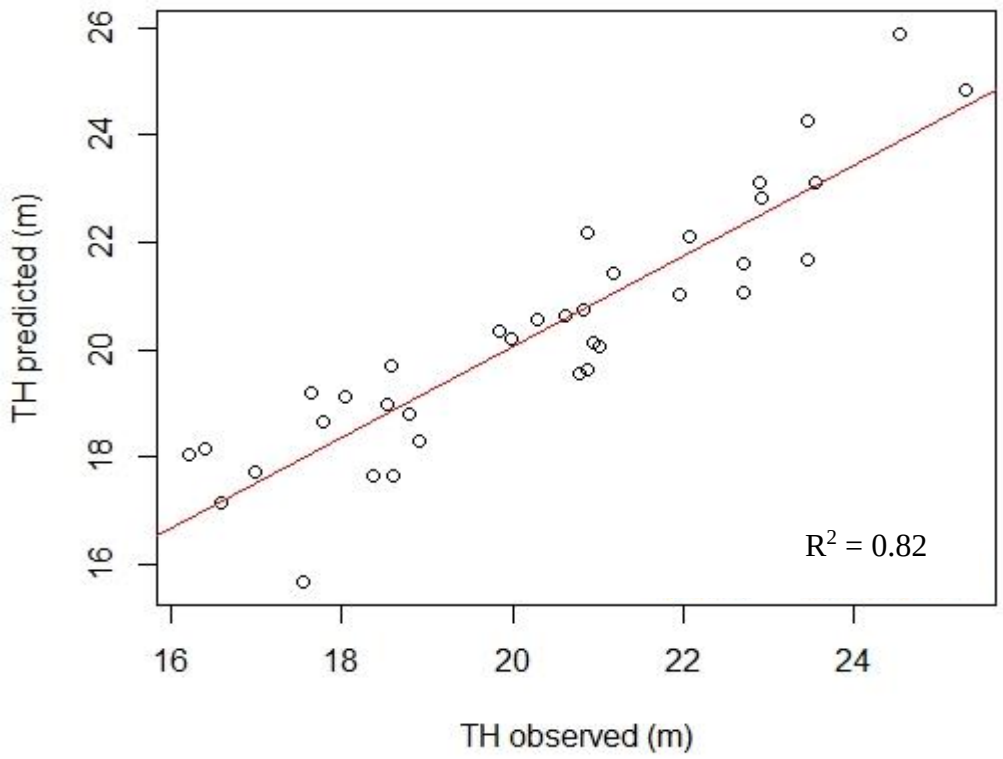

Figure 2. Relationship between observed and predicted values for total height (TH m) of *Eucalyptus dunnii* plantations by model 1.

3.2 Non-parametric Random Forest model

After applying multicollinearity analysis and the correlation test to the metrics to be included in the model, the varSelection function was used. The candidate metrics to be used in the Random Forest model as predictor variables of total height were defined with the generalized mean square distance statistics (gmsd, Table 4). The selected metrics were used together and individually within the model, with the 99th percentile model presenting a higher R^2 value (R^2 = 0.85, RMSE = 1.27, Table 5), using a value of "k" equal to 5. The value of "q" of this model corresponded to a value of 1.26, presenting some overfitting because it was above the empirical value of 1.07. In addition, the result of the cross-validation presented a RMSEcv value of 1.67, being the RMSE ratio of the model (Table 5) and RMSE of the cross-validation equal to 0.76. The value of the ratio between internal precision and external precision was 0.60 (Table 5).

Table 4. Statistics of generalized mean square distance (gmsd): sd: standard deviation; p99 and p75: percentile 99 and 75; b80 and b90: bicentile 80 and 90; dns_gap: cover density.

LiDAR metrics	**gmsd (mean)**	**gmsd (sd)**
p99	0.71	0.08
p75	0.71	0.09
b80	0.71	0.07
b90	0.64	0.07
dns_gap	0.65	0.07

Table 5. Fit values, errors and the comparison between the RMSE of the training data with the RMSE of the evaluation data of the Random Forest model.: HT (m): total height; R^2: coefficient of determination; RMSE: root mean square error (m); MAPE: absolute percentage error; Bias: model bias; RMSEt/RMSEe: RMSE of the training and the RMSE of the evaluation.

		Internal accuracy (Train)				**External accuracy (Evaluation)**	
Var.	**Models**	**R^2**	**RMSE**	**MAPE**	**BIAS**	**RMSE**	**RMSEt/RMSEe.**
HT (m)	model 4	0.85	1.27	7.20	-0.17	2.12	0.60

Figure 3 shows the relationship between the observed and predicted total height values for model 4 model. This relationship showed a determination coefficient (R^2) of 0.57. This value determined an acceptable agreement between observed and predicted values.

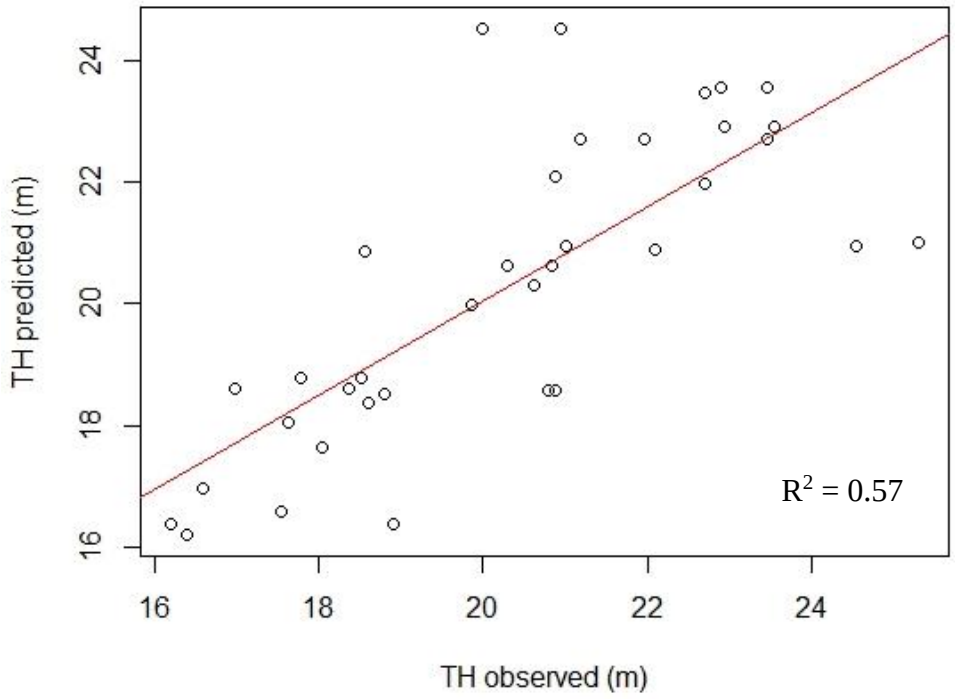

Figure 3. Relationship between observed and predicted values for total height (TH m) of *Eucalyptus dunnii* plantations by model 4.

3.3 Site index and height raster

Two raster of stand height were obtained, one using model 1 and the other using model 4. In both cases the model was applied on the LiDAR metric raster corresponding to the 99th percentile, since it was the metric included in the selected model. Two rasters had a pixel size of 17.8 m and the digital level corresponded to the height of the trees in meters. The height rasters generated with models 1 and 4 were of good quality, as they represented the reality of the forest, explained by the acceptable validity of the models.

In the case of site index (SI) estimation, a linear model using LiDAR metrics as independent variables could not be fitted satisfactorily. Nevertheless, when using the model 4, a coefficient of determination (R^2) of 0.65, an RMSE of 1.627m and Bias of -0.3 were obtained, using as independent variables the metrics corresponding to the 99th percentile and the 80th bicentile. Figure 4 shows the relationship between the observed and predicted site index (SI) values for Random Forest model. This relationship showed a determination coefficient (R2) of 0.41. This value determined a relatively low agreement between observed and predicted values. Then, in R with the stack function of the raster package and rgdal, the rasterstack of 99th percentile raster and 80th bicentile raster was carried out. As a result, a single two-band raster was obtained, keeping the

digital levels of the metrics in each band. The Random Forest model was then applied to this rasterstack to obtain the site index raster.

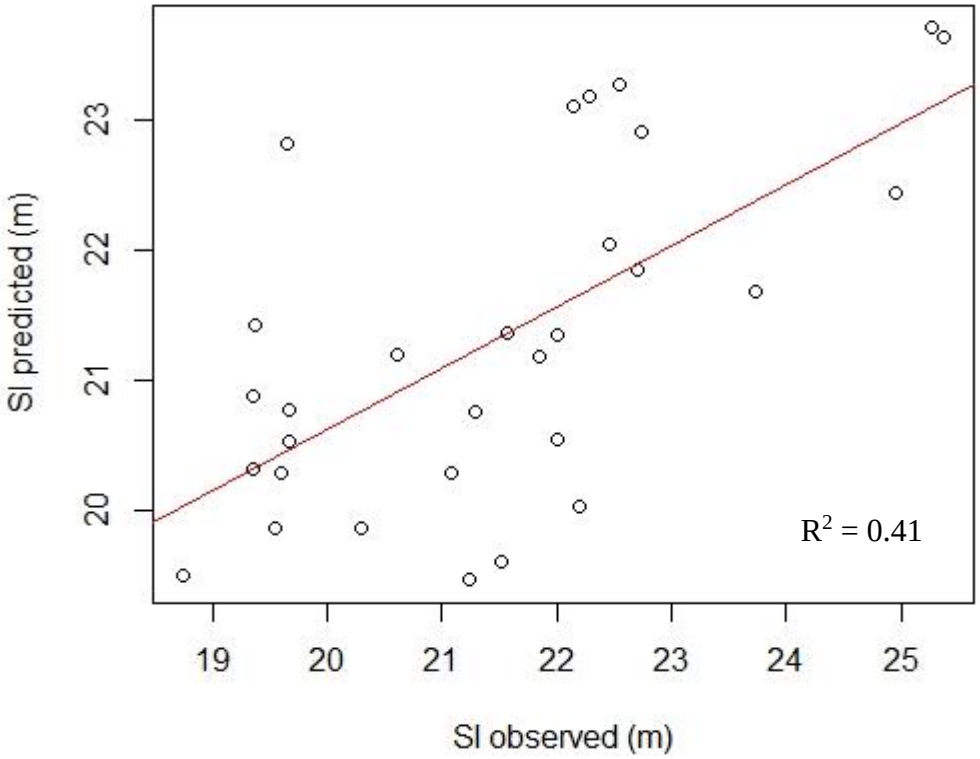

Figure 4. Relationship between observed and predicted values for index site (SI m) of *Eucalyptus dunnii* plantations by Random Forest.

3.4 Segmentation OTB

The result of the OTB segmentation for forest stand delimitation based on the site index varied according to the combination of the parameters spatial radius, range radius and minimum region size. Various combinations of these parameters were tested, and the ten best segmentations were taken for evaluation (Table 6). Table 6 shows a value of 1 for "Hete. Ranking" for the largest variance between the SI means of each segment, and with a value of 1 for "Homo. ranking" for the largest variance of the SI internal variance values of each segment. The segmentation selected corresponded to the one referred to as "i", as it was the one with the lowest intra-segment variation and the highest inter-segment variation (Table 6). The parameters of this segmentation corresponded to a value of 30 for spatial radius, a value of 1 for range radius and a value of 64 for minimum region size. The minimum region size value of 64 is equivalent to an area of 2,03 hectares. This value was obtained by multiplying the area of the SI raster pixel by the minimum region size value of 64 pixels.

Table 6. Combination of spatial radius, range radius and minimum region size parameters and evaluation of segmentations: Seg.: segmentation name; SP: spatial radius; RR: range radius; MRS: minimum region size; Var. mean SI: variance site index value of each segment (m^2).; Hete. ranking: heterogeneity ranking.; Var.Var. SI: variance of the variance values of site index presented by each segment.; Homo. Ranking: homogeneity ranking.; Nº seg.: number of segments.; Area (ha): mean area of segments in hectares.

Seg.	Parameters			Var. mean SI	Hete. ranking	Var. Var. SI	Homo. ranking	Nº seg.	Area (ha)
	SP	RR	MRS						
e	20	8	16	1.32	10	255.72	1	529	10.81
c	20	3	16	1.76	9	276.08	2	582	11.77
b	16	3	16	1.90	8	321.92	3	635	10.96
i	30	1	64	2.99	5	393.38	4	585	11.60
d	20	1	16	3.24	3	423.74	5	1558	4.39
j	35	1	64	3.10	4	426.15	6	584	11.63
a	4	3	16	2.17	7	428.78	7	413	16.63
f	20	1	64	2.95	6	435.61	8	630	10.79
g	20	1	95	3.29	2	462.38	9	499	13.55
h	35	1	95	3.44	1	473.19	10	458	14.74

Figure 5 represents the result of the delimitation of the stands based on the OTB segmentation as a function of the site index. Figure 4 and figures 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 (Annexes) represents the detail of the generated segments and location in the study area.

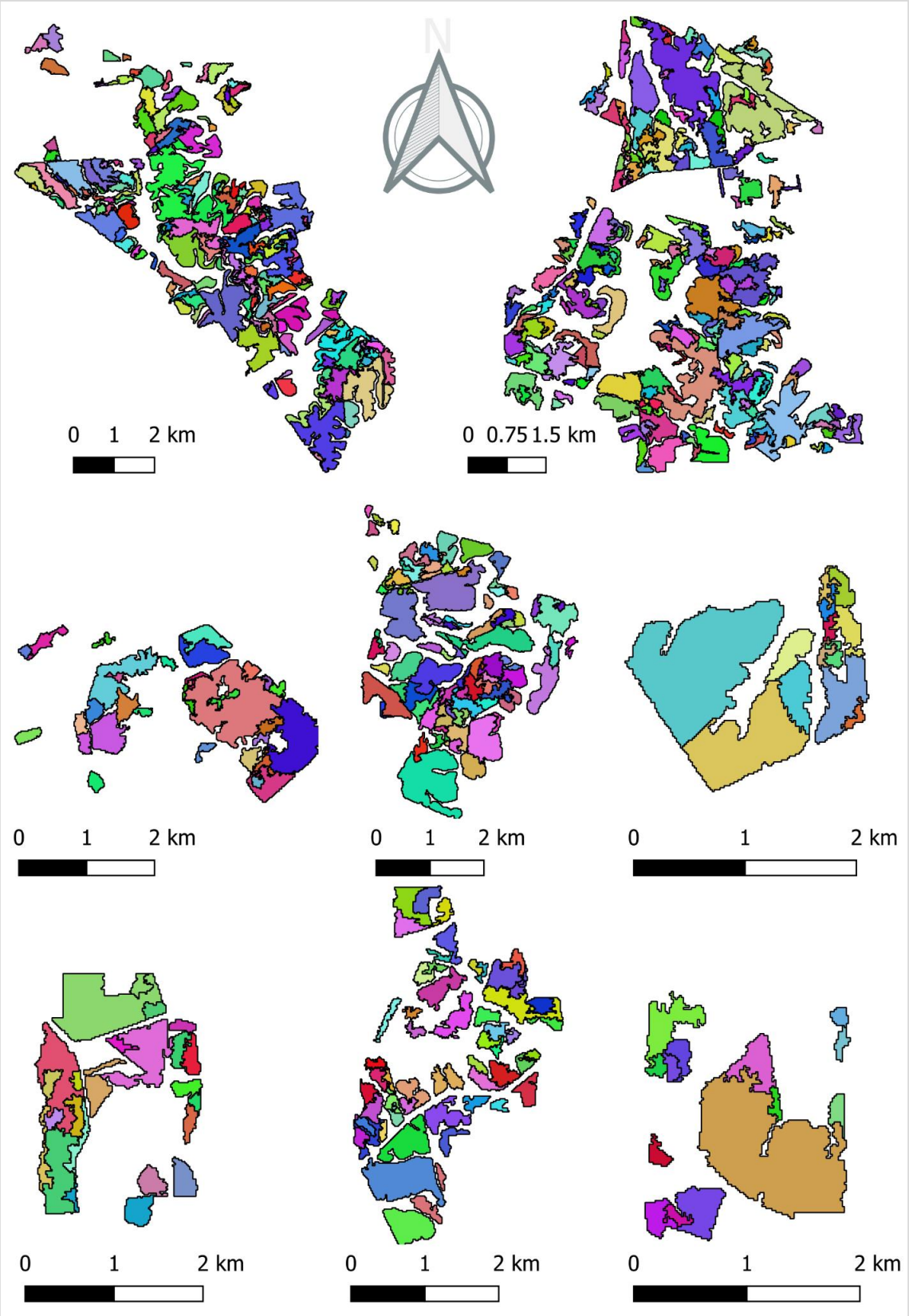

Figure 5. Representation of the detail of *Eucalyptus dunnii* plantation segments of segmentation "i" by site index.

4. Discussion

Technology involving ALS sensors has been applied in the forest environment, aiming to determine forest variables (Böck et al., 2017; Yang et al., 2019). In this work high spatial resolution site index (SI) and total height (HT) rasters were generated using models based on LiDAR metrics. Linear models and the non-parametric model using Random Forest imputation of the K-nn algorithm were compared. The linear model was used because of its simplicity and less processing when generating the rasters, and the non-parametric model was applied due to its quality of being independent of the data distribution. In addition, forest stand delimitation was achieved based on the site index raster, delimiting areas of different forest productivity for *E. dunnii* specie. The estimation of total height is closely related to the site index (SI) and used as indicators of forest site quality (Mora et al., 2013). For segmentation, the Mean Shift (MS) segmentation algorithm available in the Orpheo ToolBox (OTB) software was used, and the results of the site index segmentation were of high quality.

4.1 Height estimation models and height raster generation

The most common methods for estimating total height in forest inventories using LiDAR metrics correspond to simple regression methods and non-parametric methods (Leite et al., 2020; Coops et al., 2021; Pourshamsi et al., 2021). Our results demonstrated that the simple regression method and the Random Forest model are suitable methods to estimate total height based on LiDAR metrics. Regarding the selection of LiDAR metrics for the linear models, the best predictors of total height were the 99th, 95th and 90th percentiles. This is because these percentiles correspond to the part of the LiDAR point cloud that represents the first returns which are generated by the impact of the laser pulse on the highest part of the canopy of the trees. There are studies confirming a high correlation between total tree height and metrics corresponding to the highest percentiles (Varo et al., 2017; Shao et al., 2018; Hirigoyen et al. 2020; Arumäe, 2020). In this case, the accuracy of the height models was high because of low RMSE values and high R^2 values, with a high ratio between the RMSE of the model and RMSE of the cross-validation, determining this a good predictive value. Random Forest method has been also highlighted as one of the best k-NN models for estimating dominant height, total volume and biomass. (Zald et al. 2016; Hirigoyen et al. 2020). In the case of the Random Forest model, the best model used the 99th percentile as the only independent variable. When comparing the results between the linear models (model 1, 2 y 3) and the Random Forest model (model 4), a similar value of the coefficient of determination (R^2) and similar values of RMSE were observed. However, in the case of the model 4, the RMSE presented a slightly higher RMSE value compared to the linear models. In addition, it should be noted that the model 1, which included the 99th percentile, had a higher ratio between the RMSE of the model and the RMSE of the cross-validation, with a value of 0.94. In turn, the relationship between the observed and predicted values showed a higher determination coefficient for the model 1 with the 99th percentile (R^2=0.82) than for the model 4 with the 99th percentile (R^2=0.57). These results are in line with precedents on the comparison of linear and non-parametric models (Xu et al., 2018; Silva et al., 2018; Hirigoyen et al., 2020). However, this work highlights the inclusion in model 4 of only one LiDAR metric (99th percentile), compared to the Random Forest model generated in the study by Hirigoyen et al. (2020) which included the 75th percentile and Elev.max metrics (maximum statistic of all heights above the cut-off height of the point cloud).

The height rasters generated with model 1 (linear) and model 4 (Random Forest) presented good quality, due to the concordance between the height values of each pixel with the ground reality represented with the 99th percentile raster. In the case of the height raster generated with model 4, the presence of areas of pixels of similar height (homogeneous patches) stands out in the interior of the forest. This is explained by the use of *Eucalyptus dunnii* clones by the company Forestal Oriental S.A. The clones tend to have similar height values for the same age. In addition, in this raster, a large variation of heights (heterogeneous patches) was observed at the edge of the forest, explained by the edge effect. Trees at the forest edge tend to be more exposed to climatic conditions, such as wind damage. In addition, trees at the edge tend to use their energy for greater production of lateral branches (greater crown volume), resulting in a lower height of these individuals.

4.2 Site index estimation

Site quality determines the productive potential of an area, referring to the volume of timber generated by a stand at final rotation, and is quantified by the site index (SI). The SI is defined as the dominant height at a key age for a forest species. The SI is an indicator of stand productivity, because, for the same site and species, its value will vary in relation to the quality of the genetic material and the silviculture applied (Methol, 2008). LiDAR technology provides a scan of the surface, generating a census of the heights of the entire forest area. The SI can be estimated by means of models where the independent variables (LiDAR metrics) correspond to the highest percentiles of the point cloud. This is because there is a high correlation between the dominant height, included in the SI equation of Methol (2008), and these percentiles. The possibility of having SI mapping of the entire forest area is of utmost importance, as it increases the efficiency of stand harvesting planning.

Linear models and the Random Forest model were used to estimate the site index. The linear models presented very low values of the coefficient of determination (R^2) when the regression between the SI and the metrics of the highest percentiles of the point cloud was performed (R^2_{p99}=0.09; R^2_{p95}=0.08; R^2_{p90}=0.07). These results are concordant with those obtained in Brazil for non-linear mixed SI estimation models for *Eucalyptus urugrandis* (Packalén et al., 2011). In that study, the Chapman-Richards equation was applied, which is very similar to the SI equation used in this work (Methol, 2008). In the SI equation, the current dominant height of the plots was replaced by a linear height estimation model based on LiDAR metrics. However, in this work Random Forest model showed a worst fit (R^2=65%, RMSE=1.627m), similar to those obtained by Noordermeer et al. (2020, R^2<69%). Our model included the 99th percentile and the 80th bicentile, and the Noordermeer et al. (2020) model included the 90th, 60th percentiles and the difference in the 90th percentile of height (ΔH90). In future studies, the way of estimating the SI Packalén et al. (2011) may generate better models in future studies for *E. dunnii* in Uruguay.

4.3 OTB segmentation based on the site index

OTB segmentation to delimit eucalyptus forest stands is simpler, more efficient and of similar accuracy compared to other more complex methods (Hirigoyen et al., 2020). This automatic segmentation method is better than manual segmentation, mainly due to the faster generation of segments (stands). In this work, this segmentation method was applied on the site index (SI) raster generated with the Random Forest model. Visually and quantitatively, the segmentation results of this work were of good quality.

The segmentation obtained presented high homogeneity within each segment, in concordance with previous studies (Varo et al., 2017; Pukkala, 2020; Hirigoyen et al., 2020). Therefore, OTB segmentation is a very useful tool to automatically segment *Eucalyptus dunnii* stands. However, it is important to note that the shape and size of the resulting stands is not perfect, since there were forest areas which were not segmented and areas with small polygons. These imperfections were present in the results of several studies and were manually modified (Gutiérrez et al., 2013; Duarte et al., 2018). Such imperfections would be related to the combination of parameters used and the SI raster entered as input. This is because the unsupervised Mean Shift (MS) classifier will delineate the segments based on the digital levels of the raster in question. To overcome these imperfections, different combinations of spatial radius (SR), range radius (RR) and minimum region size (MRS) were tested. In this work we defined a value of 30 for SR, a value of 1 for RR and a value of 64 for MRS. The SR value of 30 will determine the number of neighbouring pixels to be averaged. The value of 1 for RR defines the interval in the spectral space when performing the delineation of the segments, this low value determines that the classifier finds larger differences in the SI raster. The value of 64 for MRS defines a minimum segment size of 2.03 ha, solving to a large extent the generation of very small segments that are not representative of the real terrain.

4.4 Forest management applications

The possibility of generating models to estimate forest variables based on LiDAR metrics is an important tool for good quality inventories. This technology allows obtaining information from the whole study area and requires a low number of calibration plots. In the case of traditional inventories, a larger number of plots is required, resulting in a greater need for human resources, a greater time investment and a higher economic cost. LiDAR technology provides a census of forest height, generating more accurate results than traditional inventory. When planning forest harvesting, it is essential to have the total volume of the stands, with the height of the forest being a fundamental variable for its calculation.

The use of the OTB segmentation method improves processing time by increasing the area covered and maximizing labor efficiency (Ortega, 2018; Hirigoyen et al., 2020). The segmentation based on the site index (SI) is very useful silvicultural product because it defines the productivity of different sites. Efficient forest management requires stands to be as homogeneous as possible within the forest in terms of IS. The selection of segmentation "i" was not only based on intra-segment homogeneity and inter-segment heterogeneity, but also on the number of segments and the average area of each segment. The segmentation presented 585 segments for the total surface (1995 ha) with an average area of 11.6 ha. This stand size is consistent for forest planning and harvesting planification. The minimum region size (MRS) parameter value defined a minimum stand area of 2.03 ha, allowing for relatively small stands that will determine the productivity of each site in more detail. Knowing the productivity of the forest area will not only promote efficient harvesting but will also improve the planning of future plantations.

The availability of models for estimating total height (HT) and site index (SI), using LiDAR metrics as an independent variable, allows their application in large areas. SI equation derived from LiDAR data for *Eucalyptus dunnii* can predict the production by the reference age of the species (8 years). The efficiency of forest management is improved when information is available for the entire study area, which is not the case in

traditional inventories. Therefore, the OTB segmentation of the SI is an economical option, which allows foresters to improve forest planning.

5. Conclusions

This study showed that it is possible to generate models for estimating total height (HT) and site index (SI) using LiDAR metrics as independent variables. In addition, the best models for HT and SI estimation included the metrics corresponding to the highest percentiles of the point cloud. It is worth noting that the linear model (model 1) and the Random Forest model (model 4), presented similar results; however, model 4 presented a slightly higher RMSE value. As for the estimation of SI, the Random Forest model showed a better fit. By applying these models, it was possible to obtain high spatial resolution maps of HT and SI for the entire study area. Automatic stand delineation from SI based on an unsupervised assessment allows to increase the accuracy of the SI value for each stand. This methodology provides an inexpensive and easy approach to update model for the generation of SI maps, based on raster derived from LiDAR metrics. SI maps are very useful for forest planning, as they define the productivity of the study area. The tools used in this work, based on LiDAR metrics, promote an improvement of the decision-making process on forestry activities based on the SI. This study, in addition to providing new tools for better forest management, promotes the need for further progress in the application of airborne laser scanning (ALS) data for the estimation of the SI of Eucalyptus spp. plantations in Uruguay.

6. References

Ahmed, O.; Franklin, S; Wulder, M.; White, J. Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the Random Forest algorithm. Journal of Photogrammetry and Remote Sensing. 2015, 101, 89-101.

Alberola, S.; Oliver, P.; Estornell, J.; Dopazo, C.. Estimación de variables forestales de *Pinus sylvestris* L. en el contexto de un inventario forestal aplicando tecnología lidar aeroportada. GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica. 2018, 21, 79-99.

Altman, D.; Royston, P.. What do we mean by validating a prognostic model?. Statistics in Medicine. 2000, 19, 453-473.

Andersen, H.; McGaughey, R.; Reutebuch, S.. Estimating Forest Canopy Fuel Parameters Using LIDAR Data. Remote Sensing of Environment. 2005, 94, 441–449.

Arumäe, T.. Estimating forest variables using airborne lidar measurements in hemi-boreal Forests. 2020, 197.

Bergen, K. and Dronova, I. Observing succession on aspen-dominated landscapes using a remote sensing-ecosystem approach. Landscape Ecology. 2007, 22, 1395-1410.

Böck, S.; Immitzer, M.; Atzberger, C. Sobre la objetividad de la función objetivo: problemas con la evaluación de segmentación no supervisada basada en la puntuación global y un posible remedio. Remote Sensing. 2017, 9, 769.

Castaño, J.P.; Giménez, A.; Ceroni, M.; Furest, J.; Aunchayna, R.; Bidegain, M. Caracterización Agroclimática del Uruguay 1980–2009; Serie Técnica INIA 193; Instituto de Investigaciones Agropecuarias: Montevideo, Uruguay, 2011.

Chen, Y.; Zhu, X.. Site quality assessment of a Pinus radiata plantation in Victoria, Australia, using LiDAR technology. Southern Forests: a Journal of Forest Science. 2012, 74, 217-227.

Comaniciu, D. and Meer, P. Mean shift: A robust approach toward feature space analysis. Transactions on Pattern Analysis and Machine Intelligence. 2002, 24, 603–619.

Coops, N.; Tompalski, P.; Goodbody, T.; Queinnec, M.; Luther, J.; Bolton, D.; Hermosilla, T.. Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. Remote Sensing of Environment. 2021, 260, 112477.

Crookston, N.; Finley, A.. YaImpute: An R Package for kNN Imputation. Journal of Statistical Software. 2008, 23, 16.

Cruz, I.; Valdez, R.; Ángeles, G.; Héctor, Y.; De Los Santos, M.. Modelación espacial de área basal y volumen de madera en bosques manejados de Pinus patula y P. teocote en el ejido Atopixco, Hidalgo. Madera y bosques. 2010, 16, 75-97.

Dechesne, C.; Mallet, C.; Le Bris, A.; Gouet, V.; Hervieu, A. Forest stand segmentation using airborne lidar data and very high resolution multispectral imagery. Remote Sensing and Spatial Information Sciences. 2016, 41, 207–214.

Duarte, L.; Silva, P.; Teodoro, A.. Development of a QGIS plugin to obtain parameters and elements of plantation trees and vineyards with aerial photographs. ISPRS International Journal of Geo-Information. 2018, 7, 109.

FAO. Evaluación de recursos forestales mundiales 2020. 2020. Available online: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/sites/ministerio-ganaderia-agricultura-pesca/files/2020-10/Informe%20final.pdf (accessed on 1 August 2021).

FAO. Atlas de suelos de América Latina y el Caribe. 2015. Available online: https://esdac.jrc.ec.europa.eu/Library/Maps/LatinAmerica_Atlas/Documents/LAC.pdf (accessed on 1 August 2021).

Fernández, B.. Validación interna de modelos predictivos de regresión logística. Comando Validation (Stata). Final Master's Thesis, Complutense University of Madrid.2018, 54.

Garcia, J.; Gonzalez, E.; Riquelme, J.; Miranda, D.; Dieguez, U.; Navarro, R. Evolutionary feature selection to estimate forest stand variables using LiDAR. International Journal of Applied Earth Observation and Geoinformation. 2014, 26, 119–131.

García, D.. Estimación de variables de interés forestal basada en datos LiDAR en el monte número 117 del CUP Término municipal de Cuenca. Final Master's Thesis, Polytechnic University of Madrid. 2010, 194.

Gopalakrishnan, R.; Kauffman, J.; Fagan, M.; Coulston, J.; Thomas, V.; Wynne, R.; Quirino, V.. Creating landscape-scale site index maps for the southeastern US is possible with airborne LiDAR and Landsat imagery. Forests. 2019, 10, 234.

Gutiérrez, M.; Gallego, H.; Abril, A.; Martí, M.; Fernández, S.; Tejera, R.. Delineación de rodales para la ordenación forestal a partir de información LIDAR. In Proceedings of the 6th Congreso Forestal Español. Montes: Servicios y desarrollo rural, Vitoria, España, 10-14 June 2013; pp. 1-13.

Hirigoyen, A.; Varo, M.; Rachid, C.; Franco, J.; Navarro, R. Stand Characterization of Eucalyptus spp. Plantations in Uruguay Using Airborne Lidar Scanner Technology. Remote Sensing. 2020, 12, 3947.

Knapp, N.; Fischer, R.; Huth, A.. Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. Remote Sensing of Environment. 2017, 205, 199–209.

Leite, R.; Amaral, C.; Pires, R.; Silva, C.; Soares, C.; Macedo, R.; Leite, H.. Estimating stem volume in eucalyptus plantations using airborne LiDAR: A comparison of area-and individual tree-based approaches. Remote Sensing. 2020, 12, 1513.

Lekuona, I.; Lamelas, M.; Montealegre, A.. Cartografía de la biomasa aérea total en masas de *Pinus radiata* D. Don a partir de datos públicos LIDAR-PNOA e inventario forestal nacional. GeoFocus. 2017, 20, 87-107.

Logroño, S.; López, C.; Moyano, M.; Oyague, E.. El alcance de la teledetección satelital utilizando modelos estadísticos y físicos y sus beneficios en áreas contables. Dominio de las Ciencias. 2020, 6, 25-40.

Matasci, G; Hermosilla, T.; Wulder, M.; White, J.; Coops, N.; Hobart, G.;Bater, C.. Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. Remote Sensing of Environment. 2018, 216, 697-714.

Methol, R.. SAG Eucalyptus: sistema de apoyo a la gestión de plantaciones de Eucalyptus orientadas a la producción de celulosa en Uruguay. Instituto Nacional de Investigación Agropecuaria: Montevideo, Uruguay, 2008, pp. 1-26, ISBN 978-9974-674-11-0..

MGAP. Análisis sectorial y cadenas productivas. 2019. Available online: https://descargas.mgap.gub.uy/OPYPA/Anuarios/Anuario%202019/ORIGINAL%202019%20OPYPA%20INTERACTIVO%20agregado%2018-12-2019.pdf (accessed on 1 July 2021)

MGAP. Análisis sectorial y cadenas productivas. 2018. Available online: https://descargas.mgap.gub.uy/OPYPA/Anuarios/Anuario%202018/ANUARIO%20OPYPA%202018%20WEB%20con%20v%C3%ADnculo.pdf (accessed on 1 July 2021)

Moe, T.; Owari, T.; Furuya, N.; Hiroshima, T.. Comparing individual tree height information derived from field surveys, LiDAR and UAV-DAP for high-value timber species in Northern Japan. Forests. 2020, 11, 223.

Mora, B.; Wulder, MA; White, JC; Hobart, G.. Modeling stand height, volume, and biomass from very high spatial resolution satellite imagery and samples of airborne LiDAR. Remote Sensing. 2013, 5, 2308–2326.

Montealegre, A.; Lamelas, M.; de la Riva, J.; García, A.; Escribano, F.. "Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest". Forestry. 2016, 89, 373-382.

Noordermeer, L.; Gobakken, T.; Næsset, E.; Bollandsås, O.. Predicting and mapping site index in operational forest inventories using bitemporal airborne laser scanner data. Forest Ecology and Management. 2020, 457, 117768.

O'Hara, K.; Nagel, L.. The stand: Revisiting a central concept in forestry. Journal of Forestry. 2013, 111, 335–340.

Ortega, J.. Forest stand delineation through remote sensing and Object-Based Image Analysis. Final Master's Thesis, Univeristy of Gävle. 2018, 54.

Packalén, P.; Mehtätalo, L.; Maltamo, M.. ALS-based estimation of plot volume and site index in a eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. Annals of Forest Science. 2011, 68, 1085-1092.

Palop, E.; Bañuelos, M.; Quevedo, M.. Combinando datos LiDAR e inventario forestal para identificar estados avanzados de desarrollo en bosques caducifolios. Revista Ecosistemas. 2016, 25, 35-42.

Prodan, M.; Peters, R.; Cox, F.; Real, P.. In Mensura forestal. Serie Investigación y Educación en Desarrollo Sostenible. Instituto Interaméricano de Cooperación para la Agricultura, San José, Costa Rica, 1997; pp. 329-330, ISBN 92-90-39-304-1.

Pukkala, T.. Delineating forest stands from grid data. Forest Ecosystems. 2020, 7, 14.

Qgis, D.T. QGIS Geographic Information System, Open Source Geospatial Foundation. 2009. Available online: https://scholar.google.com/scholar_lookup?title=QGIS%20Geographic%20Information%20System&author=QGIS%20Development%20Team&publication_year=2009#d=gs_cit&u=%2Fscholar%3Fq%3Dinfo%3AK9pErfIkywoJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Des (accessed on 20 November 2020).

Quinn, G.; Keough, M.. Experimental Design and Data Analysis for Biologists; Cambridge University Press: Cambridge, England, 2002; pp. 1-572, ISBN 0-521-81128-7.

Sanchez, N.; Boschetti, L.; Hudak, A.. Semi-Automated Delineation of Stands in an Even-Age Dominated Forest: A LiDAR-GEOBIA Two-Stage Evaluation Strategy. Remote Sensing. 2018, 10, 1622.

Shao, G.; Gallion, J.; Saunders, M.; Frankenberger, J.; Fei, S. Improving Lidar-based aboveground biomass estimation of temperate hardwood forests with varying site productivity. Remote Sensing of Environment. 2018, 204, 872–882.

Silva, C.; Klauberg, C. Hudak, A.; Vierling, L.; Liesenberg, V.; Bernett, L.; Schoeninger, E.. Estimating Stand Height and Tree Density in *Pinus taeda* plantations using in-situ data, airborne LiDAR and k-Nearest Neighbor Imputation. Anais da Academia Brasileira de Ciências. 2018, 90, 295-309.

Socha J.; Hawryło P.; Stereńczak K.; Miścicki S.; Tymińska-Czabańska L.; Młocek, W.; Gruba P. Assessing the sensitivity of site index models developed using bi-temporal airborne laser scanning data to different top height estimates and grid cell sizes. International Journal of Applied Earth Observation and Geoinformation. 2020, 91, 102-129.

Varo, M.; Navarro, R.; Hernández, R.. Semi-automated stand delineation in Mediterranean Pinus sylvestris plantations through segmentation of LiDAR data: The influence of pulse density. International Journal of Applied Earth Observation and Geoinformation. 2017, 56, 54–64.

Varo, M. and Navarro, R.. Stand Delineation of Pinus sylvestris L. Plantations Suffering Decline Processes Based on Biophysical Tree Crown Variables: A Necessary Tool for Adaptive Silviculture. Remote Sensing. 2021, 13, 436.

Verma, N.; Lamb, D.; Sinha, P.. Airborne LiDAR and high resolution multispectral data integration in Eucalyptus tree species mapping in an Australian farmscape. Geocarto International. 2019, 21.

White, C., Wulder, A., Varhola, A., Vastaranta, M., Coops, N., Cook, B., Pitt, D., Woods, M.. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. The Forestry Chronicle. 2013, 89, 722-723.

Weisberg, S.. Applied linear regression. University of Minnesota press: Minnesota, USA, 2005; pp. 1-317, ISBN 0-471-66379-4.

Wu, Z.; Heikkinen, V.; Hauta-Kasari, M.; Parkkinen, J.; Tokola, T.. Forest Stand Delineation Using a Hybrid Segmentation Approach Based on Airborne Laser Scanning Data. Scandinavian Conference on Image Analysis. 2013, 95–106.

Xu, C.; Manley, B.; Morgenroth, J.. Evaluation of modelling approaches in predicting forest volume and stand age for small-scale plantation forests in New Zealand with RapidEye and LiDAR. International Journal of Applied Earth Observation and Geoinformation. 2018, 73, 386-396.

Yang, L.; Mansaray, LR; Huang, J.; Wang, L. Optimal segmentation scale parameter, feature subset and classification algorithm for geographic object-based crop recognition using multisource satellite imagery. 2019, 11, 514.

Zald, H.; Wulder, M.; White, J.; Hilker, T.; Hermosilla, T.; Hobart, G.; Coops, N.. Integrating Landsat pixel composites and change metrics with lidar plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada. Remote Sensing of Environment. 2016, 176, 188–201.

7. Annexes

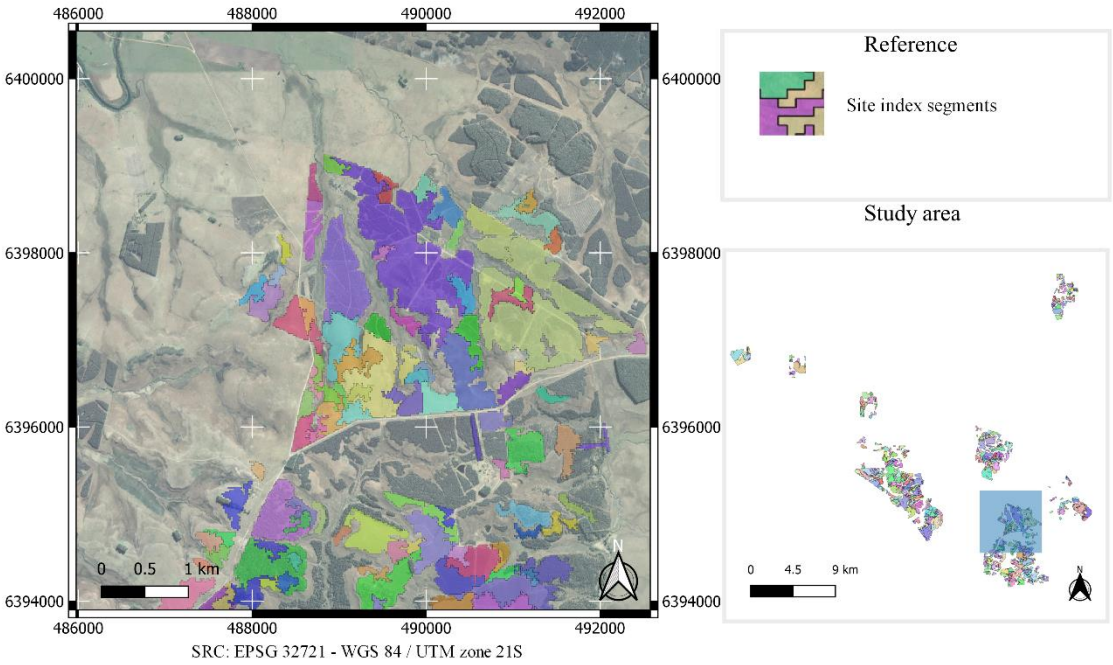

Figure 6. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 1).

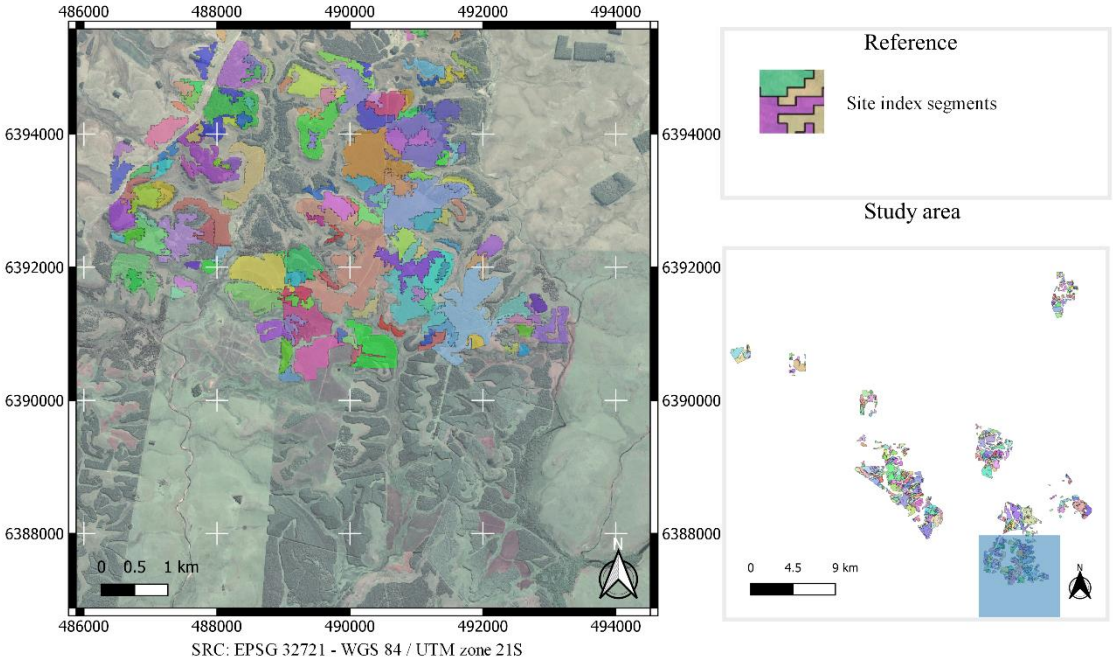

Figure 7. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 2).

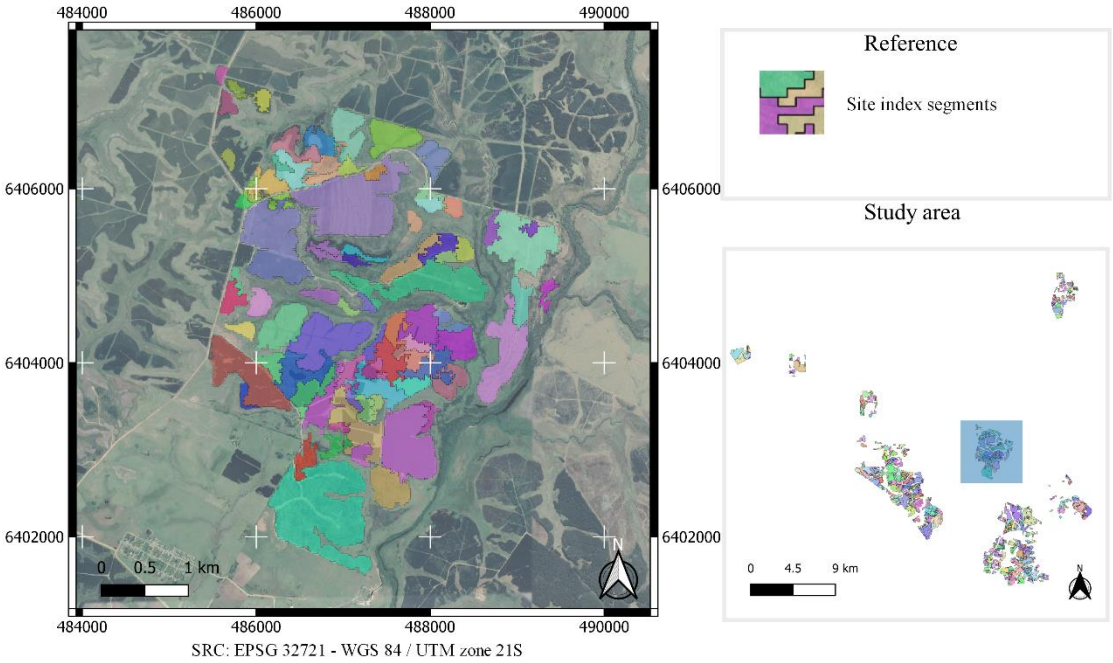

Figure 8. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 3).

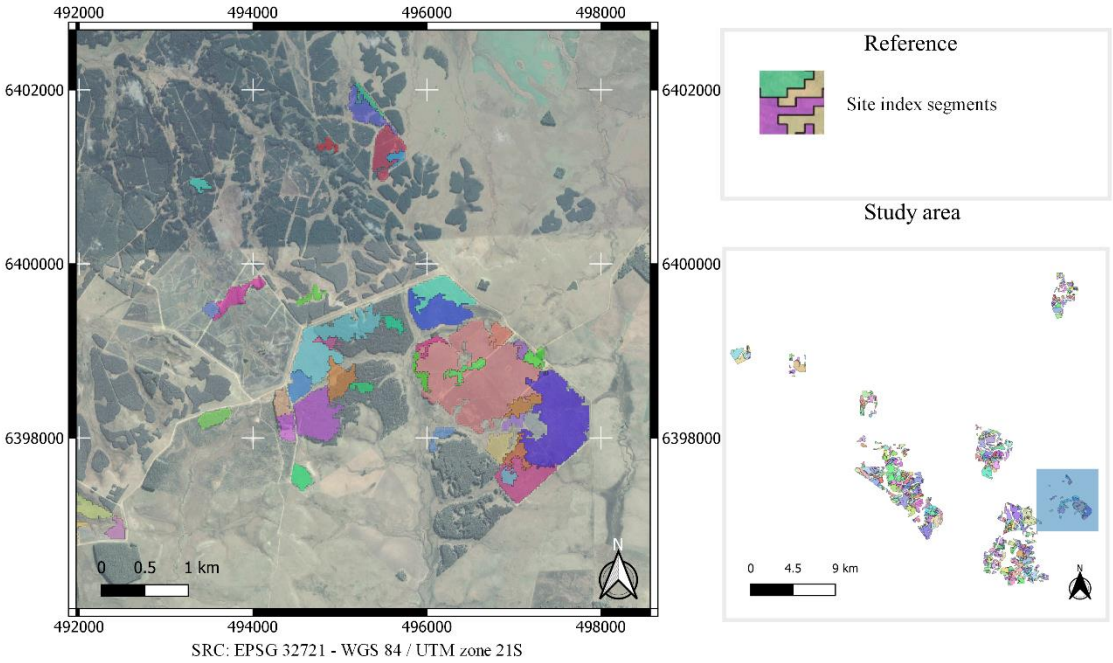

Figure 9. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 4).

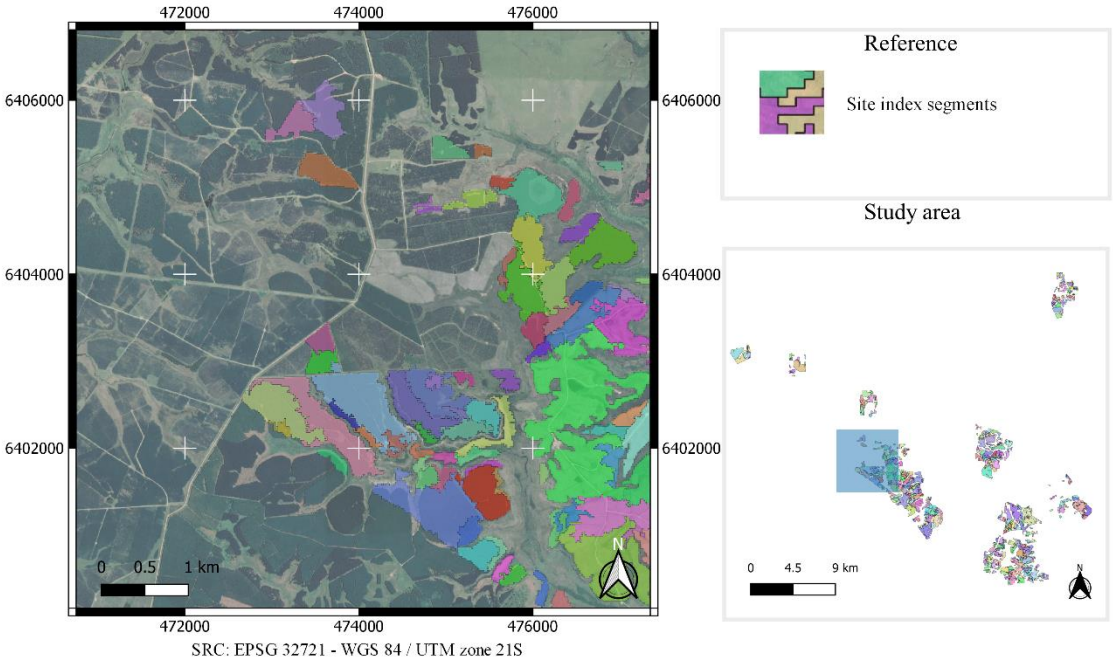

Figure 10. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 5).

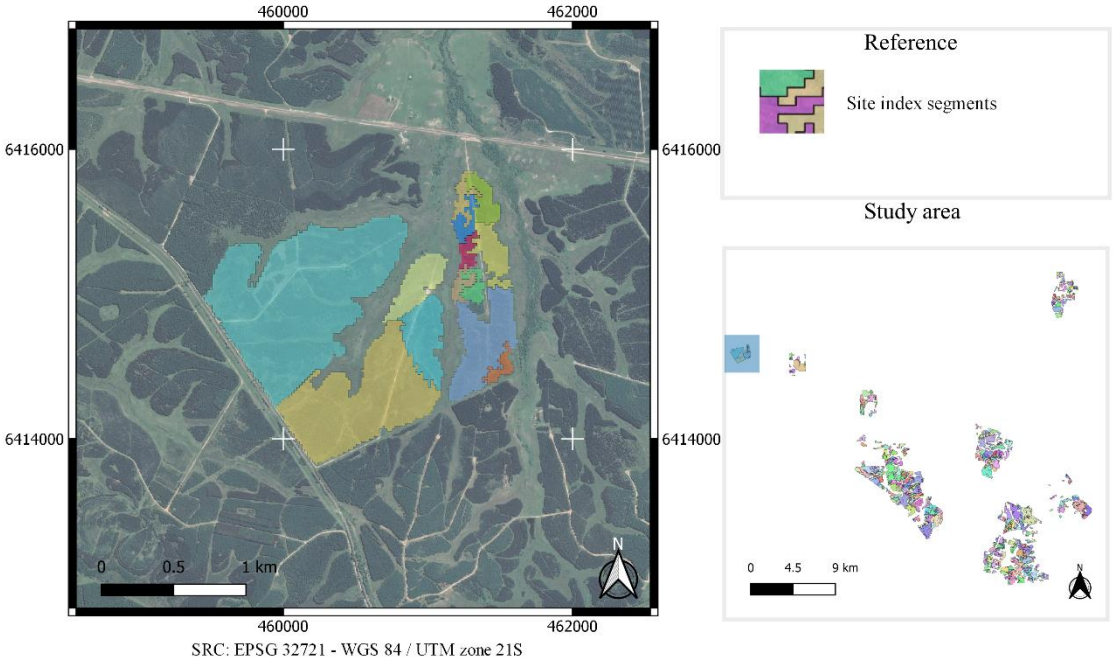

Figure 11. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 6).

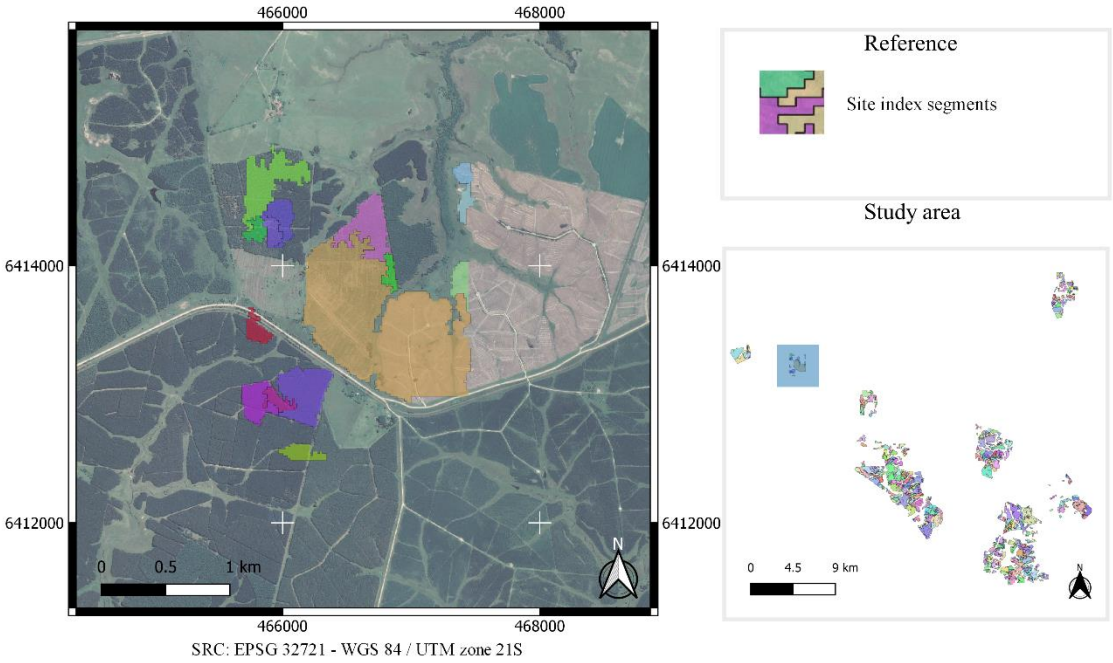

Figure 12. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 7).

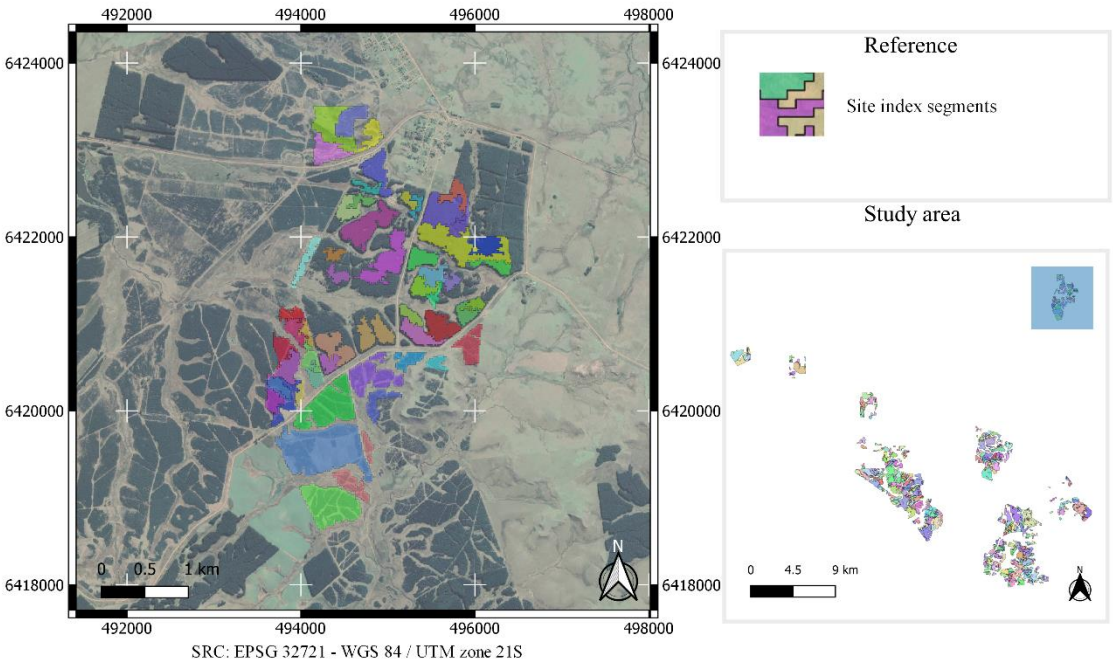

Figure 13. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 8).

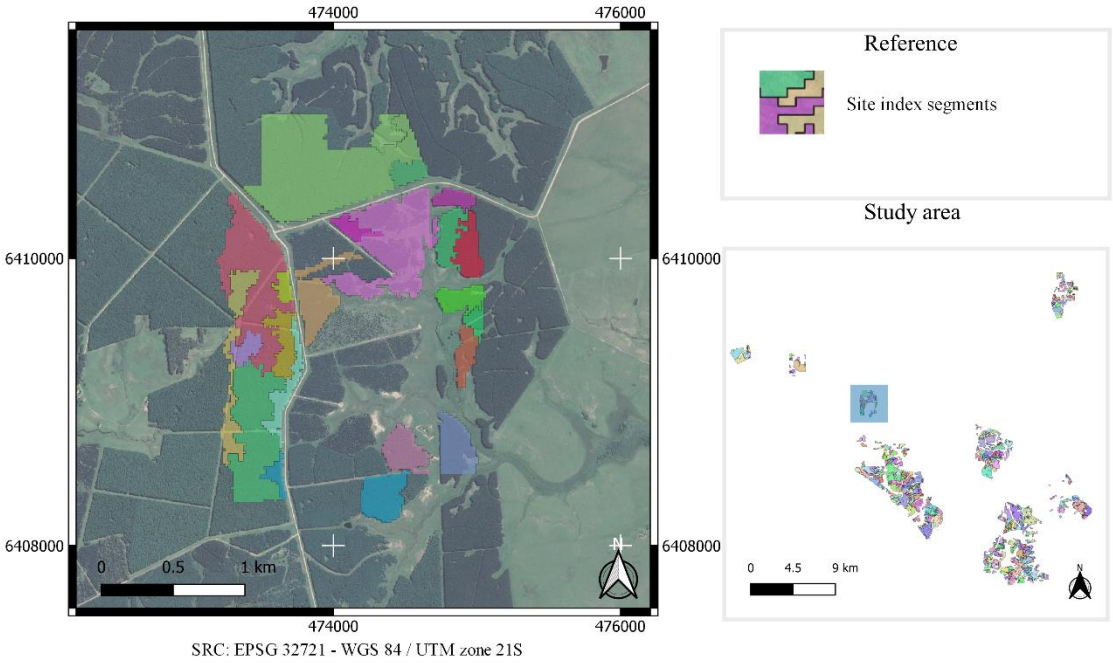

Figure 14. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 9).

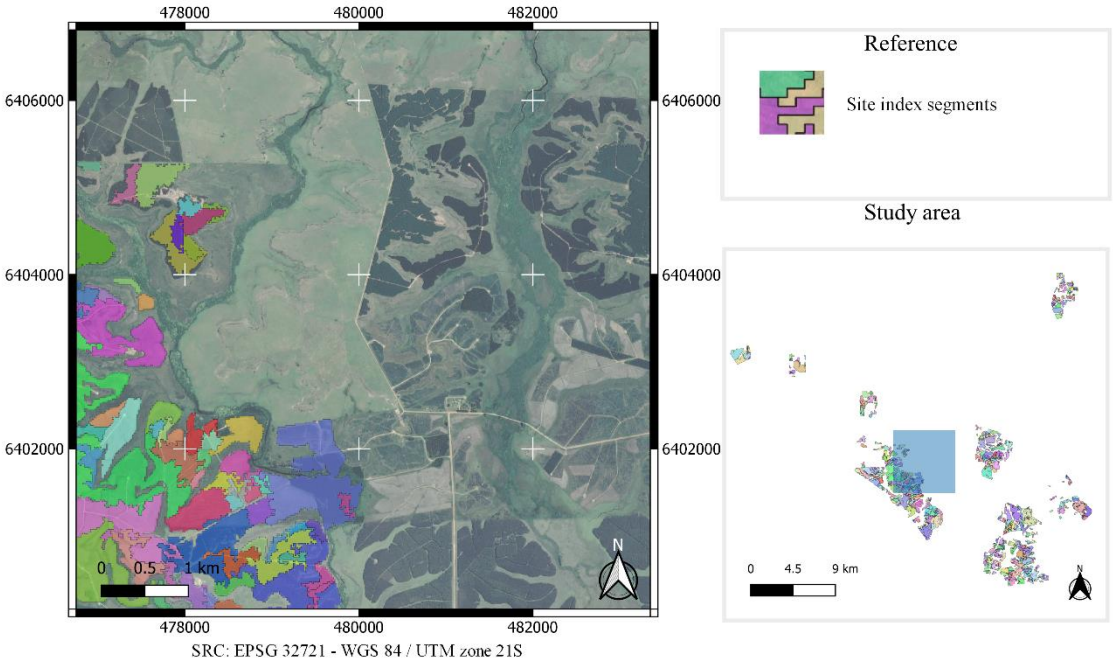

Figure 15. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 10).

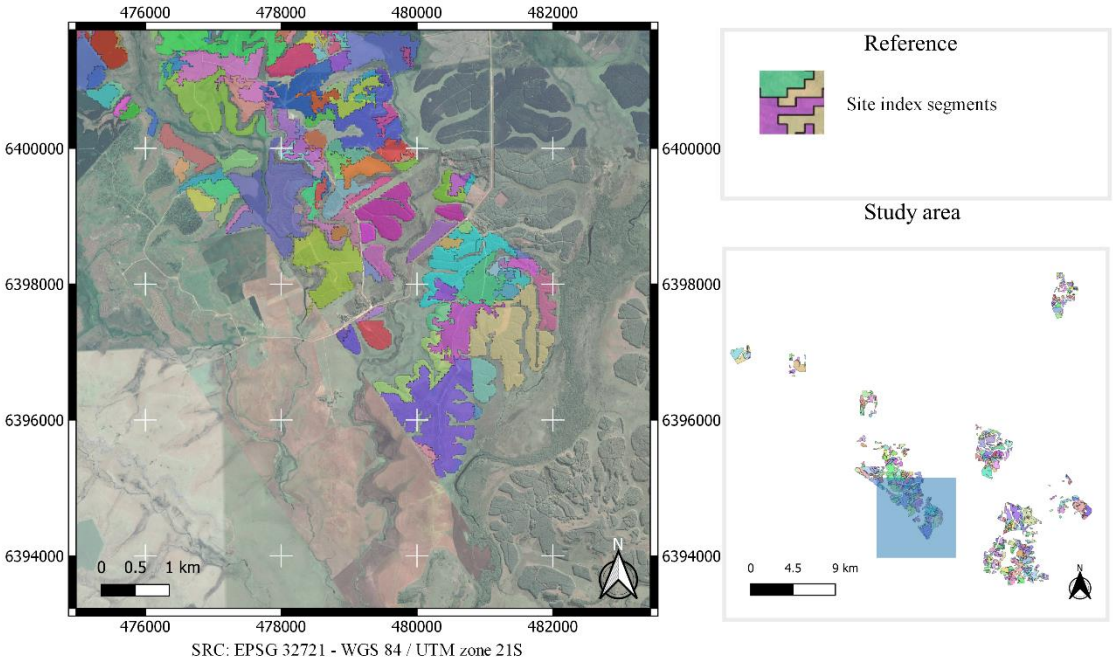

Figure 16. Detail of segments on Google satellite image of *Eucalyptus dunnii* plantations using the site index (part 11).