UNIVERSIDAD DE CONCEPCIÓN
Dirección De Postgrado
Facultad de Ciencias Forestales
Programa de Magíster de Ciencias Forestales

Efecto de densidad de puntos LiDAR y tamaño de parcela en modelación de variables dasométricas en rodales de *Pinus radiata*.

Effect of LiDAR point density and plot size on modeling of dasometric variables in Pinus radiata stands.

Tesis presentada a la facultad de Ciencias Forestales de la Universidad de Concepción para optar al grado de Magíster en Ciencias Forestales

Por

Marcelo López González

Profesor Guía: Dr. Simón Sandoval Rocha

Depto. Manejo de Bosque y Medio Ambiente

Facultad de Ciencias Forestales

Universidad de Concepción

Concepción, Chile 2023

Efecto de densidad de puntos LiDAR y tamaño de parcela en modelación de variables dasométricas en rodales de *Pinus radiata*.

Comisión Evaluadora:

Simón Sandoval Rocha
(Profesor guía)
Ingeniero Forestal, Dr. ____________________________

Eduardo Acuña Carmona
(Comisión evaluación)
Ingeniero Forestal, Dr. ____________________________

Julio Becker Guaiquil
(Comisión evaluación)
Ingeniero Forestal, Mg. ____________________________

Santiago Yepez Figueroa
(Comisión evaluación)
Ingeniero Geólogo, Dr. ____________________________

Director de Postgrado:

Regis Teixeira Mendonça
Ingeniero Químico, Dr. ____________________________

Decano Facultad de Ciencias Forestales:

Eugenio Sanfuentes Von Stowasser.
Ingeniero Forestal, Dr. ____________________________

DEDICATORIA

A mi esposa y mis padres.

AGRADECIMIENTOS

A mi familia, en especial a mi esposa por estar presente en la continuación de mis estudios, me apoyó y estuvo a mi lado en todas las decisiones que elegí durante este proceso.

Agradecer a la empresa CMPC, por el apoyo y auspicio de la información para el desarrollo de esta investigación.

Al Dr. Simón Sandoval por su buena disposición, me guío y ayudó a desarrollar mi trabajo, y además me instó a continuar en el área de la investigación.

TABLA CONTENIDOS

DEDICATORIA iii
AGRADECIMIENTOS iv
RESUMEN vi
ABSTRACT vii
INTRODUCTION 1
METHODOLOGY 4
RESULTS 8
DISCUSSION 14
CONCLUSION 16
REFERENCES 17

RESUMEN

Los inventarios forestales basados en tecnología LiDAR ofrecen ventajas significativas en términos de costos y tiempos de ejecución en comparación con los métodos tradicionales. El método de masa o ABA (Area Based-Approach) es una metodología comúnmente utilizada en los inventarios LiDAR, la cual requiere variables de entrada, como datos del terreno y nubes de puntos LiDAR, para modelar características dasométricas clave, incluyendo altura dominante, área basal, densidad de rodal y volumen. La precisión de estos predictores dasométricos obtenidos a través del enfoque ABA depende en gran medida de la calidad de los datos de entrada y las condiciones del área de estudio, especialmente la densidad del rodal. Por lo tanto, es crucial determinar tamaños de parcelas de muestreo adecuados para diferentes condiciones de rodal, con el fin de obtener estimaciones dasométricas confiables. Además, es esencial evaluar el impacto de la densidad de puntos LiDAR utilizada en el análisis, considerando que un aumento excesivo en la densidad de puntos conlleva un incremento en los costos operativos asociados a los inventarios LiDAR.

El objetivo de este estudio es investigar el efecto de la densidad de puntos LiDAR, el tamaño de la parcela de terreno y las condiciones del rodal en la modelación dasométrica utilizando el enfoque ABA. El trabajo se dividirá en tres fases. En la primera fase, se evaluarán dos condiciones de rodal, utilizando dos tamaños de parcelas de terreno, junto con una densidad de nube LiDAR de 400 puntos/m2. El objetivo es seleccionar las métricas más adecuadas para obtener modelos precisos de las variables dasométricas de interés utilizando un enfoque de regularizacion tipo LASSO. En la segunda fase, se analizará el efecto de la densidad de puntos LiDAR. Para ello, se reducirá de forma sistemática la densidad inicial de puntos en porcentajes desde el 80% hasta el 1%, generando un total de diez condiciones de densidad de puntos distintas. Se espera identificar un punto de inflexión en los resultados de las variables dasométricas, determinando así la densidad de puntos LiDAR óptima para realizar inventarios forestales sin incurrir en costos operativos excesivos asociados a densidades elevadas. Finalmente, la tercera fase tendrá como objetivo realizar modelación con el método SUR (regresión aparentemente no relacionada), metodología nueva en el área para la estimación de variables dasométricas con informacion LiDAR.

Este estudio contribuirá al avance de la metodología de inventarios forestales basados en LiDAR, al proporcionar recomendaciones prácticas para la selección adecuada de densidades de puntos de nubes y tamaños de parcelas de terreno, mejorando la eficiencia y precisión en la modelación de inventarios forestales.

ABSTRACT

Forest inventories based on LiDAR technology offer significant advantages in terms of cost and execution time compared to traditional methods. The Area Based-Approach (ABA) method is a commonly used methodology in LiDAR inventories, which requires input variables, such as terrain data and LiDAR point clouds, to model key dasometric characteristics, including dominant height, basal area, stand density and volume. The accuracy of these dasometric predictors obtained through the ABA approach is highly dependent on the quality of the input data and the conditions of the study area, especially stand density. Therefore, it is crucial to determine adequate sample plot sizes for different stand conditions to obtain reliable dasometric estimates. In addition, it is essential to evaluate the impact of the LiDAR point density used in the analysis, considering that an excessive increase in point density leads to an increase in the operational costs associated with LiDAR inventories.

The objective of this study is to investigate the effect of LiDAR point density, plot size and stand conditions on dasometric modeling using the ABA approach. The work will be divided into three phases. In the first phase, two stand conditions will be evaluated, using two plot sizes, along with a LiDAR cloud density of 400 points/m². The objective is to select the most appropriate metrics to obtain accurate models of the dasometric variables of interest using a LASSO-type regularization approach. In the second phase, the effect of LiDAR point density will be analyzed. For this purpose, the initial point density will be systematically reduced in percentages from 80% to 1%, generating a total of ten different point density conditions. It is expected to identify a turning point in the results of the dasometric variables, thus determining the optimal LiDAR point density for forest inventories without incurring excessive operational costs associated with high densities. Finally, the third phase will aim to perform modeling with the SUR (seemingly unrelated regression) method, a new methodology in the area for the estimation of dasometric variables with LiDAR information.

This study will contribute to the advancement of LiDAR-based forest inventory methodology by providing practical recommendations for the proper selection of cloud point densities and plot sizes, improving the efficiency and accuracy of forest inventory modeling.

INTRODUCTION

Currently, traditional forest inventories are widely used as the primary tool for physically assessing stands. These inventories provide information for silvicultural management planning and forest harvesting (Johnson et al. 2007). However, conducting these traditional inventories involves significant costs and execution times, especially in areas that are difficult to access and travel.

In recent years, a new methodology has emerged for conducting forest inventories using LiDAR (Light Detection and Ranging) technology. This technique is based on remote sensing to assess stands. LiDAR is a type of active remote sensing that emits beams of light towards objects and measures the time it takes to receive the light returns (Carter et al. 2012). These sensors can be mounted on manned or unmanned aerial vehicles such as drones, small planes or helicopters, they are known as ALS (Airborne Laser Sensor). During the overflight of the study area, they capture information in the form of a point cloud. This technique allows obtaining data more efficiently in terms of time and costs compared to traditional inventories in large stands with difficult accessibility (Zhang et al. 2016).

Forest inventories using LiDAR technology improve the efficiency of data capture in spatial and temporal terms. One of the notable advantages is their ability to cover the entire study area, which allows observing the spatial variability of the data. This contrasts with traditional inventories, where the sample is reduced to the plot level (Goodbody et al. 2017). In terms of safety, LiDAR inventories are safer, as they require fewer personnel in the forest. However, to implement these inventories, users must have additional skills, such as programming, knowledge of geographic information systems, statistics, and, above all, understanding of forest behavior. These skills are necessary to analyze and interpret the information obtained from the LiDAR point cloud.

One of the methods used in LiDAR forest inventories is the Area Based-Approach (ABA). This method involves establishing a relationship between forest variables obtained in the field and metrics derived from the information captured by the LiDAR cloud, with the objective of obtaining forest estimators for a specific area. These metrics, geometric in nature, are values and statistics that represent the vertical distribution of forest LiDAR returns, such as percentiles (White et al. 2013). The relationship between terrain data and LiDAR metrics is established by regression models, using approaches such as k-nearest neighbor (K-NN) (Næsset 2002) or by machine learning techniques, such as artificial neural networks and random forest (Corte et al. 2020).

The dasometric variables of interest do not present a uniform behavior when using the ABA method for their estimation, the dominant height estimation is better related to the LiDAR cloud metrics than the stand density estimation. Treitz et al. (2012) obtained an RMSE value of 3.9% when estimating height in a conifer forest with a density of 300 to 500 trees/ha, using a linear regression model and LiDAR data of 3.2 points/m². Similarly, Silva et al. (2018) obtained comparable results for height in a Pinus Taeda plantation in southern Brazil, with a density of 1667 to 2000 trees/ha, obtaining an RMSE of 5.7% using 4 points/m² and a nonparametric nearest neighbor (K-NN) model. In both studies, it was highlighted that stand height is the easiest variable to predict, as it shows a high correlation using only LiDAR metrics of height distribution, such as the 95th percentile (González-Ferreiro et al. 2012, Strunk et al. 2014, Varo-Martínez et al. 2017). In contrast, the dasometric variable of stand density presents a significant challenge. Rahlf et al. (2015) obtained an RMSE value of 51% in a conifer forest in southeastern Norway, with a variable stand density from 40 to 4100 trees/ha, using LiDAR data of 1.2 points/m² and a multivariate nearest neighbor (K-NN) model. López and Sandoval (2023) reported improvements by incorporating a mixed method between ITD and ABA in the estimation of stand density with an RMSE of 20.9%. Silva et al. (2018), obtained a value of R2=0.38 when modeling the stand density variable, which confirms its difficulty in predicting it. From these results, it can be interpreted that the performance of the estimations depends on the one hand on the characteristics of the stands and on the quality of the LiDAR information.

One of the crucial input data for the ABA workflow is the data obtained from the field plots. The quality of the data collection has a direct impact on the accuracy of the dasometric predictors of interest. An important factor to consider within this workflow is the accuracy of the plot center since a higher error in the plot center accuracy can result in mismatches when crossing the terrain information with the LiDAR information cloud. Mauro et al. (2011) determined that, when working with plots with a radius greater than 10 meters in terrain, the effects of positioning errors were negligible when studying forest tree height distributions. Furthermore, Ruiz et al. (2014) indicate that plots of at least 500-600 m² are necessary to estimate volume, biomass and basal area, and 300-400 m² for canopy cover in a low stand density conifer forest in Spain. They also point out that increasing plot size does not significantly improve the accuracy of the models and increases field operational costs.

Another factor that influences plot size variation is the edge effect. This refers to situations in which the stem of an individual is inside the plot, but part of its crown extends outside the plot; or conversely, when an individual is outside the plot, but part of its crown is inside the plot. This generates an error in accuracy when comparing the data obtained in the field and the information provided by the LiDAR cloud (Mascaro et al. 2011, Packalen et al. 2015). Studies by Frazer et al. (2011) indicate that plot size has an impact on the accuracy of estimating forest structure attributes. The researchers found that larger plot areas provide greater precision in such estimation, as the edge effect is significantly reduced.

In addition to considering plot size in field work to mitigate problems in data capture, it is crucial to consider the condition of the stand being studied. These effects may vary in different proportions depending on the stand density. It has been observed that forests with a higher tree density present predictor with a higher error in terms of RMSE compared to lower density forests (Belmonte et al. 2019). However, this increase in the error of the dasometric predictors is not only related to stand density, but also to the density of the LiDAR information cloud used. In forests with a high density of trees per hectare and when using a low density point cloud, problems arise when generating the digital terrain model. The limited number of points in the cloud is not sufficient to penetrate the coverage of the forest individuals, leading to an incorrect definition of the terrain surface. This problem is especially exacerbated in terrain with steep topography, resulting in an overestimation or underestimation of the heights of individuals when modeling (White et al. 2013).

Thus, the point density of the LiDAR cloud is a crucial factor in the ABA workflow, as it influences the metrics obtained during the modeling phase. It has been determined that working with clouds having a point density between 0.5 to 10 points per square meter (points/m²) does not generate large differences in the accuracy of dasometric predictors (Tompalski et al. 2019). In fact, it has been recommended to work with clouds having a density close to 1 point/m² to obtain good results (Jakubowski et al. 2013, Magnussen et al. 2010). Sánchez et al. (2018) managed to obtain accurate predictors in sclerophyll forests in southern Spain using a density of only 0.5 points/m² in a forest of Pinus sylvestris L. A similar case was reported by Yoga et al. (2017), who obtained predictors with low estimation error in a coniferous forest in southeastern Canada using 6 points/m². In both cases, despite having low resolution LiDAR information, the good modeling results were due to favorable stand conditions, in terms of low stand density.

Other studies have found significant differences in biomass, volume and stand density estimation when reducing LiDAR cloud point density in forests with higher tree density and irregular vertical structure, such as tropical forests (González-Ferreiro et al. 2012, Magnussen et al. 2010, Manuri et al. 2017). One of the possible factors contributing to this variation in results when using different point cloud densities is the impact on the metrics obtained from the LiDAR cloud, which are affected by the variation in point density (Roussel et al. 2017). It has been observed that the maximum height metric shows larger variations when the cloud density decreases (Gobakken and Næsset 2008, Hansen et al. 2015). However, these variations in metrics also depend on the size of the LiDAR flight footprint and the shape of the stand canopy. In other words, as a stand becomes more homogeneous, the differences in results between different LiDAR cloud densities will be smaller (Roussel et al. 2017).

Using a LiDAR point density greater than 10 points/m² is considered high resolution, leading to more demanding computational analysis and increased costs within the workflow. Most studies have worked with this point density, but today, with the increasing accessibility of equipment with higher computational power and the development of new efficient software and programming codes, it is possible to analyze large amounts of information more effectively. It is essential to understand the importance of the quality of the input information in the ABA workflow, encompassing both terrain data and LiDAR point cloud. In addition, it is crucial to consider the relationship between these information factors and stand density. Stand density should be addressed as an additional factor when developing the modeling strategy and applying it operationally in forest inventories. The main objective of this research is to evaluate the plot size within the operational forest area and the influence of LiDAR point cloud density on the estimation of dasometric variables. The aim is to identify if there is a trend or an inflection point that allows improving the accuracy of the predictive variable models without the need to excessively increase the density of the cloud, thus avoiding higher operational costs.

METHODOLOGY

Area of study

The study was carried out in two forestry estates of the CMPC company in Chile, both of which are *Pinus radiata* plantations. The Coihueco property has an area of 49 ha and a stand density of approximately 770 trees/ha for pulp production. The Nihuinco property has an area of 60 ha and a stand density of approximately 400 trees/ha for the production of debobinable and sawmill products.

Data capture

The analysis design contemplated the evaluation of three factors. First, the two forest properties with different silvicultural treatments were considered, highlighting the difference in stand density, with the Coihueco property having a higher stand density (~770 trees/ha) than the Nihuinco property (~400 trees/ha). The second factor evaluated was the size of the sampling plots, where concentric plots of 300 and 500 m² were established, dimensions used in traditional forest inventory operations. The third and last factor evaluated was the density of points in the LiDAR information cloud. Both properties were obtained a LiDAR cloud with an approximate density of 400 points/m², which was reduced by 80%, 60%, 50%, 40%, 40%, 30%, 20%, 20%, 15%, 10%, 5% and 1% of the total number of points. Each reduction process was performed randomly and each of these processes was repeated 100 times.

In each of the properties, plots were established and systematically located according to the operational prescription of the forest inventory. In Coihueco 23 plots were established and in the Nihuinco property 36 plots were established. In each of the plots, the DBH and height of all the trees were recorded, and the information was subsequently processed to estimate.

LiDAR data análisis

The LiDAR information of each stand was processed using the ABA method workflow. This process was carried out using R software and the *LidR* library (Roussel et al. 2021). A cleaning and filtering of the point cloud was performed to obtain the position of the first returns in each study stand. From this, a classification of the points was made identifying soil and vegetation, generating the digital terrain model (DTM) and the digital surface model (DSM). Subsequently, the point cloud was normalized for its projection to the horizontal axis. Once the LiDAR information of both properties was processed, the information of each plot measured in the field was extracted according to the coordinate of the plot center and the radius corresponding to the surfaces of 300 and 500 m². Once the LiDAR information was obtained for both types of plots in the two properties, the set of standard LiDAR metrics was determined (Table 1). The LiDAR metrics were obtained for each set of the reduced point cloud and for each of the 100 iterations performed.

Table 1. Standard canopy height metrics obtained with LiDAR.

Metrics	Description
Zmax	Maximum height
Zskew	Asymmetry of height distribution
Zkurt	Kurtosis of height distribution
Zentropy	Entropy of height distribution
Zmean	Mean height
Zsd	Standard deviation of height distribution
Zabove2	Percentage of returns over 2 m
Zabovezmean	Percentage of returns over mean height
Zq_5, …, Zq_{95}	Percentile value of the height distribution between 5 to 95%.
$Zpcum_1$, …, $Zpcum_9$	Cumulative percentage of the cloud divided in 10 equal parts

Base modeling and selection of metrics

The modeling was divided into three stages. In the first stage, a linear model was performed (Equation 1), where the metrics obtained from 100% of the LiDAR cloud information, selected with the LASSO selection method, were used to estimate the dasometric variables of dominant height (DH), basal area (BA), stand density (N) and volume (V). This linear model presented a matrix with the information of the metrics in the two stands with different conditions and with different plot size used for the extraction of this information (300 and 500 m²).

$$\hat{y}_k = f\left(X_{\in N} \mid \hat{\beta}_{\in n}\right) \qquad \text{Equation 1.}$$

Where ŷk represents the estimated stand dasometric variable, X is the matrix of metric values corresponding to the stand and plot size, and β are the parameters associated with the metrics used in the independent model.

As mentioned above, a selection of metrics is made for each model through LASSO type regularization (Equation 2), according to plot size and by stand. The LASSO regularization system penalizes the least relevant metrics, controlled by the hyperparameter λ. As λ increases, the penalty increases and more metrics are excluded in the model estimation of dasometric variables (James et al. 2013). The independent models fitted for each stand and plot size obtained will be compared using error indicators such as root mean square error (RMSE), Akaike index (AIC) and R^2.

$$\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda\sum_{j=1}^{p}|\beta_j|$$

Equation 2.

Where yi represents the estimated dasometric variable of the stand in the i-th plot measured in the field, β0 and βj are the linear parameters of the model, xji is the j-th metric selected from a total of J-metrics in the i-th plot measured in the field, and λ is a hyperparameter of the Lasso method.

Independent modeling with variation in LiDAR point density

Once the model results were obtained with the selection of metrics based on the LiDAR cloud from the first stage, the effect of the density of the point cloud on the models adjusted independently for DH, BA, N and V was evaluated. Thus, this second stage of modeling was performed using the LiDAR metrics obtained for each stand and plot size, systematically decreasing the total percentage of the cloud, generating an iteration of 100 random point selections for percentage reduction. This resulted in a total of 10 point density conditions, ranging from 80% to 1% with a systematic reduction of the percentage of the point cloud in each plot (Figure 1).

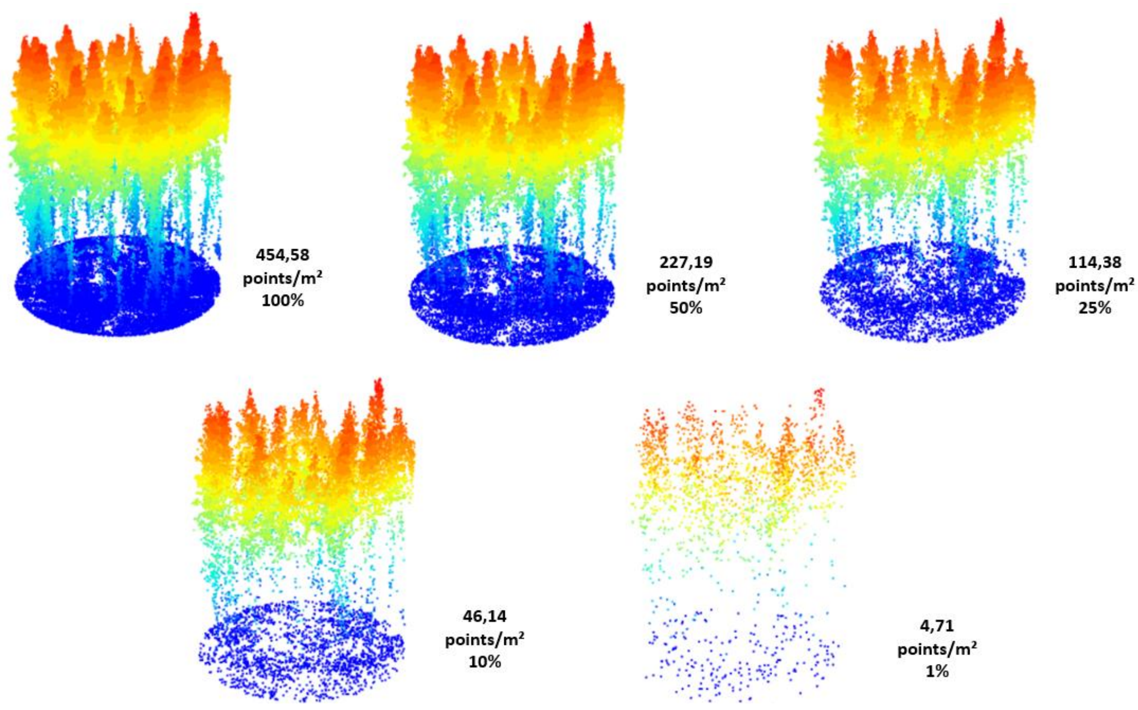

Figure 1. Example of a LiDAR point cloud systematics decrease of a 500 m² plot of the Coihueco property.

The objective of this second stage was to evaluate the effect of LiDAR point cloud density by incorporating the LiDAR metrics obtained in each case as predictors in linear models of

DH, BA, N and V adjusted independently. The effect was evaluated in relation to the residual root mean square (RMSE) obtained in each fit, i.e. for each stand and plot size.

Simultaneous modeling with LiDAR point density variation

In the third stage, a simultaneous modeling strategy was performed for each stand and plot size. Here, the modeling of DH, BA, N and V generated a single system of equations, a method known as SUR (Seemingly unrelated regressions). This method considers the relationship between the errors of each model to obtain more consistent parameters.

$$\begin{bmatrix} \hat{y}_{DH} \\ \hat{y}_{BA} \\ \hat{y}_{N} \\ \hat{y}_{V} \end{bmatrix} = \begin{bmatrix} X_{red1} & 0 & 0 & \cdots & 0 \\ 0 & X_{red2} & 0 & \cdots & 0 \\ 0 & 0 & X_{red3} & \cdots & 0 \\ 0 & 0 & 0 & \cdots & X_{redn} \end{bmatrix} + \begin{bmatrix} \beta_{DH} \\ \beta_{BA} \\ \beta_{N} \\ \beta_{V} \end{bmatrix} + \begin{bmatrix} \epsilon_{DH} \\ \epsilon_{BA} \\ \epsilon_{N} \\ \epsilon_{V} \end{bmatrix}$$

Equation 3.

Where ŷ represents the estimated dasometric variable of the stand (DH, BA, N and V), Xred is the matrix of metric values according to stand and plot size for each model obtained from the Lasso selection in the second modeling stage; β is the vector of parameters associated to the metrics used in the models and ε is the error of each model. To evaluate the inflection point where the RMSE of the models changes significantly in relation to the LiDAR cloud point density, a segmented modeling strategy was used using the *Segmented* library of the R software (Muggeo 2023).

RESULTS

The first stage of the modeling of the dasometric variables DH, BA, N and V used the full density of the LiDAR information (~400 points/m²). In most cases in the 500 m² plots, better indicators of quality of fit were observed according to RMSE, AIC and R2 values (Table 2). Only in the Coihueco property did the BA modeling show higher RMSE values, when comparing the information from 300 and 500 m² plots, increasing from 6.49 to 6.64%, respectively. In the average adjustment models, when information from both plots was used, the same tendency was observed with the 500 m² plots in relation to 300 m². The BA and V variables showed the greatest difference, improving the RMSE from 14.72 to 10.60% and 14.33 to 9.53% in the 300 to 500 m² plots, respectively.

As for the AIC values, they showed a similar behavior to the RMSE (%), being better when plots of size 500 m² were used. The DH variable presented a greater decrease in this value in the models adjusted independently for each stand and in the average adjustment of both stands. Only one exception was observed in the estimation of AB and N, where the AIC value increased in the Coihueco stand from 67.53 to 69.17 and 142.43 to 146.88, respectively. Regarding R^2, in the average adjustment, improvements were observed in all the models that used the information from the 500 m² plots; the model for N achieved a value of 0.88. The Coihueco stand, which has an average stand density of 770 trees/ha, presented higher R^2 values than the Nihuinco property, which has a density of 400 trees/ha in the estimation of all the dasometric variables.

Table 2. Results of RMSE, AIC and R^2 indicators of the models according to site and plot size for the estimation of dominant height, basal area, stand volume and stand density.

		RMSE %				AIC				R^2			
stands	plot (m²)	DH	BA	N	V	DH	BA	N	V	DH	BA	N	V
Nihuinco	300	3,39	13,69	17,48	15,15	60,69	115,14	209,60	205,87	0,31	0,33	0,15	0,30
	500	2,06	9,55	13,52	10,21	42,58	103,05	206,93	192,32	0,33	0,54	0,38	0,54
Coihueco	300	0,78	6,49	10,34	9,44	25,52	67,53	142,43	126,62	0,94	0,76	0,75	0,59
	500	0,72	6,64	10,07	6,34	19,78	69,17	146,88	124,83	0,94	0,67	0,63	0,75
Average	300	2,94	14,72	16,88	14,33	95,71	199,18	376,13	331,90	0,64	0,61	0,82	0,46
	500	1,84	10,60	16,09	9,53	65,47	185,56	372,25	312,08	0,82	0,77	0,88	0,68

Figure 2 shows the relationship between the measured dasometric variables and their estimates when using 100% of the LiDAR information, with average adjustment models. The models generate consistent estimates for the four variables evaluated and for both plot sizes. Consistent with the interpretation of the goodness-of-fit indicators, the predicted values of the four variables modeled using the information from the 500 m² plots fit better than the values estimated with a plot size of 300 m².

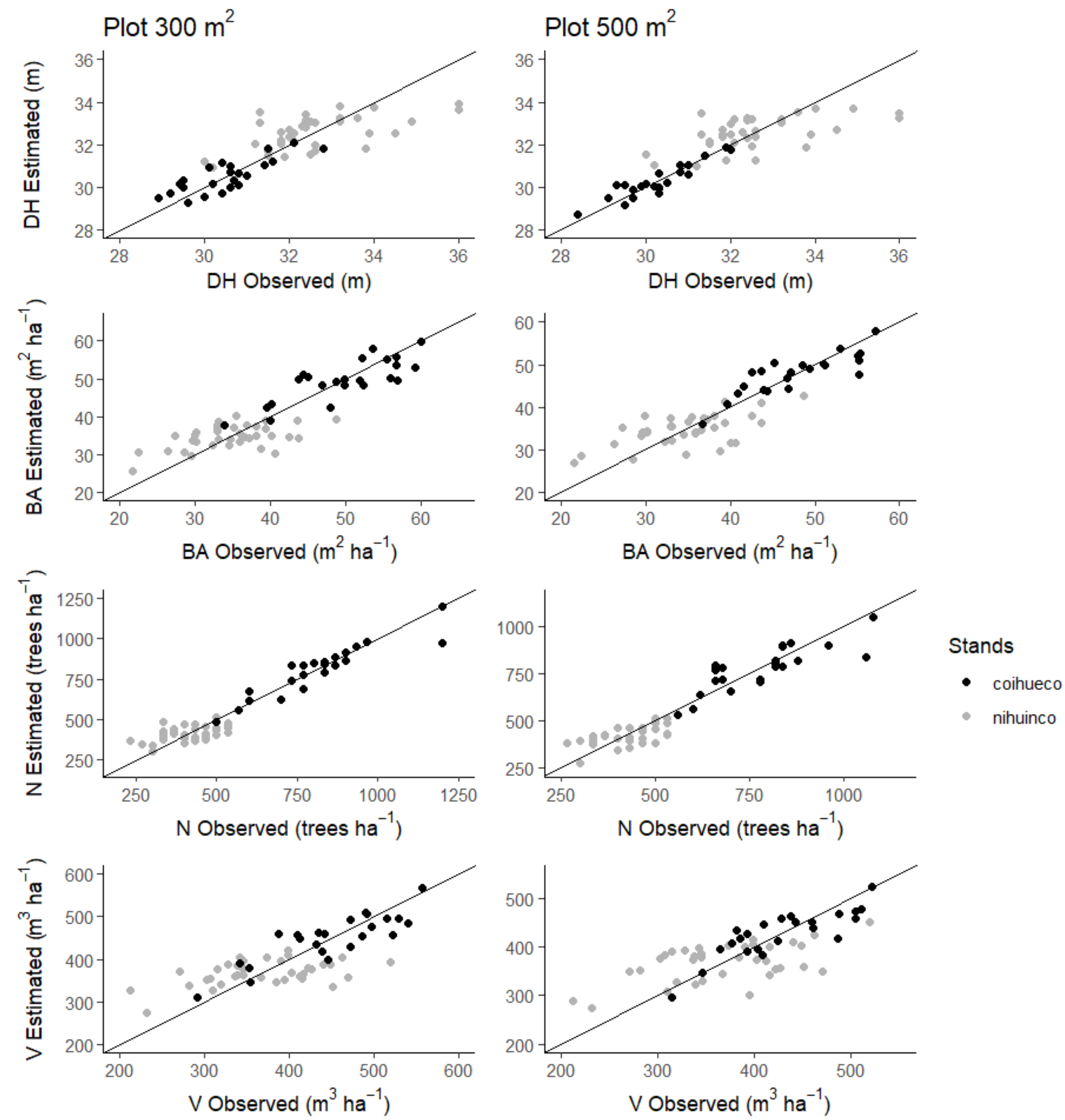

Figure 2. Plots with observed and estimated values of average fit model by plot size for the dasometric variables DH, BA, N and V.

Figure 3 presents plots showing the behavior of RMSE as the percentage of LiDAR cloud points in the DH, BA, V, and N modeling of the second modeling stage decreases. In general, the average RMSE values remain constant as the LiDAR cloud resolution decreases; this effect was observed in the modeling that used both plot sizes. Another evident effect, is that as the LiDAR cloud density decreases, greater variability was observed in the RMSE results of the lower resolution models, between 1% (4 points/ m²) and 5% (20 points/m²), after this point, the RMSE values tend to stabilize, with the exception of the volume estimation with a plot size of 500 m², since no break point is identified in comparison to the other estimations that a break point is identified at 5% according to the segmented models. Analyzing the trend of the RMSE values in the models that used the information from the 500 m² plots, it is possible to observe that here the RMSE values are lower in relation to the modeling performed with 300 m² plots and that this difference remains constant when analyzing the decrease in the density of the LiDAR cloud. The exception was evident in the N estimate, which was lower with a plot size of 300 m². The variation in values is more pronounced when a 300 m² plot is used to estimate DH, BA, N and V in the mean fit modeling and that effect is captured by our segmented models.

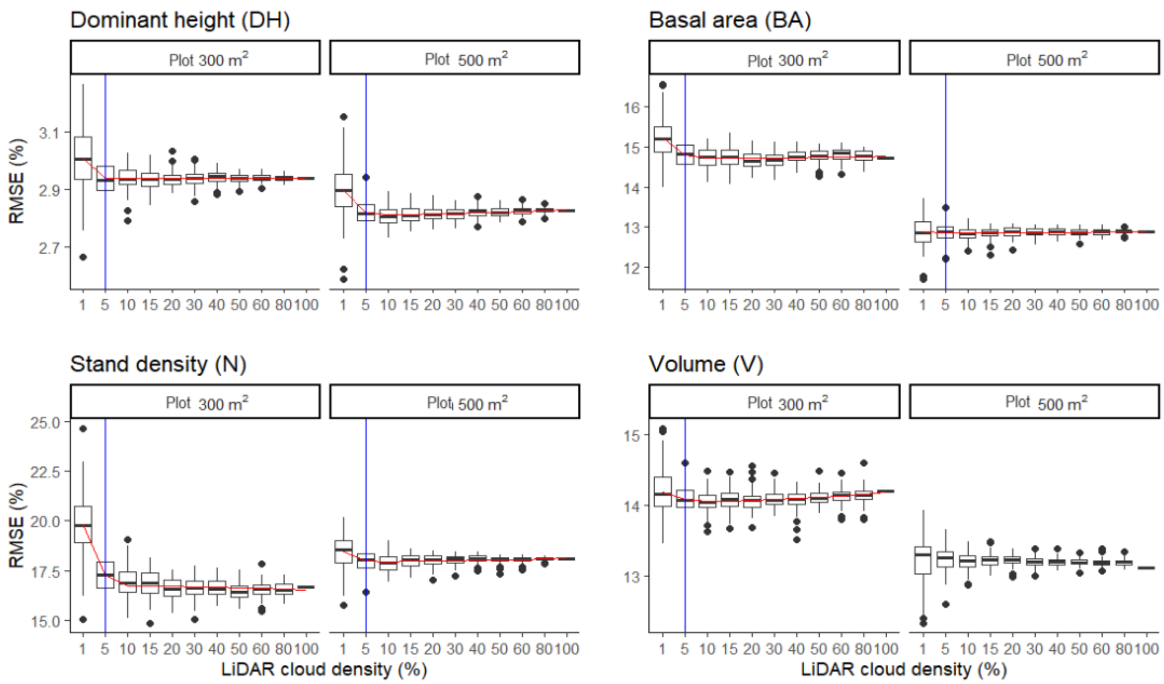

Figure 3. Plots of RMSE values of 100 iterations by LiDAR point density according to plot size of DH, BA, N and V dasometric variables of average fit models.

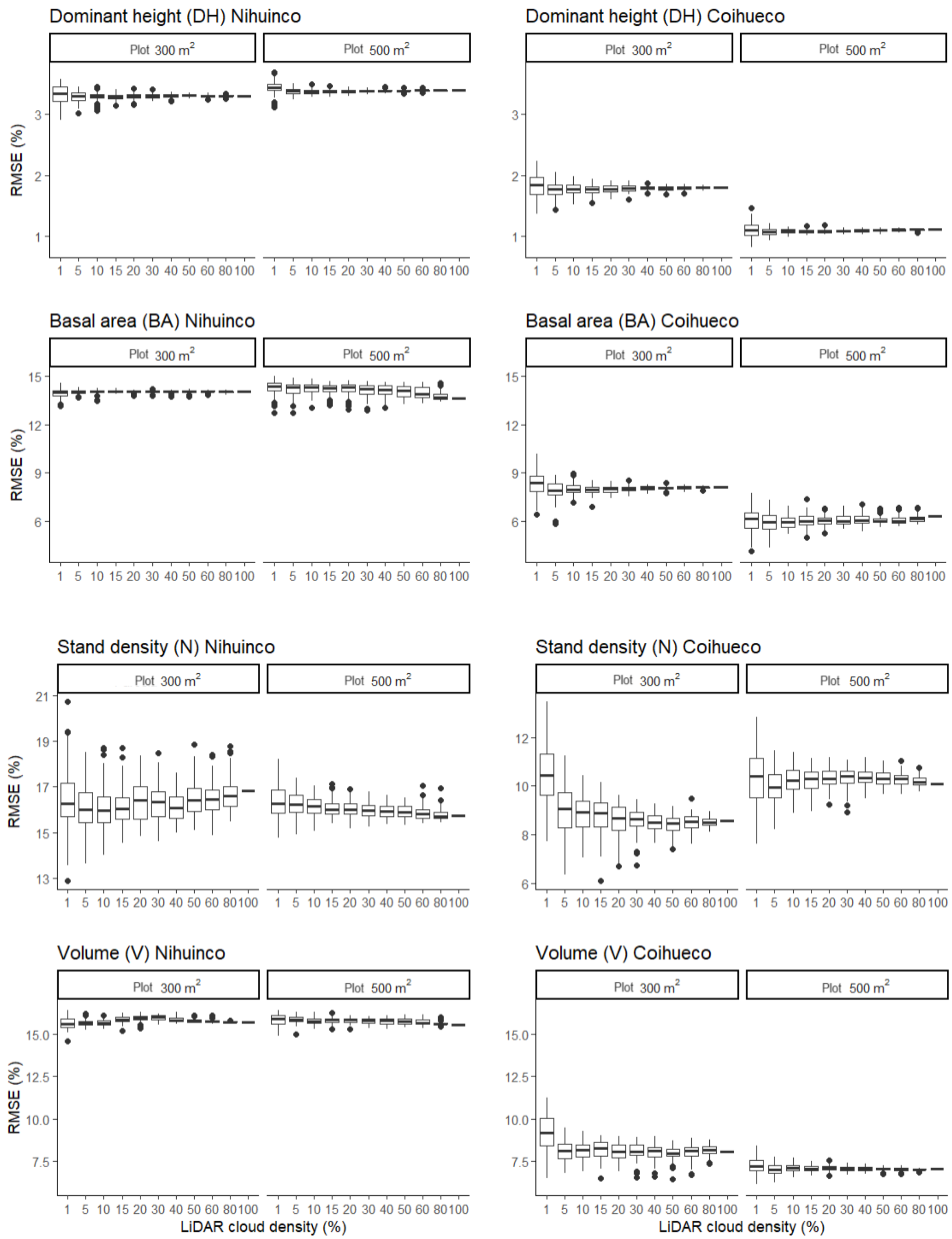

Figure 4. Plots of RMSE values of 100 iterations by LiDAR point density according to plot size for DH, BA, N and V of the model by plot.

The results of the SUR regression modeling are presented in Table 3. In terms of root mean square error (RMSE %) and coefficient of determination (R^2), the values obtained are very similar to those obtained using independent models. However, a significant improvement is observed in the estimation of V in the Coihueco stand when using a plot size of 300 m², where the RMSE % is reduced from 9.44 to 6.83 and the R^2 increases from 0.59 to 0.80. On the other hand, in the case of the stand density variable in the average adjustment condition of the study, the independent models show better results in terms of RMSE %, in this case, values of 16.88 and 16.09 are obtained, and when using plots of 300 m² and 500 m² respectively, the values of RMSE % increase to 19.84 and 17.30 with the SUR method.

Figure 5 shows the relationship between the measured dasometric variables and their estimates when using 100% of the LiDAR information, with independent and SUR models for the average adjustment. The predicted values of the four modeled variables do not show a major difference between the independent and SUR models.

Table 3. Results of RMSE (%) and R^2 indicators of the independent models and SUR method according to stand and plot size for DH, BA, V and N variables.

Variable	Stand	Plot (m²)	Independent Models		SUR Models	
			RMSE (%)	R^2	RMSE (%)	R^2
DH	Average	300	2,94	0,64	2,96	0,64
		500	1,83	0,82	1,87	0,82
	Nihuinco	300	3,39	0,31	3,44	0,28
		500	2,06	0,33	2,09	0,31
	Coihueco	300	0,78	0,94	0,59	0,80
		500	0,72	0,94	0,86	0,91
BA	Average	300	14,72	0,61	14,99	0,61
		500	10,60	0,77	11,06	0,75
	Nihuinco	300	13,69	0,33	13,79	0,33
		500	9,55	0,54	9,72	0,55
	Coihueco	300	6,49	0,76	6,55	0,78
		500	6,64	0,67	6,86	0,66
N	Average	300	16,88	0,82	19,84	0,76
		500	16,09	0,88	17,30	0,79
	Nihuinco	300	17,48	0,02	17,58	0,17
		500	13,52	0,38	13,81	0,36
	Coihueco	300	10,34	0,75	10,84	0,73
		500	10,07	0,63	9,86	0,69
V	Average	300	14,33	0,46	14,09	0,48
		500	9,53	0,68	9,66	0,68
	Nihuinco	300	15,15	0,30	15,08	0,32
		500	10,21	0,56	9,99	0,59
	Coihueco	300	9,44	0,59	6,83	0,80
		500	6,34	0,75	6,81	0,73

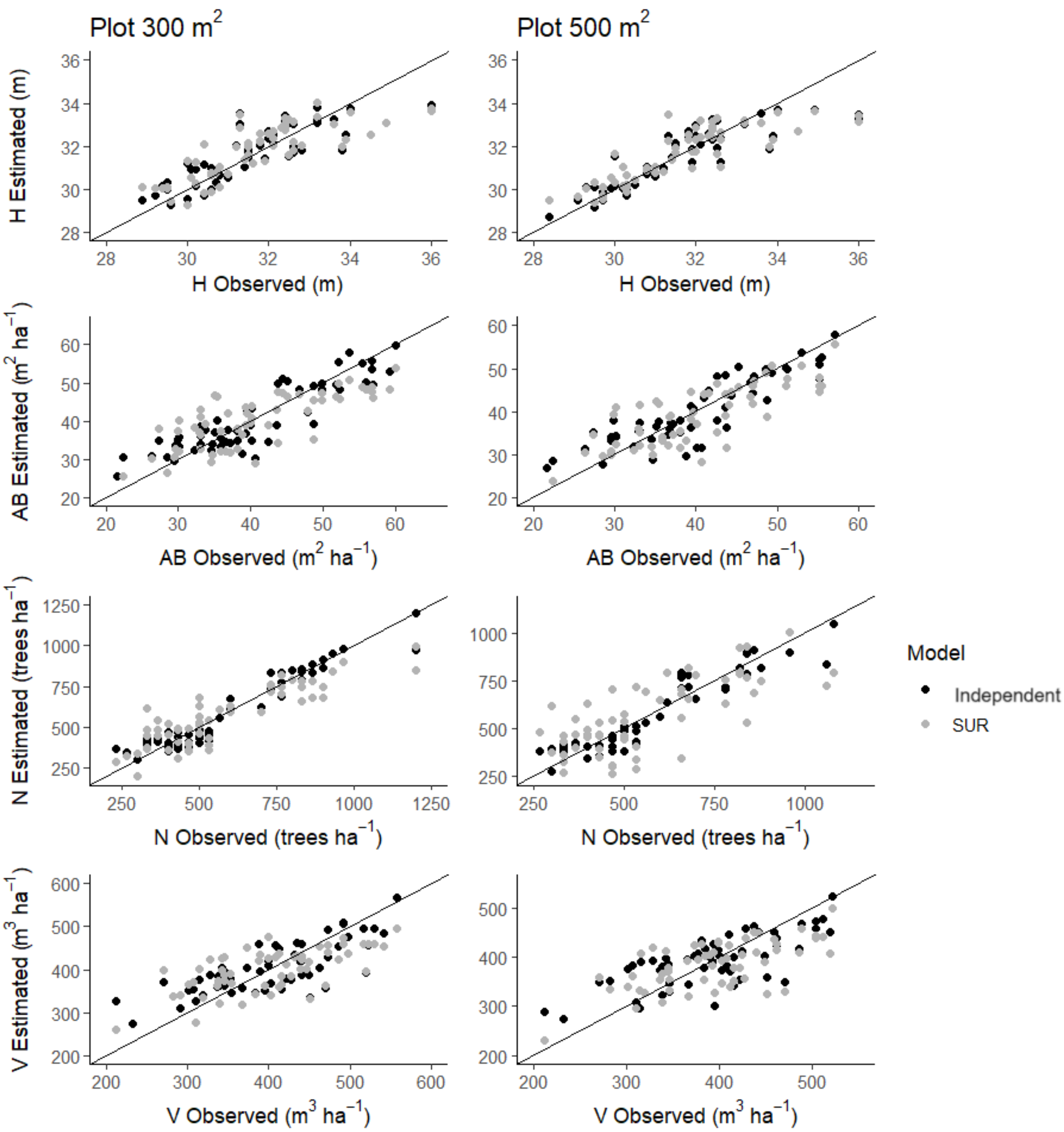

Figure 5. Plots with observed and estimated values of the independent model and SUR with average adjustment according to plot size for the dasometric variables HD, AB, N and V.

DISCUSSION

The values obtained in terms of RMSE (%) in the average fit modeling were better in all cases of estimation of dasometric variables when a plot size of 500 m² was used, highlighting the improvement in the estimation of basal area and volume, decreasing the RMSE from 14.72 to 10.60% and 14.33 to 9.53% respectively. Ruiz et al. (2014) indicate that plots of at least 500-600 m² are necessary to estimate volume, biomass and basal area, and using larger plots does not significantly increase modeling accuracy, but it does increase the cost and field work. Similarly, Li et al. (2022) found improvements when estimating dasometric variables with LiDAR information by systematically increasing the plot size from 100 to 900 m², indicating that at least a plot size of 600 m² should be used in the field plots, and that the estimates that decreased their RMSE the most were the basal area and volume variables when increasing the plot size.

The estimation of the dasometric variable with better RMSE and R2 indicators was HD in the average adjustment modeling that incorporated both plots using the plot size of 500 m² for dominant height presented an RMSE(%) of 1.84 and an R2 of 0.82, which represents results higher than those obtained by Treitz et al. (2012) with an RMSE value of 3.9% and Silva et al. (2018) with an RMSE of 5.7%, both works conducted in conifer plantations. Regarding the estimation of basal area, an RMSE(%) of 10.60 and an R2 of 0.77 were obtained, improving the estimates of the work of Treitz et al. (2012) who reported an RMSE(%) of 14.12. Regarding volume estimation, an RMSE(%) of 9.53 and an R2 of 0.68 were obtained, more accurate results when compared to the work of Pearse et al. (2019) who reported RMSE(%) values of 25 and the work of Lara-Gómez et al. (2023) with RMSE(%) values of 28.38, both works carried out in plantations of *Pinus radiata*. Finally, in the estimation of stand density, an RMSE(%) of 16.09 and an R2 of 0.88 were obtained, highlighting such estimation in comparison with the work of Rahlf et al. (2015) who obtained an RMSE(%) value of 51 or the work of Silva et al. (2018) in terms of R2 of 0.38, and even better than the work of López and Sandoval (2023) who reported improvements by incorporating a mixed method between ITD and ABA in the estimation of stand density with an RMSE of 20.9% in a *Pinus radiata* plantation.

The Coihueco stand (~770 trees/ha) showed the best RMSE values in the adjusted models for all the dasometric variables compared to the Nihuinco stand (~400 trees/ha). This behavior can be attributed to the point density of the LiDAR cloud used of ~400 points/ m², which better penetrates the upper canopy, providing better vertical vegetation information (Wallace et al. 2016), and the Nihuinco property had more understory, which may have contributed to the higher RMSE despite having a lower stand density than the Coihueco property. Lara-Gómez et al. (2023) used the ABA and ITD methods to estimate values of height, basal area, volume and radal density in a *Pinus radiata* plantation in Chile, indicating that when using the ABA method, better estimation results are obtained in stands with higher density, while the ITD method obtained better results in stands with lower density.

In this research, LiDAR metrics were selected for the estimation models of dasometric variables using regularization (LASSO). The use and implementation of this selection method in this area of research is still incipient. This methodology can select the most appropriate variables to develop more robust models. Kankare et al. (2013) used this methodology to select metrics and estimate biomass and volume in a boreal forest in Finland using ABA obtaining RMSE values of 24.9 and 26.4%, respectively. Recently, Adhikari et al. (2023) obtained R^2 values of 0.88, 0.83 and 0.87 for volume, basal area and dominant height, respectively, when using the selection of metrics with LASSO type regularization in plantations of *Eucalyptus globulus* in Chile, recommending the use of this type of metric selection in the modeling of forest dasometric variables of interest, due to its easy implementation and good results.

n this study we worked with information from the LiDAR point cloud of density ~400 points/m² which could be considered high resolution. The results presented in Figure 3, slight variations in RMSE (%) are observed, similar result found by Pearse et al. (2019), evaluating the effect of density from 1 to 280 points/m² in the estimation of dasometric variables in a *Pinus radiata* plantation in Australia. A point of 5 % of cloud density (20 points/m²) is identified, where the variation is higher if the percentage of cloud used is lower, in the estimation of dominant height, basal area, volume and stand density variables.

The results obtained with seemingly unrelated regression (SUR) have no significant differences compared to the results obtained with the independent models. Næsset et al. (2005) found minor differences when applying ordinary least squares (OLS), partial least squares (PLS) and SUR when estimating dasometric variables in a coniferous forest in Norway. Hao et al. (2022) found improvements when applying simultaneous SUR modeling to LiDAR information in estimating diameter distribution in a *Larix olgensis* plantation in China, pointing out the importance of considering spatial and inter-model correlations. SUR modeling is an underexplored procedure in estimating forest variables with LiDAR information, but the possibility of using this approach to use inter-model correlations to improve the efficiency of the estimates is highlighted (Woods et al. 2011).

CONCLUSION

In this study, it has been shown that the plot size used in ground measurements for model calibration using LiDAR information is a significant factor in obtaining estimates of dasometric variables such as dominant height, basal area, stand volume and stand density. It is recommended to use 500 m^2 plots instead of 300 m^2 plots, since more accurate estimates were obtained. In addition, stand density was found to be an important factor to consider when using the ABA approach, since more accurate estimates were obtained in stands with a greater number of trees. On the other hand, the selection of metrics with the LASSO-type regularization approach proved to be a powerful tool, as good results were obtained in the study. The density of the point cloud did not prove to be a significant factor in the estimation of the dasometric variables, but it was observed that a value of approximately 20 points/m^2 showed less variation in the estimates. Finally, it is concluded that simultaneous modeling with apparently unrelated regression does not show significant improvements compared to independent models.

REFERENCES

Adhikari A, CR Montes, A Peduzzi. 2023. A Comparison of Modeling Methods for Predicting Forest Attributes Using Lidar Metrics. *Remote Sensing* 15:1284.

Belmonte A, T Sankey, JA Biederman, J Bradford, SJ Goetz, T Kolb, T Woolley, N Horning, S Bohlman. 2019. UAV-derived estimates of forest structure to inform ponderosa pine forest restoration. *Remote Sensing in Ecology and Conservation* 6:181-197.

Carter J, K Schmid, K Waters, L Betzhold, B Hadley, R Mataosky, J Halleran. 2012. Lidar 101: An introduction to Lidar technology, data, and applications. Charleston, SC. National oceanic and atmospheric administration (NOAA), Coastal Services Center. 76 p.

Corte APD, DV Souza, FE Rex, CR Sanquetta, M Mohan, CA Silva, AMA Zambrano, G Prata, DR Alves de Almeida, JW Trautenmüller, C Klauberg, A de Moraes, MN Sanquetta, B Wilkinson, EN Broadbent. 2020. Forest inventory with high-density UAV-Lidar: Machine learning approaches for predicting individual tree attributes. *Computers and Electronics in Agriculture* 179.

Frazer GW, S Magnussen, MA Wulder, KO Niemann. 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment* 115:636-649.

Gobakken T, E Næsset. 2008. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research* 38:1095-1109.

González-Ferreiro E, U Diéguez-Aranda, D Miranda. 2012. Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry* 85:281-292.

Goodbody TR, NC Coops, PL Marshall, P Tompalski, P Crawford. 2017. Unmanned aerial systems for precision forest inventory purposes: A review and case study. *The Forestry Chronicle* 93:71-81.

Hansen E, T Gobakken, E Næsset. 2015. Effects of Pulse Density on Digital Terrain Models and Canopy Metrics Using Airborne Laser Scanning in a Tropical Rainforest. *Remote Sensing* 7:8453-8468.

Hao Y, FRA Widagdo, X Liu, Y Quan, Z Liu, L Dong, F Li. 2022. Estimation and calibration of stem diameter distribution using UAV laser scanning data: A case study for larch (Larix olgensis) forests in Northeast China. *Remote Sensing of Environment* 268:112769.

Jakubowski MK, Q Guo, M Kelly. 2013. Tradeoffs between LiDAR pulse density and forest measurement accuracy. *Remote Sensing of Environment* 130:245-253.

James G, D Witten, T Hastie, R Tibshirani. 2013. An introduction to statistical learning. New York. Springer.

Johnson N, SN Gordon, S Duncan, D Lach, B McComb, K Reynolds. 2007. Conserving creatures of the forest: A guide to decision making and decision models for forest biodiversity. Corvallis, OR. Oregon State University Press. 88 p.

Kankare V, M Vastaranta, M Holopainen, M Räty, X Yu, J Hyyppä, H Hyyppä, P Alho, R Viitala. 2013. Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR. *Remote Sensing* 5:2257-2274.

Lara-Gómez MÁ, RM Navarro-Cerrillo, I Clavero Rumbao, G Palacios-Rodríguez. 2023. Comparison of Errors Produced by ABA and ITC Methods for the Estimation of Forest Inventory Attributes at Stand and Tree Level in Pinus radiata Plantations in Chile. *Remote Sensing* 15:1544.

Li C, X Lin, H Dai, Z Li, M Zhou. 2022. Effects of Plot Size on Airborne LiDAR-Derived Metrics and Predicted Model Performances of Subtropical Planted Forest Attributes. *Forests* 13:2124.

López M, S Sandoval. 2023. Estimación de la densidad de rodal a partir de información LiDAR aérea integrando el método de masa y árbol individual en plantaciones de *Pinus radiata. Bosque (Valdivia)* 44:377-386.

Magnussen S, E Næsset, T Gobakken. 2010. Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sensing of Environment* 114:700-712.

Manuri S, H-E Andersen, RJ McGaughey, C Brack. 2017. Assessing the influence of return density on estimation of lidar-based aboveground biomass in tropical peat swamp forests of Kalimantan, Indonesia. *International Journal of Applied Earth Observation and Geoinformation* 56:24-35.

Mascaro J, M Detto, GP Asner, HC Muller-Landau. 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment* 115:3770-3774.

Mauro F, R Valbuena, JA Manzanera, A García-Abril. 2011. Influence of Global Navigation Satellite System errors in positioning inventory plots for tree-height distribution studiesThis article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Canadian Journal of Forest Research* 41:11-23.

Næsset E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 80:88-99.

Næsset E, OM Bollandsås, T Gobakken. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sensing of Environment* 94:541-553.

Packalen P, JL Strunk, JA Pitkanen, H Temesgen, M Maltamo. 2015. Edge-tree correction for predicting forest inventory attributes using area-based approach with airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8:1274-1280.

Pearse GD, MS Watt, JP Dash, C Stone, G Caccamo. 2019. Comparison of models describing forest inventory attributes using standard and voxel-based lidar predictors across a range of pulse densities. *International Journal of Applied Earth Observation and Geoinformation* 78:341-351.

Rahlf J, J Breidenbach, S Solberg, R Astrup. 2015. Forest Parameter Prediction Using an Image-Based Point Cloud: A Comparison of Semi-ITC with ABA. *Forests* 6:4059-4071.

Roussel J-R, D Auty, F De Boissieu, A Meador. 2021. LidR : Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. Disponible en https://cran.r-project.org/web/packages/lidR/index.html.

Roussel J-R, J Caspersen, M Béland, S Thomas, A Achim. 2017. Removing bias from LiDAR-based estimates of canopy height: Accounting for the effects of pulse density and footprint size. *Remote Sensing of Environment* 198:1-16.

Ruiz L, T Hermosilla, F Mauro, M Godino. 2014. Analysis of the Influence of Plot Size and LiDAR Density on Forest Structure Attribute Estimates. *Forests* 5:936-951.
Sánchez J, P Oliver, J Estornell, C Dopazo. 2018. 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* 21:79-99.
Silva CA, C Klauberg, AT Hudak, LA Vierling, V Liesenberg, LG Bernett, CF Scheraiber, ER Schoeninger. 2018. Estimating stand height and tree density in pinus taeda plantations using in-situ data, airborne LiDAR and k-nearest neighbor imputation. *An Acad Bras Cienc* 90:295-309.
Strunk J, H Temesgen, H-E Andersen, JP Flewelling, L Madsen. 2014. Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing* 38:644-654.
Tompalski P, JC White, NC Coops, MA Wulder. 2019. Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote Sensing of Environment* 227:110-124.
Treitz P, K Lim, M Woods, D Pitt, D Nesbitt, D Etheridge. 2012. Lidar sampling density for forest resource inventories in Ontario, Canada. *Remote Sensing* 4:830-848.
Varo-Martínez MÁ, RM Navarro-Cerrillo, R Hernández-Clemente, J Duque-Lazo. 2017. 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* 56:54-64.
Wallace L, A Lucieer, Z Malenovský, D Turner, P Vopěnka. 2016. Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests* 7:62.
White JC, MA Wulder, A Varhola, M Vastaranta, NC Coops, BD Cook, D Pitt, M Woods. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *The Forestry Chronicle* 89:722-723.
Woods M, D Pitt, M Penner, K Lim, D Nesbitt, D Etheridge, P Treitz. 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle* 87:512-528.
Yoga S, J Bégin, B St-Onge, M Riopel. 2017. Modeling the effect of the spatial pattern of airborne lidar returns on the prediction and the uncertainty of timber merchantable volume. *Remote Sensing* 9.
Zhang J, J Hu, J Lian, Z Fan, X Ouyang, W Ye. 2016. Seeing the forest from drones: Testing the potential of lightweight drones as a tool for long-term forest monitoring. *Biological Conservation* 198:60-69.