Forestry An International Journal of Forest Research

Institute of Chartered Foresters

*Forestry* 2017; **00**, 1–14, doi:10.1093/forestry/cpx021

# Generalized allometric equations of total volume generated by meta-analysis for roble, raulí and coigüe in second growth forests in Chile

**Carlos Valenzuela[1]*, Eduardo Acuña[1], Fernando Muñoz[1], Alicia Ortega[2] and Jorge Cancino[1]**

[1]*Facultad de Ciencias Forestales, Universidad de Concepción, Concepción 4070386, Chile*
[2]*Facultad de Ciencias Forestales y Recursos Naturales, Universidad Austral de Chile, Valdivia 5110566, Chile*

*Corresponding author. Tel: +56 994741027; E-mail: carlovalenzuela@udec.cl

*Received 13 September 2016*

A modified meta-analysis was used to develop generalized allometric equations of total individual volume underbark for second growth forests of roble, raulí and coigüe. From a set of total volume equations compiled from the literature, pseudo-data were generated to fit generalized equations for each species. The meta-analysis was supplemented with an observed database, which, besides contributing data for validation was used to fit the height–diameter relationship and to model the variance of total volume. This variance was used to simulate random pseudo-data of volume with variability similar to that of the observed data, in order to avoid residual autocorrelation problems. It was also used to define weights for the fitting of volume equations. In order to determine zonal effects, volume equations compiled from the literature were assigned to agro-climatic zones defined by ODEPA (ODEPA. 2000 Oficinas De Estudios y Políticas Agrarias. Clasificación de las explotaciones agrícolas del VI censo nacional agropecuario según tipo de productor y localización geográfica. Ministerio de Agricultura. Documento de trabajo N°5. I.S.S.N. 0717-0378. Santiago, Chile. 91 p.), according to the location of the sample. In the fitting, which included *dummy* variables, no significant zonal effects were detected in the regression parameters in any of the species. The generalized allometric equations of total volume showed highly precision and accuracy, i.e. $FI < 0.0852\,m^3\,tree^{-1}$ and $E < 0.0674\,m^3\,tree^{-1}$. Thus, obtained equations are considered valid for widespread use in the study zones.

## Introduction

In Chile, there are 3.8 million hectares of second growth forests, of which 1.2 million ha correspond to Roble-Raulí-Coigüe Forest Type, distributed from the Maule Region (35°25′S–71°40′W) to Los Lagos Region (43°28′S–72°56′W) (CONAF, 2011). Second growth forests of roble (*Nothofagus obliqua* (Mirb.) Oerst.), raulí (*N. alpina* (Poepp et Endl.) Oerst.) and coigüe (*N. dombeyi* (Mirb.) Oerst.) represent a high value economic resource, because they have high growth rates and high timber quality (Donoso *et al.*, 1993; Lara *et al.*, 1999). Sustainable use of this resource requires evaluations at individual tree level based on allometric equations of total volume. Although there are currently local equations of total volume in specialized publications and other bibliographic sources for these three species, the use of these models outside areas for which they were designed is risky, because it may result in erroneous estimations (Fournier *et al.*, 2003; Henry *et al.*, 2011). Because the local equations come from limited geographic zones, their widespread use in different environmental conditions where stands of varied structures, densities and ages are intermingled, requires at least a validation based on independent samples (Wirth *et al.*, 2004). Usually, local equations are not suited to perform regional or large scale estimates, resulting in the need to develop generalized allometric equations of total volume (Muukkonen, 2007).

In recent years, generalized allometric equations of total volume have been developed for different species that grow in a wide range of environmental conditions (Case and Hall, 2008; Henry *et al.*, 2011). These equations can be developed from two methods. The first method consists in using observed data, compiled through destructive sampling over a large area (Zianis and Mencuccini, 2003; Lambert *et al.*, 2005; Wutzler *et al.*, 2008; Návar, 2009; Henry *et al.*, 2011); however, because of the high cost of this method the application of this method has been limited. The second method corresponds to the meta-analysis, which develops generalized equations from previously published equations (Pastor *et al.*, 1983/1984; Jenkins *et al.*, 2003; Muukkonen, 2007; Chojnacky *et al.*, 2014; Wayson *et al.*, 2015); this method has a lower cost but requires an independent set of data for validation.

There are two techniques of meta-analysis to develop generalized equations for volume, formal and modified (Jenkins *et al.*, 2003). The formal meta-analytical technique combines the regression coefficients from different equations and all equations used in such meta-analysis must have identical structure and identical variable transformations (Peña, 1997). While the modified meta-analysis fits generalized equations from pseudo-data, generated with previously published equations, it does not present constraints in structure and variables, unlike the formal technique, making it more applicable in practice (Pastor et al., 1983/1984; Jenkins et al., 2003). Since original databases used in the fitting of the published equations are not usually available, the modified meta-analysis can expand the availability of data from different locations, because in the absence of observed data, the information provided by specific data of each location can be recovered under the form of specific pseudo-data, which are generated using the published equations. This is a more practical method that requires no field sampling for fitting one or more models.

Several authors have used this meta-analytical technique to develop generalized equations for different species in North America and Europe (Pastor et al., 1983/1984; Jenkins et al., 2003; Wirth et al., 2004; Zianis et al., 2005; Muukkonen, 2007). Some authors only use the diameter at breast height as the predictor variable in the generation of pseudo-data. Others, consider that both diameter at breast height and total tree height should be included, because these data provide better estimates of volume (Zianis and Mencuccini, 2003; Montagu et al., 2005; António et al., 2007; Gonzalez-Benecke et al., 2014). However, before to the generation of pseudo-data, the problem of allocating the value of total height to each specific diameter should be faced, and can be solved by fitting the height–diameter relationship with observed data (Muukkonen, 2007).

Another problem to be solved is that the direct generation of pseudo-data from the compiled equations violates the assumption of independence of residuals. The pseudo-data set on which the generalized equations are based is highly autocorrelated (Lambert et al., 2005). Additionally, in total height and total volume data, it is usual to observe heteroscedasticity; there is a greater variability in these variables at larger size of the trees (Wayson et al., 2015). Thus, the estimates of the regression coefficients of the generalized allometric equations obtained using ordinary least squares are linear, unbiased and consistent, but the estimators not efficient. The estimated variances are biased, so the statistical tests normally used and the confidence intervals generated to verify the validity of the estimates are not valid. Moreover, the reported coefficient of determination ($R^2$) values are meaningless (Kmenta, 1986).

Several studies published on generalized equations based on pseudo-data have not considered the autocorrelation problem and heteroscedasticity (Jenkins et al., 2003; Muukkonen, 2007; Chojnacky et al., 2014). The main problem is often a lack of statistical information, which generally accompanies the published equations, which in some cases is restricted to the regression parameters and sample size. Thus, when carrying out a meta-analysis for the preparation of generalized equations, a method to generate pseudo-data free from autocorrelation should be applied as well, with a distribution similar to that of the observed data.

This study was aimed to develop generalized allometric equations of total volume underbark for second growth forests of roble, raulí and coigüe using the modified meta-analysis, supplemented with observed data. The observed database, besides being used in the validation of the generalized fitted equations from pseudo-data, was used to model the height–diameter relationship ($H = f(D)$), to model the variance of the total volume underbark ($\sigma_V^2 = f(D^2H)$), and to model the total volume underbark from observed data. The height–diameter equation was used to generate volume pseudo-data from the combined variable $D^2H$, with the allometric equations of the total volume underbark compiled from different bibliographic sources. The variance equation, besides being used to randomize volume pseudo-data and provide a variability similar to that of the observed data, with the purpose of avoiding residual autocorrelation problems, it was used to define weights for the fit of volume by weighted least squares method. The allometric equations of total volume underbark fitted to observed data were used to generate pseudo-data, from which new equations were fitted, in order to have direct evidence on the validity of the method used in this study.

## Methods

### *Compilation equations*

The allometric equations of the total volume underbark used in this study were compiled through a literature review. First, an online review was conducted, consulting databases and meta-searchers available in the main Chilean university libraries, including books, scientific and technical magazines, undergraduate and graduate theses. Then, authors of the documents that were not freely available were contacted by e-mail and digital and printed copies of their studies were obtained. The literature review allowed us to identify the authors of the studies and contact them to request their permission to access and use their data for this study.

The compilation was focused on allometric equations of total volume underbark obtained from models, whose predicting variables included diameter at breast height, total height or combination thereof (i.e. $V = f(D, H, D^2H)$). All equations that clearly informed about the structure of the model, value of estimated parameters, sample size, range of values of independent variables (i.e. $D$ and $H$), geographic location of the sample used in the fitting (i.e. Region, Province, Commune, locality or farm, and geographic coordinates), and source were considered useful to the study.

The analysis of the compiled allometric equations showed the frequent use of a reduced number of models (Table 1). For roble, three models were found (a, b and c), while for raulí and coigüe, only one model was found (a).

The detected models have been used to fit a variable number of equations. Of all allometric equations of total volume underbark compiled, a set of 18 equations met the requirements outlined in previous paragraphs and they were used in this study (Table 2). In order to determine zonal effects, the equations were allocated to five agro-climatic zones defined by ODEPA (2000), according to the geographic location of the sample used in the fitting of each equations. The zones were: Coastal rainfed, Inner rainfed, Rainfed valley, Foothills and Mountains. The nine equations for roble considered useful for the study cover four

**Table 1** Models of total volume underbark detected in the literature review, by species.

| Species | Model structure | Model |
|---|---|---|
| Roble, raulí and coigüe | $V = a_0 + a_1D^2H$ | (a) |
| Roble | $V = a_0 + a_1D^2H + a_2D^2$ | (b) |
| Roble | $V = a_0 + a_1D^2H + a_2D^2 + a_3H$ | (c) |

$V$: total volume underbark ($m^3$); $D$: diameter at breast height (cm); $H$: total height (m); $a_{0,1,2,3}$: parameters of the models.

**Table 2** Parameters estimates and fit statistics for total volume underbark equations used in this study, by species and zone.

| Species | Zone | Model | Ranges | | $n$ | Parameters | | | | $R^2$ | *RMSE* (m$^3$) | *CV* (%) | Source |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *D* (cm) | *H* (m) | | $a_0$ | $a_1$ | $a_2$ | $a_3$ | | | | |
| Roble | Inner rainfed | (a) | 10.0–55.0 | 5.0–35.0 | 134 | 5.0000E−03 | 3.1510E−05 | | | 0.9742 | 0.0344 | n.i. | Donoso *et al.* (1984)* |
| | | (b) | 10.0–55.0 | 5.0–35.0 | 134 | 9.4000E−03 | 3.4130E−05 | −6.3500E−05 | | 0.9722 | 0.0346 | n.i. | Donoso *et al.* (1984)* |
| | | (c) | 10.0–55.0 | 5.0–35.0 | 134 | 1.7890E−02 | 3.6110E−05 | −1.2990E−04 | −0.0012 | 0.9742 | 0.0346 | n.i. | Donoso *et al.* (1984)* |
| | Rainfed valley | (a) | 10.0–60.0 | 5.0–40.0 | 30 | 2.4623E−02 | 3.0644E−05 | | | n.i. | n.i. | n.i. | Nuñez and Real (1992)* |
| | | (a) | 10.0–60.0 | 5.0–40.0 | 30 | 2.5957E−02 | 3.0292E−05 | | | n.i. | n.i. | n.i. | Nuñez and Real (1992)* |
| | Foothills | (a) | 10.0–60.0 | 5.0–35.0 | 24 | 3.2300E−03 | 3.2140E−05 | | | 0.9870 | 0.0430 | n.i. | Puente *et al.* (1981) |
| | | (a) | 10.0–60.0 | 5.0–35.0 | 24 | 3.5238E−03 | 3.8263E−05 | | | 0.9742 | n.i. | n.i. | UACH (1982)* |
| | Mountains | (a) | 15.0–60.0 | 10.0–35.0 | 15 | 7.1120E−02 | 3.3890E−05 | | | 0.9890 | 0.0386 | n.i. | Puente *et al.* (1981) |
| | | (a) | 10.6–38.1 | 7.2–22.6 | 50 | 2.5828E−02 | 2.8502E−05 | | | 0.9900 | n.i. | 9.40 | Grosse and Cubillos (1991) |
| Raulí | Foothills | (a) | 10.0–60.0 | 5.0–35.0 | 26 | 1.3710E−02 | 2.8899E−05 | | | 0.9930 | 0.0233 | n.i. | Puente *et al.* (1981) |
| | | (a) | 16.1–32.2 | 18.1–25.3 | 13 | 1.4110E−02 | 2.6890E−05 | | | 0.9900 | n.i. | 8.9 | Cubillos (1988a) |
| | | (a) | 14.5–35.9 | 19.2–26.4 | 21 | 8.5000E−04 | 2.8390E−05 | | | 0.9900 | n.i. | 14.3 | Cubillos (1988a) |
| | | (a) | 9.1–40.3 | 10.6–27.7 | 80 | 4.6000E−02 | 2.9056E−05 | | | 0.9850 | 0.0842 | n.i. | Grosse and Cubillos (1991) |
| | Mountains | (a) | 5.2–33.3 | 6.6–24.2 | 75 | 2.0700E−03 | 3.0000E−05 | | | 0.9801 | n.i. | 13.3 | Grosse and Cubillos (1991) |
| | | (a) | 10.0–60.0 | 5.0–30.0 | 28 | 1.3370E−02 | 2.9931E−05 | | | 0.9940 | 0.0204 | n.i. | Puente *et al.* (1981) |
| | | (a) | 5.2–33.3 | 6.6–24.2 | 75 | 2.0700E−03 | 3.0000E−05 | | | 0.9900 | n.i. | 13.13 | Cubillos (1988a) |
| Coigüe | Rainfed valley | (a) | 10.0–60.0 | 5.0–40.0 | 30 | 2.4623E−02 | 3.0644E−05 | | | n.i. | n.i. | n.i. | Nuñez and Real (1992)* |
| | Mountains | (a) | 12.1–39.2 | 13.7–26.2 | 50 | 1.2105E−02 | 2.9462E−05 | | | 0.9900 | n.i. | 6.4 | Cubillos (1988b) |

*D*: diameter at breast height; *H*: total height; $n$: sample size; $R^2$: coefficient of determination; *RMSE*: root mean square residual; *CV*: residual coefficient of variation; n.i.: no information; *: as cited by Drake *et al.* (2003).

of the five zones with presence of second growth forests of the species (i.e. 3 for Inner rainfed, 2 for Rainfed valley, 2 for Foothills and 2 for Mountains). The seven equations for raulí only cover two zones (i.e. 4 for Foothills and 3 for Mountains), as do the two equations for coigüe (i.e. 1 for Rainfed valley and 1 for Foothills).

## Database

The observed database was structured from a sampling performed in the framework of this study and data provided by other studies. Similar to what was carried out with compiled equations, the observed data were allocated to the agro-climatic zones defined by ODEPA (2000), according to the geographic location of each sample (Table 3).

The database was compiled in the range of the natural distribution of second growth forests of roble, raulí and coigüe (Figure 1). The database includes measurements of total height ($H$) and diameter at breast height ($D$) of a total of 2044 trees, including 753 robles, 916 raulies and 375 coigües. Of them, 1380 (i.e. 635 robles, 459 raulies and 286 coigües) correspond to trees that were felled in order to measure diameters over and underbark along the stem. From these data, the total volume underbark was determined. Volume underbark ($V$) on each tree section was calculated using Smalian's formula, i.e. $V_i = (A_l + A_u)L/2$, where $V_\mathrm{i}$ is the volume underbark (m$^3$) for section $i$; $L$ is the section length (m) and $A_{l,u}$ are the upper and lower-end cross-sectional areas (m$^2$) for section $i$, respectively. Total stem volume was obtained by adding the volume of the stem sections from a stump height equal to 0.3 m up to tree top. The database observed for each species and zone was divided for fitting and validation purposes with independent samples. The *HD* sample was systematically divided into two parts. The first consisted of 634 trees and was intended for the fitting height–diameter relationship (*H*/*D*), whereas the second contained 1431 trees and was used for validation. The database of total volume (*VDH*) was systematically divided into three datasets. The first was of 415 trees intended for the fitting of variance equations of the total volume. The second, consisted of 676 trees for the fitting of allometric equations of total volume from observed data. The third contained 289 trees and was intended for the validation of these relationships. The allometric equations of total volume fitted to observed data were used to generate pseudo-data, which in turn were used to fit equations in order to have direct evidence on the validity of the method used in this study.

## Model development

For the modelling of the total volume underbark, the non-linear model proposed by Spurr (1952) (1) was used. The fit was performed from pseudo-data generated by using the compiled allometric equations of total volume underbark. The model proposed by Spurr (1952) also was fitted to observed data and then was fitted with pseudo-data generated from these equations, in order to have direct evidence on the validity of the method used in this study.

$$V = a_1(D^2H)^{a_2} \tag{1}$$

where $V$ is the total volume underbark (m$^3$), $D$ is the diameter at breast height (cm), $H$ is the total height (m) and $a_{1,2}$ are the parameters of the model.

With each total volume equation compiled, an amount of pseudo-data equal to the sample size used in the fitting ($n$) was generated in the diameter range ($D_{\mathrm{min}}$, $D_{\mathrm{max}}$) of the sample used in the fitting of each equation and at uniform intervals of range $(D_{\mathrm{min}} - D_{\mathrm{max}})/n$. This ensures a greater weight to the equations fitted with larger sampling size. In order to estimate the total height ($H_i$) corresponding to a specific value of diameter ($D_i$), the allometric model proposed by Stage (1963) (2) was fitted. The equations resulting from the fitting were used to estimate heights in the range of heights ($H_{\mathrm{min}}$, $H_{\mathrm{max}}$) of the sample used in the fitting of each volume equation compiled. In this case, three options were analysed; the first consisted of positioning the estimation line on the minimum height (relationship 3); in the second, the estimation line was positioned on the mean height (relationship 4); in the third option, the estimation line was positioned on the maximum height (relationship 5).

$$H = 1.3 + b_1D^{b_2} \tag{2}$$

$$H_i = H_{\mathrm{min}} + b_1(D_i^{b2} - D_{\mathrm{min}}^{b2}) \tag{3}$$

$$H_i = H_{\mathrm{mean}} + b_1(D_i^{b2} - D_{\mathrm{mean}}^{b2}) \tag{4}$$

$$H_i = H_{\mathrm{max}} + b_1(D_i^{b2} - D_{\mathrm{max}}^{b2}) \tag{5}$$

where $H_{\mathrm{min,mean,max}}$ are the minimum, medium and maximum total height (m), respectively; $D_{\mathrm{min,mean,max}}$ are the minimum, medium and maximum diameter at breast height (cm), respectively and $b_{1,2}$ are the parameters of the model.

The generation of non-autocorrelated pseudo-data required the modelling of the variance of the total volume underbark, as an equation of the predicting variable of the model (1). In order to achieve this, the observed data available for fitting (i.e. *VDH* database) were arranged in eight classes of $D^2H$ of constant amplitude and with class midpoints defined from the following relationship $mc = [0.5 + int(0.99D^2H/w)]w$, where $mc$ is the $D^2H$ class midpoint, $int$ is the equation of the integer part and $w$ is the class amplitude, which in this study was set in $w = 5000$. In each class, the arithmetic average of the predicting variable ($\overline{D^2H_i}$) and the volume variance ($\sigma^2_{v\_i}$) were calculated. In the modelling of the volume variance, the linear model (6) was used.

$$\sigma^2_{v\_i} = c_0 + c_1\overline{D^2H_i} \tag{6}$$

where $c_{0,1}$ are the parameters of the model.

The generation of pseudo-data of total volume underbark using each compiled equation was performed as follows: given a value of the diameter at breast height ($D_i$), the total height ($\hat{H}_i$) was estimated using the best relationship (3, 4 or 5). Volume variance ($\hat{\sigma}^2_{v\_i} = f(D_i^2\hat{H}_i)$) was estimated using model (7), and the total volume underbark was estimated from the respective compiled equation ($\hat{V}_i = g(D_i, \hat{H}_i)$). The respective pseudo-datum was randomly obtained from a normal distribution, whose mean and variance are $\hat{V}_i$ and $\hat{\sigma}^2_{v\_i}$, respectively. In turn, 10 000 pseudo-data were generated using this methodology and 10 000 pseudo-data were generated using a Monte-Carlo simulation, similar to the proposed by Wayson *et al.* (2015). Both simulations were used to validate the generalized allometric total volume equations.

## Fitting and evaluation of models

All statistical analyses were performed using SAS 9.2 software (SAS, 2009). The non-linear models (1 and 2) and linear model (7) were fitted using the NLIN and REG procedures, respectively. The fitting of the model of variance of the total volume (6) was performed by species. While the fitting of the models of total volume underbark (1) and total height (2) were performed by species and agro-climatic zone, using weighted least squares to homogenize the variance. For the model (1), the weighting factor used was the reciprocal of the variance equation (i.e. $1/\sqrt{f(\sigma^2_{v\_i})}$). While for the model (2), the weighting factor used was the reciprocal of the predicting variable (i.e. $1/\sqrt{D}$).

The determination of the significance of the zonal effects in the models of total volume underbark and total height was performed by incorporating *dummy* variables associated to each parameter of the models (1 and 2) (Hardy, 1993; Bergerud, 1994; Ott, 1997). In a hypothetical situation with data available for $z$ zones, a total of $z - 1$ *dummy* variables

**Table 3** Descriptive statistics of the database used to fit and validation of equations, by species and zone.

| Species | Database | Variable | Zone | $n_a$ | $n_v$ | *D* (cm) | | | *CV* (%) | *H* (m) | | | *CV* (%) | *V* ($m^3$) | | | *CV* (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Min. | Mean | Max. | | Min. | Mean | Max. | | Min. | Mean | Max. | |
| Roble | *HD* | *H* | Coastal rainfed | 11 | 27 | 9.8 | 22.1 | 64.5 | 51.4 | 10.7 | 19.8 | 26.3 | 22.4 | | | | |
| | | | Inner rainfed | 21 | 49 | 5.4 | 20.0 | 51.0 | 51.6 | 6.0 | 15.0 | 34.0 | 35.1 | | | | |
| | | | Rainfed valleys | 17 | 41 | 5.0 | 20.0 | 44.5 | 53.7 | 9.0 | 20.9 | 37.3 | 38.8 | | | | |
| | | | Foothills | 84 | 196 | 5.0 | 24.8 | 66.0 | 37.6 | 9.0 | 21.1 | 41.3 | 26.9 | | | | |
| | | | Mountains | 92 | 215 | 4.9 | 20.1 | 59.1 | 55.3 | 4.8 | 17.9 | 36.8 | 36.2 | | | | |
| | | | Total | 225 | 528 | 4.9 | 22.1 | 66.0 | 48.0 | 4.8 | 19.0 | 41.3 | 32.6 | | | | |
| | *VDH* | $\sigma_V^2$ | Coastal rainfed | 8 | 7 | 11.2 | 22.2 | 43.9 | 44.8 | 13.3 | 18.1 | 23.6 | 20.7 | 0.0109 | 0.3764 | 1.5299 | 122.4 |
| | | | Inner rainfed | 21 | 14 | 5.4 | 20.4 | 49.1 | 58.8 | 6.0 | 15.0 | 29.3 | 35.1 | 0.0075 | 0.3901 | 1.8694 | 136.7 |
| | | | Rainfed valleys | 17 | 12 | 5.0 | 20.0 | 44.5 | 53.7 | 9.0 | 19.4 | 32.1 | 38.4 | 0.0083 | 0.4036 | 1.5559 | 114.1 |
| | | | Foothills | 79 | 56 | 7.2 | 24.8 | 42.0 | 32.9 | 11.6 | 20.6 | 35.6 | 21.2 | 0.0224 | 0.4764 | 2.1762 | 64.4 |
| | | | Mountains | 65 | 45 | 4.9 | 21.7 | 47.9 | 40.9 | 6.3 | 18.7 | 32.4 | 30.6 | 0.0068 | 0.3968 | 1.7132 | 95.2 |
| | | | Total | 190 | 134 | 5.0 | 22.7 | 49.1 | 40.9 | 6.0 | 19.1 | 35.6 | 28.7 | 0.0068 | 0.4289 | 2.1762 | 88.7 |
| | | *V* | Coastal rainfed | 14 | 7 | 9.8 | 19.5 | 43.9 | 36.9 | 12.0 | 17.8 | 26.3 | 24.6 | 0.0109 | 0.2709 | 1.5299 | 84.8 |
| | | | Inner rainfed | 35 | 14 | 5.4 | 21.1 | 51.0 | 50.5 | 8.0 | 16.8 | 34.0 | 37.2 | 0.0228 | 0.4440 | 3.1375 | 144.1 |
| | | | Rainfed valleys | 29 | 12 | 6.0 | 23.1 | 43.5 | 43.1 | 9.2 | 21.2 | 35.5 | 35.2 | 0.0109 | 0.4898 | 1.4319 | 89.0 |
| | | | Foothills | 128 | 56 | 5.3 | 25.1 | 51.1 | 37.0 | 9.0 | 20.6 | 33.7 | 23.3 | 0.0104 | 0.5189 | 2.8803 | 81.4 |
| | | | Mountains | 105 | 45 | 4.9 | 21.8 | 54.7 | 49.8 | 4.9 | 18.8 | 32.5 | 33.7 | 0.0064 | 0.4547 | 2.7317 | 109.1 |
| | | | Total | 311 | 134 | 4.9 | 23.1 | 54.7 | 43.7 | 4.9 | 19.5 | 35.6 | 30.3 | 0.0064 | 0.4749 | 3.1375 | 99.4 |
| Raulí | *HD* | *H* | Coastal rainfed | 14 | 35 | 8.8 | 21.6 | 43.5 | 42.0 | 10.9 | 21.4 | 31.2 | 23.1 | | | | |
| | | | Inner rainfed | 5 | 9 | 10.4 | 20.4 | 32.0 | 43.7 | 12.7 | 19.4 | 30.0 | 36.8 | | | | |
| | | | Foothills | 56 | 132 | 5.4 | 22.4 | 52.0 | 44.6 | 7.0 | 19.9 | 33.4 | 29.1 | | | | |
| | | | Mountains | 202 | 463 | 4.6 | 18.9 | 54.8 | 47.6 | 4.2 | 17.2 | 29.5 | 32.5 | | | | |
| | | | Total | 277 | 639 | 4.6 | 19.8 | 54.8 | 46.9 | 4.2 | 18.0 | 33.4 | 32.0 | | | | |
| | *VDH* | $\sigma_V^2$ | Coastal rainfed | 15 | 10 | 8.9 | 23.7 | 36.5 | 42.1 | 11.5 | 21.5 | 28.1 | 25.4 | 0.0502 | 0.5247 | 1.5057 | 82.6 |
| | | | Inner rainfed | 3 | 2 | 10.4 | 15.0 | 21.9 | 40.6 | 14.2 | 16.6 | 18.7 | 13.6 | 0.0542 | 0.1527 | 0.2859 | 78.4 |
| | | | Foothills | 57 | 39 | 6.2 | 22.9 | 52.0 | 40.5 | 8.0 | 20.1 | 33.4 | 27.6 | 0.0119 | 0.4601 | 2.3675 | 92.7 |
| | | | Mountains | 63 | 45 | 5.6 | 22.0 | 47.9 | 41.8 | 8.8 | 20.2 | 31.2 | 24.6 | 0.0102 | 0.4208 | 1.7425 | 83.0 |
| | | | Total | 138 | 96 | 5.6 | 22.4 | 52.0 | 41.3 | 8.0 | 20.2 | 33.4 | 25.9 | 0.0102 | 0.4425 | 2.3675 | 88.1 |
| | | *V* | Coastal rainfed | 24 | 10 | 8.8 | 22.8 | 36.5 | 35.4 | 10.9 | 22.0 | 31.2 | 23.9 | 0.0337 | 0.4722 | 1.1140 | 66.3 |
| | | | Inner rainfed | 6 | 2 | 10.4 | 14.9 | 23.6 | 33.3 | 14.2 | 16.5 | 20.4 | 15.6 | 0.0542 | 0.1448 | 0.3191 | 66.4 |
| | | | Foothills | 92 | 39 | 5.4 | 22.7 | 52.0 | 41.9 | 8.0 | 20.3 | 33.4 | 28.1 | 0.0104 | 0.4652 | 2.3675 | 87.4 |
| | | | Mountains | 103 | 45 | 4.6 | 21.1 | 47.9 | 43.5 | 7.1 | 19.5 | 31.2 | 26.5 | 0.0072 | 0.3810 | 2.0573 | 92.7 |
| | | | Total | 225 | 96 | 4.6 | 21.8 | 52.0 | 42.2 | 7.1 | 20.0 | 33.4 | 27.1 | 0.0072 | 0.4189 | 2.3675 | 88.7 |
| Coigüe | *HD* | *H* | Coastal rainfed | 7 | 10 | 11.0 | 21.4 | 42.5 | 39.2 | 9.0 | 16.8 | 30.7 | 42.8 | | | | |
| | | | Inner rainfed | 8 | 15 | 5.3 | 21.1 | 36.5 | 48.7 | 9.7 | 19.4 | 33.7 | 37.1 | | | | |
| | | | Rainfed valleys | 9 | 13 | 5.0 | 18.9 | 33.0 | 44.7 | 7.0 | 15.9 | 28.0 | 32.0 | | | | |
| | | | Foothills | 17 | 37 | 8.1 | 25.4 | 49.1 | 37.9 | 10.6 | 21.8 | 31.7 | 22.5 | | | | |
| | | | Mountains | 79 | 180 | 4.9 | 19.4 | 60.2 | 58.1 | 4.2 | 16.6 | 33.7 | 36.4 | | | | |
| | | | Total | 120 | 255 | 4.9 | 19.9 | 60.2 | 51.7 | 4.2 | 17.6 | 33.7 | 35.6 | | | | |
| | *VDH* | $\sigma_V^2$ | Coastal rainfed | 5 | 3 | 11.0 | 17.3 | 42.5 | 56.6 | 9.5 | 14.8 | 30.7 | 45.0 | 0.0375 | 0.2598 | 1.7325 | 160.6 |
| | | | Inner rainfed | 7 | 5 | 13.9 | 20.2 | 32.0 | 26.6 | 12.0 | 18.8 | 31.8 | 28.0 | 0.0685 | 0.4578 | 1.1095 | 86.7 |
| | | | Rainfed valleys | 7 | 5 | 5.0 | 15.3 | 24.0 | 48.5 | 7.0 | 14.1 | 20.0 | 34.7 | 0.0067 | 0.1712 | 0.4045 | 90.0 |
| | | | Foothills | 17 | 11 | 8.5 | 25.9 | 42.4 | 34.6 | 9.0 | 22.6 | 31.7 | 23.8 | 0.0252 | 0.6347 | 1.7741 | 74.5 |
| | | | Mountains | 51 | 35 | 5.1 | 19.6 | 53.2 | 59.3 | 5.6 | 17.5 | 31.8 | 34.3 | 0.0068 | 0.3869 | 2.8389 | 140.3 |
| | | | Total | 87 | 59 | 5.0 | 20.4 | 53.2 | 52.1 | 5.6 | 18.2 | 31.8 | 34.0 | 0.0067 | 0.4025 | 2.8389 | 122.5 |
| | | *V* | Coastal rainfed | 9 | 3 | 11.0 | 18.3 | 34.1 | 44.2 | 9.5 | 15.6 | 30.7 | 31.6 | 0.0375 | 0.2504 | 1.7325 | 125.8 |
| | | | Inner rainfed | 11 | 5 | 10.0 | 20.6 | 32.0 | 34.3 | 12.0 | 18.9 | 31.8 | 26.3 | 0.0388 | 0.3164 | 1.1095 | 74.1 |
| | | | Rainfed valleys | 11 | 5 | 5.0 | 20.7 | 30.1 | 39.0 | 7.0 | 17.7 | 27.4 | 29.6 | 0.0290 | 0.3588 | 0.9308 | 84.7 |
| | | | Foothills | 26 | 11 | 6.9 | 23.7 | 45.3 | 41.4 | 7.5 | 20.9 | 31.8 | 29.1 | 0.0175 | 0.5153 | 1.8176 | 89.8 |
| | | | Mountains | 83 | 35 | 5.1 | 20.4 | 51.3 | 52.6 | 6.3 | 17.8 | 31.8 | 31.5 | 0.0068 | 0.3976 | 2.5931 | 125.5 |
| | | | Total | 140 | 59 | 5.0 | 20.9 | 51.3 | 47.7 | 6.3 | 18.3 | 31.8 | 31.0 | 0.0067 | 0.4006 | 2.5931 | 113.4 |

*HD*: height–diameter database; *VDH*: total volume underbark database; $n_{f,v}$: sample size for fitting and validation, respectively, of total height equations (*H*), variance of the total volume underbark ($\sigma_V^2$) and total volume underbark (*V*); *CV*: coefficient of variation; Max.: maximum; Min.: minimum.

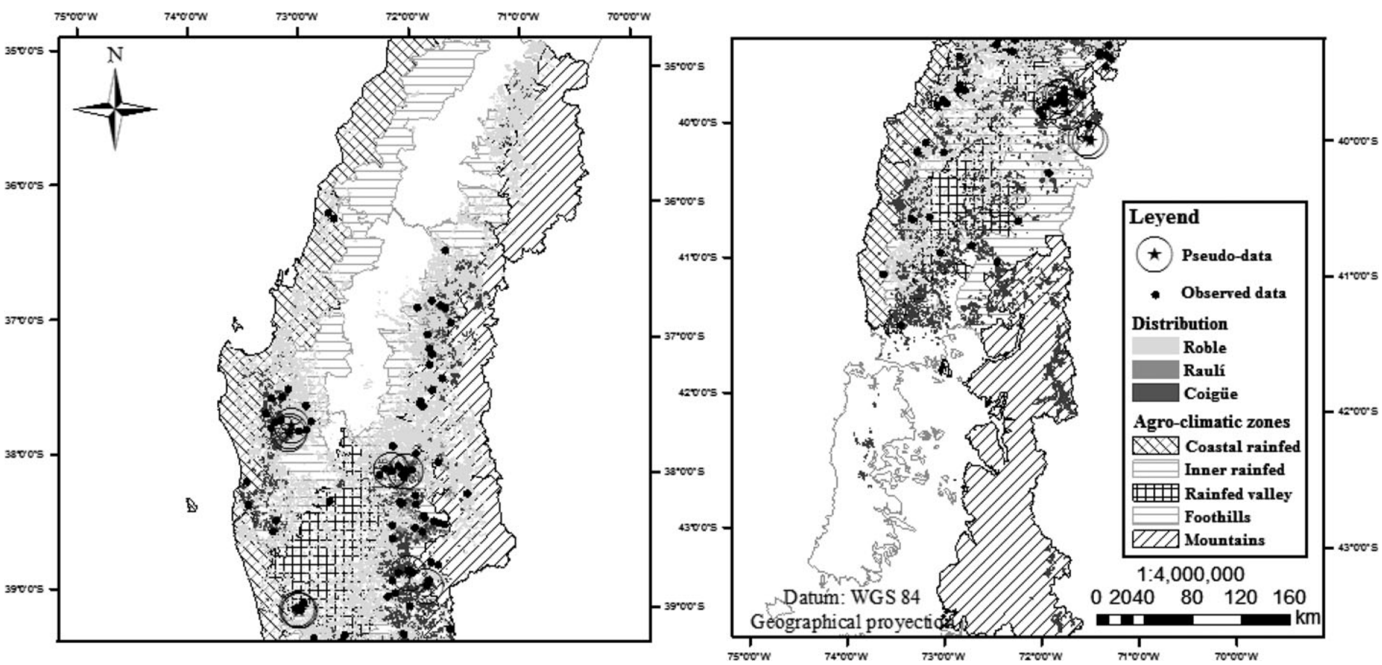


**Figure 1** Agro-climatic zones, geographic distribution of the Roble-Raulí-Coigüe Forest Type, location of sampling points associated to allometric equations of total volume underbark compiled and sampling points associated to the observed database.

are required in order to identify data of $z - 1$ in those zones. In the case of data from five zones, model (1) with *dummy* variables assumes the following form (7):

$$V = (a_1 + a_{11}d_1 + a_{12}d_2 + a_{13}d_3 + a_{14}d_4)(D^2H)^{(a_2+a_{21}d_1+a_{22}d_2+a_{23}d_3+a_{24}d_4)} \qquad (7)$$

where $V$ is the total volume underbark ($m^3$), $a_{1,2}$ are parameters of total volume model for that zone not marked with *dummy* variables, $d_{1,\ldots,4}$ are the *dummy* variables, with $d_i = 1$ for those data belonging to the $i$th zone, and $d_i = 0$, otherwise; $a_{1i,2i}$ represent the difference between $a_1$ and respectively, $a_2$ obtained in the $i$th zone and the respective values obtained from the fitting for the zone whose data were not marked ($a_{1,2}$). In this study, the zone not marked corresponded to the zone that presented the greatest sample size intended to the fit, i.e. Mountains for the fitting of total height relationship for the three species. Inner rainfed for roble; and Mountains for raulí and coigüe, for the fitting of the volume model from pseudo-data. Foothills for roble, Mountains for raulí and coigüe, for the fitting of the volume model from observed data.

The Furnival's index (*FI*) (Furnival, 1961) and the residual coefficient of variation (*CV*) were used to evaluate the precision of the fitting of total volume model (1), total height model (2) and variance of total volume model (6). *FI* was used to transform the residuals of the weighted fitting of the allometric equations of total volume and total height to express them in their original scale. The accuracy index (*E*) (Bruce, 1975) was used in the evaluation of the three options to estimate the total height from the diameter at breast height (relationships 3, 4 and 5), as well as in the validation of the allometric equations of total height and total volume.

$$FI = RMSE/f'(y)_g \qquad (8)$$

$$CV\,(\%) = 100(IF/\bar{y}) \qquad (9)$$

$$E = \sqrt{\bar{\varepsilon}^2 + s_\varepsilon^2} \qquad (10)$$

where *RMSE* is the root mean square residual (i.e. $RMSE = \sqrt{SSE/(n-p)}$); *SSE* is the sum of squared residuals (i.e. $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$), $f'(y)_g$ is the geometric mean of the first derivative of the dependent variable with respect to the untransformed dependent variable (i.e. $e^{\left[\sum_{i=1}^{n}\ln(f'(y))\right]/n}$); $e$: exponential; ln: natural logarithm; $\bar{\varepsilon}$: average deviation (i.e. $\sum_{i=1}^{n}\varepsilon_i/n$); $s_\varepsilon^2$: variance of residuals (i.e. $\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon}_i)^2/(n-1)$); $y_i$, $\hat{y}_i$ and $\bar{y}$ observed value, value estimated by the model and average value of the dependent variable; $p$: number of parameters of the model; $n$: sample size.

## Results

### *Modelling total height*

For all the species and zones, the fitting of the total height model (2) showed all the parameters statistically significant at $\alpha = 0.05$ (Table 4). The model explained a high proportion of total height variability, despite every zone bringing together data from stands of varied structure and site conditions (Figure 2, upper). For roble, the values of *FI* and residual *CV* were lower than 3.4 m and 18.6 per cent, respectively. For raulí, values of *FI* and *CV* were lower than 3.5 m and 19.4 per cent, respectively. Finally, for coigüe, values of *FI* and *CV* were lower than 3.3 m and 17.6 per cent, respectively. The weighted fitting solved the problem of heteroscedasticity in the three species (Figure 2, lower).

In the fitting including *dummy* variables, no zonal effects were detected in the regression parameters ($P(t) \geq 0.05$). No significant zone differences were detected in the parameters $b_1$ and $b_2$ of the equations obtained in Coastal rainfed, Inner rainfed, Rainfed valley and Foothills with respect to the equation obtained in the Mountains for the tree species. Therefore, the equation obtained from the general fitting for

**Table 4** Parameter estimates and fit statistics for height–diameter equations, by species and zone.

| Species | Zone | $n_f$ | Parameters | | | | *FI* (m) | *CV* (%) | $E_1$ (m) | $E_2$ (m) | $E_3$ (m) | $n_v$ | *E* (m) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $b_1$ | $b_1d_i$ | $b_2$ | $b_2d_i$ | | | | | | | |
| Roble | General fitting | 225 | 3.1616** | | 0.5631** | | 3.3497 | 17.6 | 4.9307 | 3.1287 | 3.4079 | 528 | 3.4980 |
| | Coastal rainfed | 11 | 5.1806* | 2.2720ns | 0.3936** | −0.1999ns | 2.3566 | 11.9 | 2.6129 | 2.2780 | 2.3261 | 27 | 2.5314 |
| | Inner rainfed | 21 | 2.3053** | −0.6033ns | 0.6024** | 0.0089ns | 2.1077 | 14.1 | 3.2199 | 2.3753 | 2.8953 | 49 | 3.2096 |
| | Rainfed valley | 17 | 1.7023* | −1.2063ns | 0.7928** | 0.1993ns | 2.5151 | 13.5 | 3.1172 | 2.5478 | 4.8663 | 41 | 3.7818 |
| | Foothills | 84 | 4.0103* | 1.1080ns | 0.4989** | −0.0946ns | 3.3977 | 16.1 | 3.5927 | 3.5735 | 4.3541 | 215 | 3.4503 |
| | Mountains | 92 | 2.9086** | | 0.5935** | | 3.3083 | 18.5 | 4.6797 | 3.4369 | 3.7287 | 215 | 3.3480 |
| Raulí | General fitting | 277 | 2.9762** | | 0.5878** | | 3.3028 | 18.3 | 5.5260 | 3.3717 | 3.5365 | 639 | 3.2868 |
| | Coastal rainfed | 14 | 5.4415* | 2.3639ns | 0.4324** | −0.1363ns | 3.4879 | 16.3 | 5.3579 | 3.1823 | 3.3173 | 35 | 3.6731 |
| | Inner rainfed | 7 | 1.5227* | −1.5551ns | 0.8238** | 0.2551ns | 2.9377 | 15.2 | 3.4705 | 2.5547 | 3.0440 | 9 | 2.8966 |
| | Foothills | 56 | 3.0882** | 0.0106ns | 0.5856** | 0.0169ns | 2.9258 | 14.7 | 3.4705 | 2.9517 | 3.0440 | 132 | 2.8754 |
| | Mountains | 200 | 3.0776* | | 0.5687** | | 3.3143 | 19.3 | 5.4808 | 3.1954 | 3.3614 | 463 | 3.3088 |
| Coigüe | General fitting | 120 | 2.4294** | | 0.6329** | | 2.9060 | 17.1 | 4.5362 | 3.0333 | 3.5898 | 255 | 3.3724 |
| | Coastal rainfed | 7 | 1.2771* | −1.3472ns | 0.8202** | 0.2206ns | 1.6664 | 9.8 | 1.7902 | 1.5375 | 2.3127 | 10 | 2.9374 |
| | Inner rainfed | 8 | 0.7759** | −1.8485ns | 1.0279** | 0.4283ns | 0.8821 | 4.7 | 1.2196 | 0.8148 | 1.3188 | 15 | 3.7454 |
| | Rainfed valley | 9 | 2.3074** | −0.3170ns | 0.6350** | 0.0353ns | 2.3870 | 15.0 | 2.5281 | 2.4535 | 4.4909 | 13 | 2.7394 |
| | Foothills | 17 | 1.8710* | −0.7534ns | 0.7525** | 0.1529ns | 3.2709 | 16.2 | 4.4551 | 3.1552 | 3.2103 | 37 | 3.6879 |
| | Mountains | 79 | 2.6243* | | 0.5996** | | 2.8606 | 17.5 | 4.5979 | 2.9608 | 2.8882 | 180 | 3.1795 |

*$P(t)$ < 0.05, **$P(t)$ < 0.0001; ns: non-significant ($P(t)\geq 0.05$); $n_f$: sample size for fitting; $b_1d_i$,$b_2d_i$: difference between parameters $b_1$ and, respectively, $b_2$ of each zone with respect to those obtained in mountains; *FI*: Furnival's index; *CV*: residual coefficient of variation; $E_{1,2,3}$: accuracy of the option positioning the estimation line on minimum, medium and maximum height, respectively; $n_v$: sample size for validation; *E*: accuracy index.

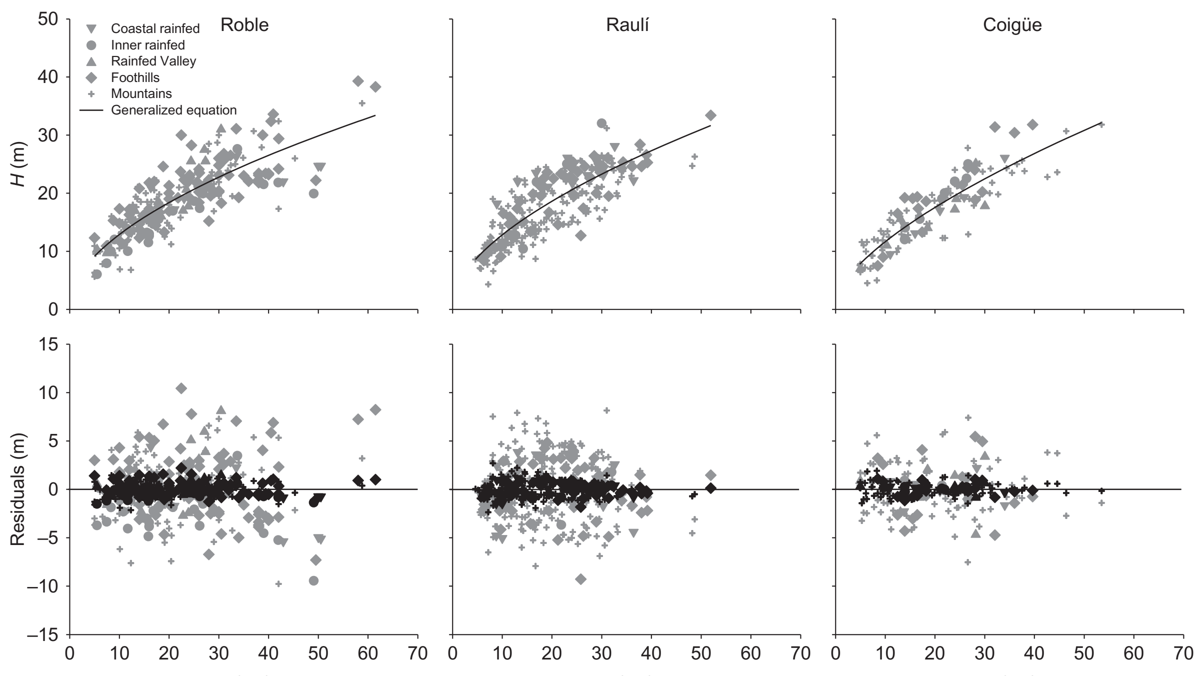


**Figure 2** The upper graphs represent the relationship between *D* (cm) and *H* (m) for roble, raulí and coigüe, in the sample used for fitting. Symbols in grey represent observed data from different agro-climatic zones. Solid line represents the generalized equation. Lower graphs represent the observed residuals (in grey) and the weighted residuals (in black).

each species is valid for all zones, i.e. $H = 1.3 + 3.1616D^{0.5631}$, $H = 1.3 + 2.9762D^{0.5878}$ and $H = 1.3 + 2.4294D^{0.6329}$, whose *FI* was 3.3497, 3.3028 and 2.9060 m for roble, raulí and coigüe, respectively. In the validation, the greatest accuracy was presented by raulí, followed by coigüe and roble, i.e. $E < 3.3$, $E < 3.4$ and $E < 3.5$ m, respectively. The estimation line for each species

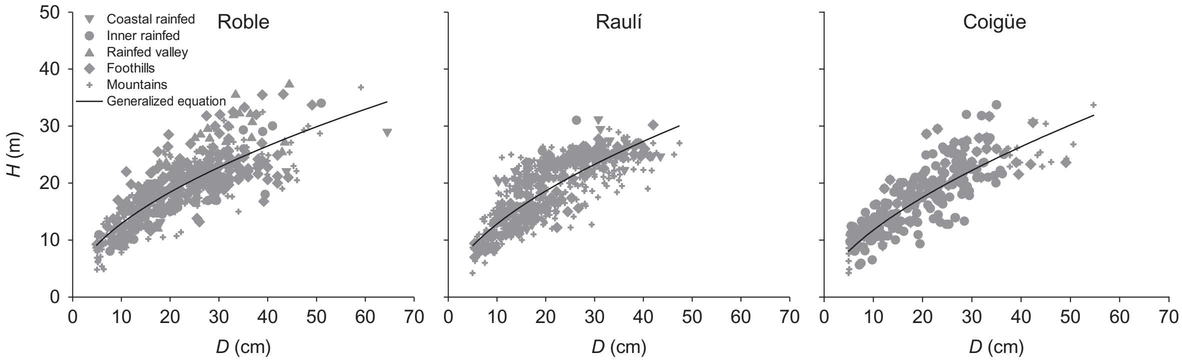


**Figure 3** Relationship between $D$ (cm) and $H$ (m) for roble, raulí and coigüe, in the sample used for validation. Symbols in grey represent observed data from different agro-climatic zones. Solid line represents the generalized equation.

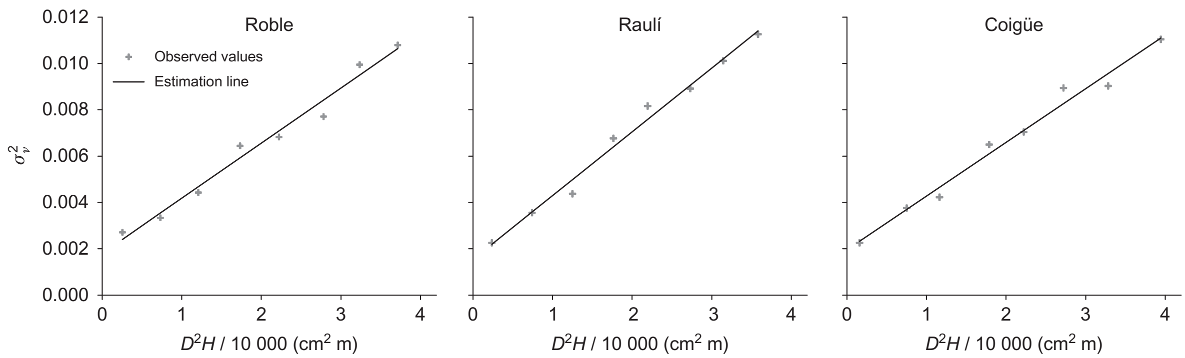


**Figure 4** Relationship between the variance of total volume and $D^2H$ (cm$^2$ m) for roble, raulí and coigüe, in the sample used for fitting. Cross symbols represent observed values in each class. Solid line represents the estimation line.

is centrally positioned at the total set of observed validation data coming from different zones (Figure 3).

In the estimation of the total height, which corresponds to a specific diameter value, the most accurate results were obtained by positioning the estimation line at the average height (relationship 4). This option presented the greatest accuracy in all zones, for the three analysed species (column $E_2$, in Table 4). Thus, relationship (4) was used to estimate the heights used to generate pseudo-data.

### Modelling of the variance of total volume

For the three species, a linear trend of increase of the variance of total volume was observed, as the independent variable increases ($D^2H$) (Figure 4). The fitting of the variance model of total volume (7), by species provided equations in which all parameters were significant ($P(t) < 0.05$) (Table 5). The greatest accuracy in terms of fitting was achieved for raulí, followed by coigüe and roble, with residual coefficients of variation 5.6, 6.5 and 7.1 per cent, respectively.

### Modelling of total volume from pseudo-data

For all the species and zones, the weighted fitting of the total volume model (1) showed all the parameters statistically significant at $\alpha = 0.05$ (Table 6). In the fitting including *dummy* variables, no zonal effects were detected in the regression parameters for the three species. Therefore, equations obtained from the general fitting for each species are valid for the zones that provided pseudo-data for the fitting. The accuracy of the equations obtained from the general fitting is similar or even higher than the original equations used in the generation of pseudo-data. Weighted fitting solved the problem of heteroscedasticity, and for the three species the weighted residuals presented a constant trend against the predictor variable (Figure 5, lower).

For roble, *FI* and *CV* values were lower than 0.0879 m$^3$ tree$^{-1}$ and 14.5 per cent, respectively. In the fitting, no significant zone differences were detected in the parameters

**Table 5** Parameter estimates and fit statistics for variance of total volume equations, by species.

| Species | $n_c$ | Parameters | | *FI* | *CV* (%) |
|---|---|---|---|---|---|
| | | $c_0$ | $c_1$ | | |
| Roble | 8 | 1.7997E−03* | 2.3796E−07** | 4.6129E−04 | 7.1 |
| Raulí | 8 | 1.5435E−03* | 2.7510E−07** | 3.8719E−04 | 5.6 |
| Coigüe | 8 | 1.9584E−03* | 2.3163E−07** | 4.2885E−04 | 6.5 |

*$P(t) < 0.05$; **$P(t) < 0.0001$; $n_c$: total number of classes of $D^2H$.

**Table 6** Parameter estimates and fit statistics for total volume equations, by species and zone.

| Species | Zone | $n_{fu}$ | $n_p$ | Parameters | | | | $FI$ ($m^3$) | $CV$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $a_1$ | $a_1d_i$ | $a_2$ | $a_2d_i$ | | |
| Roble | General fitting | 9 | 575 | 3.4915E–05** | | 0.9896** | | 0.0851 | 9.4 |
| | Inner rainfed | 3 | 402 | 3.2821E–05** | | 0.9948** | | 0.0817 | 9.4 |
| | Rainfed valley | 2 | 60 | 3.9440E–05** | 6.6193E–06ns | 0.9777** | −1.7158E–02ns | 0.0878 | 6.8 |
| | Foothills | 2 | 48 | 3.9264E–05** | 6.4429E–06ns | 0.9804** | −1.4404E–02ns | 0.0833 | 7.8 |
| | Mountains | 2 | 65 | 2.7208E–05** | −5.6120E–06ns | 1.0186** | 2.3763E–02ns | 0.0876 | 14.4 |
| Raulí | General fitting | 7 | 318 | 3.4704E–05** | | 0.9879** | | 0.0651 | 14.2 |
| | Foothills | 4 | 140 | 3.6155E–05** | 2.5546E–06ns | 0.9843** | −0.0061ns | 0.0734 | 12.6 |
| | Mountains | 3 | 178 | 3.3600E–05** | | 0.9905** | | 0.0589 | 16.3 |
| Coigüe | General fitting | 2 | 80 | 3.8854E–05** | | 0.9808** | | 0.0679 | 8.8 |
| | Rainfed valley | 1 | 30 | 5.3112E–05** | 1.7145E–05ns | 0.9534** | −0.0334ns | 0.0831 | 7.6 |
| | Mountains | 1 | 50 | 3.5967E–05** | | 0.9868** | | 0.0504 | 10.2 |

**$P(t) < 0.0001$; ns: non-significant (i.e. $P(t) \geq 0.05$); $n_{fu}$: number of equations for the generation of pseudo-data; $n_p$: total number of generated pseudo-data; $a_1d_i$,$a_2d_i$: difference between parameters $a_1$ and, respectively, $a_2$ of each zone with respect to data obtained from the non-marked zone.

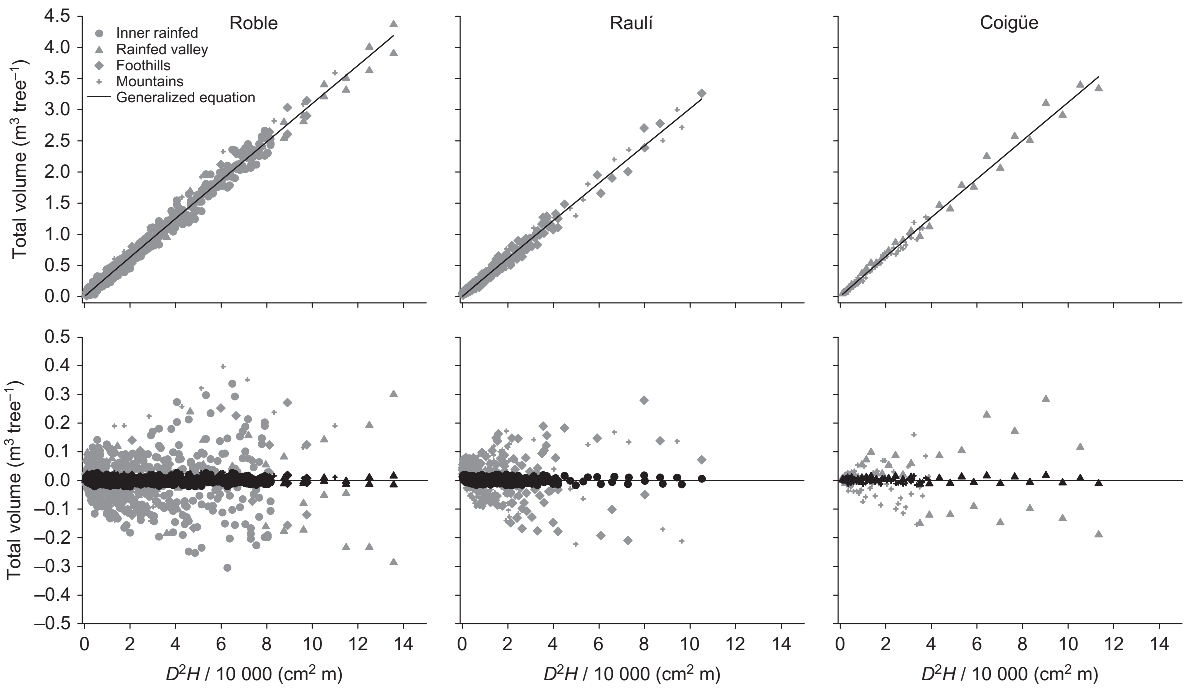


**Figure 5** The upper graphs represent the relationship between total volume ($m^3$) and $D^2H$ ($cm^2$ m) for roble, raulí and coigüe, in the sample used for fitting. Symbols in grey represent pseudo-data from different agro-climatic zones. Solid line represents the generalized equation. Lower graphs represent the observed residuals (in grey) and weighted residuals (in black).

$a_1$ and $a_2$ of the equations obtained in Rainfed valley, Foothills and Mountains with respect to the equation obtained for Inner rainfed, i.e. $V = 0.000032821(D^2H)^{0.9948}$. Therefore, the equation obtained from the general fitting is valid for the four zones that contributed with pseudo-data for the fitting, i.e. $V = 0.000034915(D^2H)^{0.9896}$. For the general fitting, an $FI$ = 0.0851 $m^3$ $tree^{-1}$, equivalent to a $CV$ = 9.4 per cent was obtained. The estimation line is centrally positioned on the total set of pseudo-data of the general fitting, showing that the generalized equation is suitable for the zones that contributed with pseudo-data for fitting (Figure 5, upper, left).

For raulí, $FI$ and $CV$ values were lower than 0.0735 $m^3$ $tree^{-1}$ and 16.4 per cent, respectively. The equation of total volume of the Mountains zone, i.e. $V = 0.000036000(D^2H)^{0.9905}$, is not different from that of the Foothills zone, i.e. $V = 0.000036155(D^2H)^{0.9843}$. Therefore, the equation obtained from the general fitting is valid for the two zones that contributed with pseudo-data for the fitting, i.e. $V = 0.000034704(D^2H)^{0.9879}$. For the general fitting, an $FI$ = 0.0651 m$^3$ tree$^{-1}$, equivalent to a $CV$ = 14.2 per cent was obtained. The estimation line is centrally positioned on the pseudo-data of the general fitting, showing that the generalized equation is suitable for both zones that contributed pseudo-data for the fitting (Figure 5, upper, centre).

For coigüe, $FI$ and $CV$ values were lower than 0.0832 m$^3$ tree$^{-1}$ and 10.3 per cent, respectively. Because in the fitting no significant zonal effects were detected, the equation obtained from the general fitting is valid for the two zones that contributed with pseudo-data for the fitting, i.e. $V = 0.000038854(D^2H)^{0.9808}$. For the general fitting an $FI$ = 0.0679 m$^3$ tree$^{-1}$, equivalent to a $CV$ = 8.8 per cent was obtained. The estimation line is centrally positioned on pseudo-data of the general fitting, showing that the generalized equation is suitable for both zones that contributed with pseudo-data for the fitting (Figure 5, upper, right).

### *Checking the validity of the equations fitting method from pseudo-data*

For all the species and zones, the weighted fitting of the total volume model (1), both from observed data and from pseudo-data, showed all the parameters were statistically significant at $\alpha = 0.05$ (Table 7). As expected, for the three species, the fitting

**Table 7** Parameter estimates and fit statistics for total volume equations, by species, type of data and zone.

| Species | Type of data | Zone | $n_f$ | Parameters | | | | $FI$ (m$^3$) | $CV$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $a_1$ | $a_1d_i$ | $a_2$ | $a_2d_i$ | | |
| Roble | Observed | General fitting | 311 | 3.7897E−05** | | 0.9847** | | 0.0586 | 12.3 |
| | | Coastal rainfed | 14 | 6.8519E−05** | 2.5130E−05ns | 0.9161** | −5.2802E−02ns | 0.0213 | 7.9 |
| | | Inner rainfed | 35 | 3.9058E−05** | −4.3316E−06ns | 0.9916** | 2.2765E−02ns | 0.0432 | 9.7 |
| | | Rainfed valley | 29 | 7.7563E−05** | 3.4174E−05ns | 0.9075** | −6.1380E−02ns | 0.0377 | 7.7 |
| | | Foothills | 128 | 4.3390E−05** | | 0.9689** | | 0.0541 | 10.4 |
| | | Mountains | 105 | 4.1401E−05** | −1.9878E−06ns | 0.9792** | 1.0340E−02ns | 0.0562 | 12.4 |
| | Pseudo-data | General fitting | 311 | 3.8423E−05** | | 0.9837** | | 0.0608 | 10.5 |
| | | Coastal rainfed | 14 | 7.3224E−05** | 3.4606E−05ns | 0.9125** | −6.8160E−02ns | 0.0423 | 7.9 |
| | | Inner rainfed | 35 | 3.6235E−05** | −2.3838E−06ns | 0.9977** | 1.7030E−02ns | 0.0664 | 7.0 |
| | | Rainfed valley | 29 | 1.6088E−04** | 1.2226E−04ns | 0.8415** | −1.3914E−01ns | 0.0795 | 11.8 |
| | | Foothills | 128 | 3.8618E−05** | | 0.9806** | | 0.0763 | 9.3 |
| | | Mountains | 105 | 4.1869E−05** | 3.2507E−06ns | 0.9777** | −2.9465E−03ns | 0.0816 | 9.0 |
| Raulí | Observed | General fitting | 225 | 3.7442E−05** | | 0.9867** | | 0.0452 | 10.7 |
| | | Coastal rainfed | 24 | 7.5797E−05** | 4.2754E−05ns | 0.9180** | −7.9873E−02ns | 0.0404 | 8.6 |
| | | Inner rainfed | 6 | 5.8025E−05** | 2.4983E−05ns | 0.9450** | −5.2872E−02ns | 0.0289 | 6.2 |
| | | Foothills | 92 | 3.7577E−05** | 4.5342E−06ns | 0.9869** | −1.0964E−02ns | 0.0415 | 8.8 |
| | | Cordillera | 103 | 3.3043E−05** | | 0.9979** | | 0.0498 | 13.1 |
| | Pseudo-data | General fitting | 225 | 3.8387E−05** | | 0.9840** | | 0.0734 | 10.8 |
| | | Coastal rainfed | 24 | 3.4293E−05** | 3.5870E−07ns | 0.9972** | 1.3494E−03ns | 0.0634 | 9.7 |
| | | Inner rainfed | 6 | 2.2378E−05** | −1.1556E−05ns | 1.0576** | 6.1767E−02ns | 0.0327 | 12.2 |
| | | Foothills | 92 | 3.6878E−05** | 2.9438E−06ns | 0.9871** | −8.7794E−03ns | 0.0782 | 9.9 |
| | | Mountains | 103 | 3.3935E−05** | | 0.9959** | | 0.0732 | 11.9 |
| Coigüe | Observed | General fitting | 140 | 3.9937E−05** | | 0.9825** | | 0.0403 | 10.1 |
| | | Coastal rainfed | 9 | 3.3637E−05** | −2.701E−06ns | 0.9977** | 4.1216E−03ns | 0.0164 | 6.5 |
| | | Inner rainfed | 11 | 3.4867E−05** | −1.471E−06ns | 0.9898** | −3.8617E−03ns | 0.0223 | 7.0 |
| | | Rainfed valley | 11 | 1.0973E−05** | −2.537E−05ns | 1.1249** | 1.3127E−01ns | 0.0309 | 8.6 |
| | | Foothills | 26 | 6.4066E−05** | 2.773E−05ns | 0.9317** | −6.1943E−02ns | 0.0439 | 8.5 |
| | | Mountains | 83 | 3.6338E−05** | | 0.9936** | | 0.0347 | 8.7 |
| | Pseudo-data | General fitting | 140 | 4.1482E−05** | | 0.9772** | | 0.0656 | 9.4 |
| | | Coastal rainfed | 9 | 4.0249E−05** | −9.6108E−06ns | 0.9741** | 1.2753E−02ns | 0.0533 | 8.6 |
| | | Inner rainfed | 11 | 2.5273E−05** | −2.4586E−05ns | 1.0246** | 6.3271E−02ns | 0.0667 | 14.2 |
| | | Rainfed valley | 11 | 2.1439E−05** | −2.8420E−05ns | 1.0534** | 9.2100E−02ns | 0.0386 | 11.1 |
| | | Foothills | 26 | 5.9484E−05** | 9.6249E−06ns | 0.9361** | −2.5161E−02ns | 0.0595 | 11.2 |
| | | Mountains | 83 | 4.9859E−05** | | 0.9613** | | 0.0657 | 7.9 |

**$P(t) < 0.0001$; ns: not significant ($P(t) \geq 0.05$); $n_f$: sample size for fitting; $a_1d_i$, $a_2d_i$: difference between parameters $a_1$ and, respectively, $a_2$ of each zone with respect to those obtained in the non-marked zone.

**Table 8** Confidence intervals for estimated parameters of total volume equations, by species and type of data.

| Species | Type of data | $n_f$ | Parameters | | | |
|---|---|---|---|---|---|---|
| | | | $a_1$ | | $a_2$ | |
| | | | $L_l$ | $L_u$ | $L_l$ | $L_u$ |
| Roble | Observed | 311 | 3.0481E–05 | 4.5313E–05 | 0.9658 | 1.0036 |
| | Pseudo-data[d] | 311 | 3.0628E–05 | 4.8552E–05 | 0.9600 | 1.0020 |
| | Pseudo-data[f] | 575 | 3.0094E–05 | 3.9736E–05 | 0.9770 | 1.0022 |
| | Pseudo-data[f] | 10 000 | 3.0660E–05 | 3.3947E–05 | 0.9947 | 1.0043 |
| | Pseudo-data[fM] | 10 000 | 3.9184E–05 | 4.3167E–05 | 0.9754 | 0.9840 |
| Raulí | Observed | 225 | 2.9698E–05 | 4.5187E–05 | 0.9663 | 1.0071 |
| | Pseudo-data[d] | 225 | 3.0257E–05 | 5.0455E–05 | 0.9550 | 1.0015 |
| | Pseudo-data[f] | 318 | 2.8141E–05 | 4.1267E–05 | 0.9700 | 1.0058 |
| | Pseudo-data[f] | 10 000 | 3.3122E–05 | 3.3639E–05 | 0.9876 | 0.9946 |
| | Pseudo-data[fM] | 10 000 | 3.0784E–05 | 3.1882E–05 | 0.9929 | 0.9961 |
| Coigüe | Observed | 140 | 3.1595E–05 | 4.8278E–05 | 0.9622 | 1.0027 |
| | Pseudo-data[d] | 140 | 3.0136E–05 | 5.2829E–05 | 0.9516 | 1.0028 |
| | Pseudo-data[f] | 80 | 2.8136E–05 | 4.9572E–05 | 0.9555 | 1.0062 |
| | Pseudo-data[f] | 10 000 | 3.2160E–05 | 3.5082E–05 | 0.9946 | 1.0029 |
| | Pseudo-data[fM] | 10 000 | 3.7365E–05 | 3.9492E–05 | 0.9784 | 0.9833 |

[d]: Pseudo-data generated from the equation fitted in this study with observed data; [f]: pseudo-data generated with compiled equations; [fM]: pseudo-data generated with compiled equations using Monte-Carlo simulation; $n_f$: sample size for fitting; $L_{l,u}$: lower and upper limit of the confidence intervals, respectively.

performed from observed data presented the lowest *FI* values. For both types of data, no significant zone effects were detected in any of the regression parameters. Therefore, the generalized equations obtained from the general fitting from observed data are $V = 0.000037897(D^2H)^{0.9847}$, $V = 0.000037442(D^2H)^{0.9867}$ and $V = 0.000039937(D^2H)^{0.9825}$, for roble, raulí and coigüe, respectively. While the respective generalized equations obtained from pseudo-data are $V = 0.000038423(D^2H)^{0.9837}$, $V = 0.000038387(D^2H)^{0.9840}$ and $V = 0.000041482(D^2H)^{0.9772}$, for roble, raulí and coigüe, respectively. Then, comparing both types of data, the estimated parameters of the total volume equations developed from pseudo-data not presented statistically significant differences with respect to the estimated parameters of the total volume equations developed from observed data. These results were verified in Table 8, where the estimated parameter confidence intervals overlap each other for the three species. The result showed that the method used in this study to generate pseudo-data from compiled equations and then developing generalized allometric equations of total volume is suitable for the three species. In turn, the generalized allometric equations of total volume developed from the methodology proposed by this study and from a Monte-Carlo simulation (i.e. 10 000 pseudo-data, respectively) not present significant differences with the parameters generated with observed data.

### *Validation of allometric equations of total volume underbark*

For the three species, the generalized allometric equations fitted in this study from observed data presented a small advantage in accuracy with respect to the equations fitted to pseudo-data (i.e. $E < 0.0642$, $E < 0.0658$ and $E < 0.0674$ m$^3$ tree$^{-1}$, respectively) (Table 9 and Figure 6). The compiled equations presented a lower accuracy ($E < 0.1116$ m$^3$ tree$^{-1}$).

## Discussion

In this study, the first generalized allometric equations of total volume for roble, raulí and coigüe in second growth in Chile forests are presented. These equations were developed using a modified meta-analysis technique that has been used in North America and Europe (Pastor *et al.*, 1983/1984; Jenkins *et al.*, 2003; Muukkonen, 2007; Chojnacky *et al.*, 2014). This technique involves collecting published equations which are used to generate pseudo-data, which then are used in the fitting of new equations. These are then analysed in order to determine their validity as generalized equations.

The meta-analysis used in this study was supplemented with a base of observed data (Table 3). This database was used to fit the height–diameter relationship, as well as a variance equation of total volume for each species, in order to solve the problem of heteroscedasticity and autocorrelation of the pseudo-data generated from the allometric equations of total volume compiled. Results show that the generalized allometric equations of total volume for each species obtained through non-linear least squares are precise and consistent (Table 6). An alternative way to generate non-autocorrelated data has been proposed by Wayson *et al.* (2015), who used the determination coefficient ($R^2$), *n* and *D* range with a Monte-Carlo simulation. In this study, both methodologies were tested simulating 10 000 pseudo-data, to validate the generalized

**Table 9** Validation of generalized and local equations to estimate total volume, by species.

| Species | Zone | $n$ | Equations | $n_v$ | $E$ (m$^3$) |
|---|---|---|---|---|---|
| Roble | Ir-Rv-Fo-Mo | 575 | $V^p = 0.000034915(D^2H)^{0.9896}$ | 134 | **0.0673** |
| | Cr-Ir-Rv-Fo-Mo | 311 | $V^d = 0.000037897(D^2H)^{0.9847}$ | | **0.0641** |
| | | | $V^{pd} = 0.000038423(D^2H)^{0.9837}$ | | **0.0657** |
| | Ir-Rv-Fo-Mo | 10 000 | $V^p = 0.000034331(D^2H)^{0.9936}$ | | **0.0652** |
| | | | $V^{pM} = 0.000041175(D^2H)^{0.9797}$ | | **0.0663** |
| | Ir | 134 | $V^l = 0.005000 + 0.000031510D^2H$ | | 0.0698 |
| | | | $V^l = 0.009400 + 0.000034130D^2H - 0.00006350D^2$ | | 0.0664 |
| | | | $V^l = 0.001789 + 0.000036110D^2H - 0.00012990D^2 - 0.0012H$ | | 0.0788 |
| | Rv | 30 | $V^l = 0.024623 + 0.000030644D^2H$ | | 0.0662 |
| | | | $V^l = 0.025957 + 0.000030292D^2H$ | | 0.0677 |
| | Fo | 24 | $V^l = 0.003230 + 0.000032140D^2H$ | | 0.0697 |
| | | | $V^l = 0.003524 + 0.000038263D^2H$ | | 0.0752 |
| | Mo | 15 | $V^l = 0.071120 + 0.000033890D^2H$ | | 0.1084 |
| | | 50 | $V^l = 0.025828 + 0.000028502D^2H$ | | 0.0841 |
| Raulí | Fo-Mo | 318 | $V^p = 0.000038854(D^2H)^{0.9808}$ | 96 | **0.0652** |
| | Cr-Ir-Fo-Mo | 225 | $V^d = 0.000037442(D^2H)^{0.9867}$ | | **0.0638** |
| | | | $V^{pd} = 0.000038387(D^2H)^{0.9840}$ | | **0.0642** |
| | Fo-Mo | 10 000 | $V^p = 0.000037442(D^2H)^{0.9867}$ | | **0.0631** |
| | | | $V^{pM} = 0.000031333(D^2H)^{0.9945}$ | | **0.0735** |
| | Fo | 26 | $V^l = 0.013710 + 0.000028899D^2H$ | | 0.0845 |
| | | 13 | $V^l = 0.014110 + 0.000026890D^2H$ | | 0.1115 |
| | | 21 | $V^l = 0.000850 + 0.000028390D^2H$ | | 0.0986 |
| | | 80 | $V^l = 0.046000 + 0.000029056D^2H$ | | 0.0790 |
| | Mo | 75 | $V^l = 0.002272 + 0.000038871D^2H$ | | 0.0783 |
| | | 28 | $V^l = 0.013370 + 0.000029931D^2H$ | | 0.0767 |
| | | 75 | $V^l = 0.002070 + 0.000030000D^2H$ | | 0.0783 |
| Coigüe | Rv-Mo | 80 | $V^p = 0.000038698(D^2H)^{0.9801}$ | 59 | **0.0456** |
| | Cr-Ir-Rv-Fo-Mo | 140 | $V^d = 0.000039937(D^2H)^{0.9825}$ | | **0.0432** |
| | | | $V^{pd} = 0.000041482(D^2H)^{0.9772}$ | | **0.0442** |
| | Rv-Mo | 10 000 | $V^p = 0.000031329(D^2H)^{0.9987}$ | | **0.0548** |
| | | | $V^{pM} = 0.000038428(D^2H)^{0.9808}$ | | **0.0478** |
| | Rv | 30 | $V^l = 0.024623 + 0.000030644D^2H$ | | 0.0583 |
| | Mo | 50 | $V^l = 0.012105 + 0.000029462D^2H$ | | 0.0646 |

Cr: Coastal rainfed; Ir: Inner rainfed; Rv: Rainfed valley; Fo: Foothills; Mo: Mountains; $n$: sample size; $V^p$: generalized equation fitted with pseudo-data using compiled equations; $V^d$: generalized equation fitted with observed data; $V^{pd}$: generalized equation fitted with pseudo-data, generated with the equation $V^d$; $V^{pM}$: generalized equation fitted with pseudo-data using compiled equations from Monte-Carlo simulation; $V^1$: compiled local equation; $n_v$: sample size for validation; $E$: accuracy index; the highest accuracies are highlighted in bold.

allometric equations of total volume, resulting that the parameters generated with pseudo-data not present significant differences with the parameters generated with observed data (Table 8).

The fitted equations in this study include as a predictive variable the combined variable $D^2H$. This facilitates their utilization in varied conditions of site, management and age (Montagu *et al.*, 2005; Ung *et al.*, 2008; Gilabert and Paci, 2010; Alegría, 2011), because the variability in site quality is reflected in the variability of the stem height and the variability in silvicultural management or natural distribution of trees is reflected in variability of diameters. In turn, the combined variable, i.e. $D^2H$, is an indicator of cilindricity of the tree, so that these equations would have greater capability to explain the variability of the volume, resulting in more precise estimates (Prodan *et al.*, 1997; Vallet *et al.*, 2006; Corral *et al.*, 2007; Neumann *et al.*, 2016). This variability of the volume is due to the shift of the cilindricity over the length of the stem, which also shifts as a function of several variables (Návar, 2010).

Although all compiled total volume equations presented linear structures, in this study the non-linear model proposed by Spurr (1952) was fitted. This is because of the curvature presented by the relationship between the observed volume and the combined variable $D^2H$. The Spurr (1952) model is frequently used in studies of volume and biomass equations (Chave *et al.*, 2005; Wutzler *et al.*, 2008; Vejpustková *et al.*, 2015).

Although the available local equations for roble, raulí and coigüe are based on models that use the $D$ and $H$ as independent variables, enabling greater accuracy using models whose only variable is $D$, the statistical information reported together with the compiled equations is scarce. In extreme cases, this information is restricted to regression parameters and sample size, making it difficult to

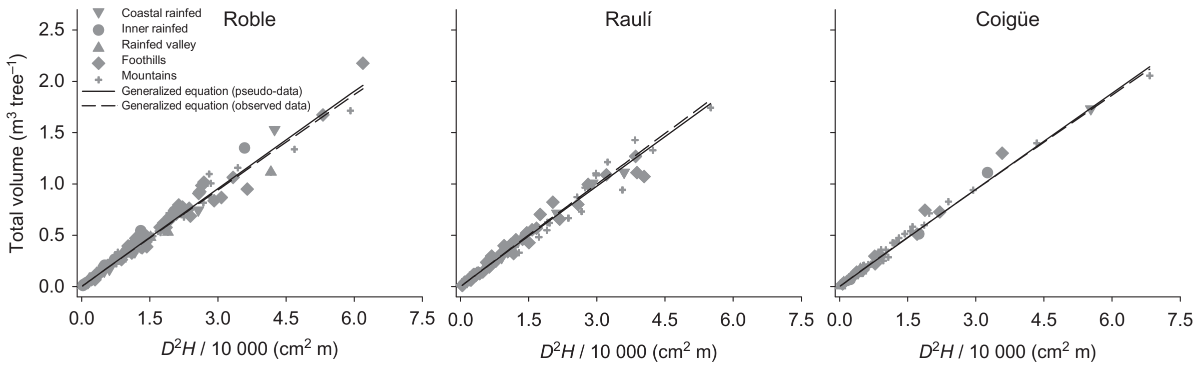


**Figure 6** Relationship between total volume ($m^3$) and $D^2H$ ($cm^2$ m) for roble, raulí and coigüe, in the sample used for validation. Symbols in grey represent observed data from different zones. Solid line represents the generalized equation fitted to pseudo-data generated with compiled equations. Dash line represents the generalized equation fitted to observed data.

generate independent pseudo-data from equations that include $H$ among predictors (Magnussen and Carillo, 2015). The obtained results in this study are similar to the results obtained by Wayson *et al.* (2015). It is probable that meta-analytical studies on allometric equations of volume or biomass will be frequent in the future. Thus, it should be noted that it would be of great help that authors of such studies should report the behaviour of the variance of the dependent variable and the height–diameter relationship of the sample used.

In the fitting of the allometric equations of total height and total volume using *dummy* variables, no significant zonal effects were detected on the parameters of the equations in any of the species under study (Table 4 and 7). One likely explanation is the sample size has larger variability than the zonal effect variability. Undoubtedly, the residual variability within each zone is relatively wide, due to a mixture of densities and stand ages. Still, the fitted equations in each zone are useful to estimate total height and generate pseudo-data for the fitting of total volume equations. Evidence not presented in the present document support the conclusion that obtaining generalized allometric equations of total volume are not depend on the sample size. The statistical tests carried out with small sample sizes (e.g. $n = 6$) and large sample sizes (e.g. $n = 10\,000$) in each zone detected no statistical zonal effects on the parameters of the fitted total volume model in any of the species. However, as the sample size increases the variability of the parameters reduces (Table 9).

## Conclusions

The generalized allometric equations of total volume underbark for second growth forests of roble, raulí and coigüe generated in this study using meta-analysis are valid for general use in the study zones. This technique, supplemented with data observed to model the height–diameter relationship and model the variance of the total volume allows for the generating non-autocorrelated random pseudo-data and with homogeneous variance. The allometric equations of total volume underbark obtained present high precision and accuracy.

It is likely that the use of modified meta-analysis will increase in the future, in studies that involve equations of volume or biomass. Therefore, it would be of great help that each author provides information on the behaviour of the variance of the dependent variable and the height–diameter relationship of the sample used in the fitting of this type of equations. Thus, the greatest advantage of this technique could be achieved, which in practice is an alternative to direct sampling. This technique even would allow for the elimination of direct sampling, with the exception of samples intended for validation.

## Acknowedgements

Authors also thank the Forest Company MASISA S.A. for granting access to farms of its assets in order to collect data. In addition, we would like to record our thanks to the Editor of the paper, Professor Mark Ducey, and two anonymous reviews of the paper, who all provided useful comments and suggestions that helped improve the paper.

## Conflict of interest statement

None declared.

## Funding

The authors thank the Corporación Nacional Forestal (CONAF) (Project Name: 025/2012 Desarrollo de herramientas de cuantificación biométrica generalizadas para el manejo y uso integral sustentable de renovales de *Nothofagus spp.*) III Concurso del Fondo de Investigación del Bosque Nativo for the financing of this research.

## References

Alegría, C. 2011 Modelling merchantable volumes for uneven aged maritime pine (*Pinus pinaster* Aiton) stands established by natural regeneration in the central Portugal. *Ann. For. Res.* **54** (2), 197–214.

António, N., Tomé, M., Tomé, J., Soares, P. and Fontes, L. 2007 Effect of tree, stand, and site variables on the allometry of *Eucalyptus globulus* tree biomass. *Can. J. For. Res.* **37** (5), 895–906.

Bergerud, W. 1994 GLM: Comparing regression lines. B.C. Biometrics Information. Ministry of Forests Research Program, British Columbia. Pamphlet 46.

Bruce, D. 1975 Evaluating accuracy of tree measurements made with optical instruments. *For. Sci.* **21** (4), 421–426.

Case, B. and Hall, R. 2008 Assessing prediction errors of generalized tree biomass and volume equations for the boreal forest region of west-central Canada. *Can. J. For. Res.* **38** (4), 878–889.

CONAF. 2011 Corporación Nacional Forestal, CL. Catastro de los recursos vegetacionales nativos de Chile. Sección Monitoreo de Ecosistemas Forestales. Santiago, Chile. 28 p.

Corral, J., Barrio, M., Aguirre, O. and Diéguez, U. 2007 Use of stump diameter to estimate diameter at breast height and tree volume for major pine species in El Salto, Durango (Mexico). *Forestry* **80** (1), 29–40.

Cubillos, V. 1988a Funciones de volumen y factor de forma para renovales de raulí. *Ciencia Investigación Forestal* **2** (3), 103–113.

Cubillos, V. 1988b Funciones de volumen y factor de forma para renovales de coigüe. *Ciencia Investigación Forestal* **2** (4), 62–68.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D. *et al.* 2005 Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia.* **145** (1), 87–99.

Chojnacky, D.C., Heath, L.S. and Jenkins, J.C. 2014 Updated generalized biomass equations for North American tree species. *Forestry* **87** (1), 129–151.

Donoso, P., Donoso, C. and Sandoval, V. 1993 Proposición de zonas de crecimiento de renovales de roble (*Nothofagus obliqua*) y raulí (*Nothofagus alpina*) en su rango de distribución natural. *Bosque* **14** (2), 37–55.

Drake, F., Emannuelli, P. and Acuña, E. 2003 Compendio de funciones dendrométricas del bosque nativo. Universidad de Concepción - Proyecto Conservación y Manejo Sustentable del Bosque Nativo CONAF-KFWDED - GTZ. Santiago, Chile. 197 p.

Fournier, R., Luther, J., Guindon, L., Lambert, M., Piercey, D., Hall, R. *et al.* 2003 Mapping aboveground tree biomass at the stand level from inventory information: test cases in Newfoundland and Quebec. *Can. J. For. Res.* **33** (10), 1846–1863.

Furnival, G.M. 1961 An index for comparing equations used in constructing volume tables. *For. Sci.* **7**, 337–341.

Gilabert, H. and Paci, C. 2010 An assessment of volume-ratio functions for *Eucalyptus globulus* and *E. nitens* in Chile. *Ciencia Investigación Agraria* **37**, 5–15.

Gonzalez-Benecke, C.A., Gezan, S.A., Martin, T.A., Cropper, W.P., Samuelson, L.J. and Leduc, D.J. 2014 Individual tree diameter, height, and volume functions for longleaf pine. *For. Sci.* **60** (1), 43–56.

Grosse, H.W. and Cubillos, V.D. 1991 Antecedentes generales para el manejo de renovales de raulí, roble, coigüe y tepa. Informe Informe Técnico N° 127. INFOR-CORFO. Santiago, Chile. 47 p.

Hardy, M. 1993 *Regression with Dummy Variables*. Sage. Sage university paper series on quantitative applications in the Social Sciences, 07-093.

Henry, M., Picard, N., Trotta, C., Manlay, R., Valentini, R., Bernoux, M. *et al.* 2011 Estimating tree biomass of Sub-Saharan african forests: a review of available allometric equations. *Silva Fenn.* **45** (3B), 477–569.

Jenkins, J., Chojnacky, D., Heath, L. and Birdsey, R. 2003 National-scale biomass estimators for United States tree species. *For. Sci.* **49** (1), 12–35.

Kmenta, J. 1986 *Elements of Econometrics*. 2nd edn. MacMillan Publishing Company.

Lambert, M., Ung, C. and Raulier, F. 2005 Canadian national tree aboveground biomass equations. *Can. J. For. Res.* **35** (8), 1996–2018.

Lara, A., Donoso, C., Donoso, P., Nuñez, P. and Cavieres, A. 1999 Normas de manejo para raleo de renovales del tipo forestal roble-raulí-coigüe. En: Silvicultura de los bosques nativos de Chile. 1ª Ed. Universitaria. Santiago. Chile. pp. 129–144.

Magnussen, S. and Carillo, O. 2015 Model errors in tree biomass estimates computed with an approximation to a missing covariance matrix. *Carbon Balance Manage.* **10** (1), 1–14.

Montagu, K.D., Düttmer, K., Barton, C.V.M. and Cowie, A.L. 2005 Developing general allometric relationships for regional estimates of carbon sequestration—an example using *Eucalyptus pilularis* from seven contrasting sites. *For. Ecol. Manage.* **204** (1), 115–129.

Muukkonen, P. 2007 Generalized allometric volume and biomass equations for some tree species in Europe. *Eur. J. For. Res.* **126** (2), 157–166.

Návar, J. 2009 Biomass component equations for Latin American species and groups of species. *Ann. For. Sci.* **66** (2), 1–21.

Návar, J. 2010 (Chapter 6) Measurement and assessment methods of forest aboveground biomass: a literature review and the challenges ahead. In: *Biomass*. Momba, M. and Bux, F. (eds). Sciyo, p. 202. ISBN: 978-953-307-113-B.

Neumann, M., Moreno, A., Mues, V., Härkönen, S., Mura, M., Bouriaud, O. *et al.* 2016 Comparison of carbon estimation methods for European forests. *For. Ecol. Manage.* **361**, 397–420.

ODEPA. 2000 Oficinas De Estudios y Políticas Agrarias. Clasificación de las explotaciones agrícolas del VI censo nacional agropecuario según tipo de productor y localización geográfica. Ministerio de Agricultura. Documento de trabajo N°5. I.S.S.N. 0717-0378. Santiago, Chile. 91 p.

Ott, P. 1997 The Use of indicator variables in non-linear regression. Biometrics Information. Ministry of Forests Research Program, British Columbia. Pamphlet 56.

Pastor, J., Aber, J. and Melillo, J. 1983/1984 Biomass prediction using generalized allometric regressions for some northeast tree species. *For. Ecol. Manage.* **7**, 265–274.

Peña, D. 1997 Combining information in statistical modeling. *Am. Stat.* **51**, 326–332.

Prodan, M., Peters, R., Cox, F. and Real, P. 1997 Mensura forestal. IICA/GTZ. 586 p.

Puente, E.M., Peñaloza, R.W., Donoso, C.Z., Paredes, R.O., Nuñez, P.M., Morales, R.A. *et al.* 1981 Estudio de raleo y otras técnicas para el manejo de renovales de Raulí y Roble. Instalación de ensayos de raleo. Investigación y Desarrollo Forestal: Docto. de Trabajo No 41. FO: DP/CHI/76 / 003. Santiago de Chile. 74 p.

SAS 2009 *SAS/STAT® 9.2 User's Guide*. 2d edn. SAS Institute Inc.

Spurr, S. 1952 *Forest Inventory*. Wiley. 472 p.

Stage, A. 1963 A mathematical approach to polymorphic site index curves for grand fir. *For. Sci.* **9** (2), 167–180.

Ung, C., Bernier, P. and Gou, X. 2008 Canadian national biomass equations: new parameter estimated that include British Columbia data. *Can. J. For. Res.* **38**, 1123–1132.

Vallet, P., Dhôte, J.-F., Moguédec, G.L., Ravart, M. and Pignard, G. 2006 Development of total aboveground volume equations for seven important forest tree species in France. *For. Ecol. Manag.* **229** (1–3), 98–110.

Vejpustková, M., Zahradník, D., Čihák, T. and Šrámek, V. 2015 Models for predicting aboveground biomass of European beech (*Fagus sylvatica* L.) in the Czech Republic. *J. For. Sci.* **61** (2), 45–54.

Wayson, C., Johnson, K., Cole, J., Olguín, M., Carrillo, O. and Birdsey, R. 2015 Estimating uncertainty of allometric biomass equations with incomplete fit error information using a pseudo-data approach: methods. *Ann. For. Sci.* **72**, 825–834.

Wirth, C., Schumacher, J. and Schulze, E.-D. 2004 Generic biomass functions for Norway spruce in Central Europe—a meta-analysis approach toward prediction and uncertainty estimation. *Tree Physiol.* **24** (2), 121–139.

Wutzler, T., Wirth, C. and Schumacher, J. 2008 Generic biomass functions for Common beech (*Fagus sylvatica*) in Central Europe: predictions and components of uncertainty. *Can. J. For. Res.* **38**, 1661–1675.

Zianis, D. and Mencuccini, M. 2003 Aboveground biomass relationships for beech (*Fagus moesiaca* Cz.) trees in Vermio Mountain, Northern Greece, and generalised equations for *Fagus* spp. *Ann. For. Sci.* **60**, 439–448.

Zianis, D., Muukkonen, P., Mäkipää, R. and Mencuccini, M. 2005 Biomass and stem volume equations for tree species in Europe. *Silva Fenn. Monogr.* **4**, 63.