ORIGINAL PAPER

Stem taper function for sweet chestnut (*Castanea sativa* Mill.) coppice stands in northwest Spain

**María Menéndez-Miguélez · Elena Canga ·
Pedro Álvarez-Álvarez · Juan Majada**

Abstract
• *Context* Despite the economic importance of *Castanea sativa* Mill. in northwest Spain, studies of its growth and yield are practically non-existent.
• *Aims* A compatible system formed by a taper function, a total volume equation, and a merchantable volume equation was developed for chestnut (*C. sativa* Mill.) coppice stands in northwest Spain.
• *Methods* Data from 203 destructively sampled trees were used for the adjustment. Outliers were removed with a non-parametric local adjustment, providing a final data set of measurements taken from 3,188 sections which was used to test five taper models (compatible and non-compatible). A second-order continuous autoregressive error structure was used to model the error term and account for autocorrelation. Presence of multicollinearity was evaluated with the condition number. Comparison of the models was carried out using overall goodness-of-fit statistics and graphical analysis.
• *Results* Results show that the models developed by Fang et al. in For Sci 46: 1–12, 2000 and Kozak in For Chron 80, N

4: 507–515, 2004 were superior to other equations in predicting diameter for chestnut coppice stands.
• *Conclusion* The compatible volume system developed by Fang et al. in For Sci 46: 1–12, 2000 was finally selected as it provided the best compromise between describing stem profile and also estimating merchantable height, merchantable volume, and total volume and therefore provides the first specific tool for more effective management of chestnut coppice stands.

1 Introduction

Sweet chestnut (*Castanea sativa* Mill.) covers more than 2.5 million hectares in Europe, with a distribution reaching from the Southern Mediterranean to central, Atlantic, and Eastern Europe (Conedera et al. 2004). Chestnut forests have been recognized as habitats of interest in the European Natura 2000 network and are considered characteristic cultural landscapes of the Mediterranean and Atlantic regions (Díaz Varela et al. 2009). In northwest Spain, chestnut is the most important forest species, covering over 100,000 ha, mainly as coppice stands (DGCONA 2013). This area accounts for over 95 % of the potential area for chestnut coppice stands in Spain.

Although chestnut fruit production has traditionally driven management in the region, changes in markets and local economies have resulted in timber production becoming the main objective in most exploitation nowadays (Álvarez et al. 2000). The vitality of the chestnut root system, with stools capable of sustainably producing an abundance of shoots, and high productivity (8–16 m^3 ha^{-1} year^{-1} depending on site conditions) facilitate management under a coppice system (Giudici et al. 2000). Chestnut coppice produces valuable

M. Menéndez-Miguélez (✉) · E. Canga · J. Majada
Forest and Wood Technology Research Centre (CETEMAS), Finca Experimental "La Mata" s/n, Grado, Asturias, Spain
e-mail: mmenendez@cetemas.es
URL: http://www.cetemas.es

P. Álvarez-Álvarez
Research Group in Atlantic Forests (GIS-Forest), Department of Organisms and Systems Biology, Escuela Politécnica de Mieres (EPM), University of Oviedo, Campus Universitario de Mieres, C/Gonzalo Gutiérrez Quirós, 33600 Mieres, Spain

timber in relatively short rotations (20–40 years) compared to other hardwoods (Gallardo et al. 2000; Kerr and Evans 1993). The total volume (with bark) of sweet chestnut stands (high forest and coppice stands together) harvested in Spain during 2011 was 58,090 m^3 (MARM 2011), with more than 42.46 % of this total volume being formed by trees from coppice stands in northwest Spain.

Estimating timber volume stocks as accurately as possible is essential in forest management. It is therefore necessary to develop tools that allow the reliable estimation of tree volume using variables which are easy to measure in the field, such as diameter at breast height (D) and total height (H). One such tool is individual tree volume equations. However, these equations have the disadvantage of not being able to predict tree volume for wood products which are classified by merchantable size depending on log dimensions.

There are a number of ways to address this issue, the two most important of which are developing volume-ratio equations that predict merchantable volume as a percentage of total volume (Burkhart 1977; Clutter 1980; Reed and Green 1984) or using taper functions.

Taper functions describe stem taper (Brink and Gadow 1986; Kozak 1988; Riemer et al. 1995) and provide forest managers with estimates of (a) diameter at any point along the stem, (b) total stem volume, (c) merchantable volume and merchantable height to any top diameter and from any stump height, and (d) individual volumes for logs of any length at any height above the ground (Kozak 2004). Such functions can be implemented in different computer software specially developed for this type of

calculation, such as GesMO (Diéguez-Aranda et al. 2009) or CubiFOR (Rodríguez et al. 2008). To develop this type of function, it is necessary to have a longitudinal data structure, that is, multiple measurements for each individual (Lindstrom and Bates 1990).

Ideally, a taper equation should be compatible, meaning that the volume computed by integration of the taper function should be equal to that calculated by a total volume equation (Clutter 1980; Demaerschalk 1972; Fang et al. 2000). Examples of compatible volume-estimating systems are the works carried out by Demaerschalk (1972), Goulding and Murray (1976), and Fang et al. (2000).

Prediction tools are essential to understand the development of forest stands and subsequently decide on the best management strategy. In Spain, many taper functions have been developed for different forest species (Barrio-Anta et al. 2007; Crecente-Campo et al. 2009; Diéguez-Aranda et al. 2006); however, there is currently no taper function available for chestnut coppice, either in Spain or elsewhere in the world. This work is a result of looking to remedy this gap in provision, and its main objective is to develop a taper function able to correctly describe the profile of and ensure appropriate estimates of stem volume using chestnut coppice stands in northwest Spain as a baseline. Specifically, we wish to focus on two questions: (a) Is it possible to correctly describe the huge variability of stem profiles in chestnut coppice stands given the high number of stems which may grow from a single stool, and (b) which model best describes this type of profile and its variability?

Fig. 1 Map showing cover rates for chestnut coppice stands in the study area. Fitting plots are indicated by *red dots*
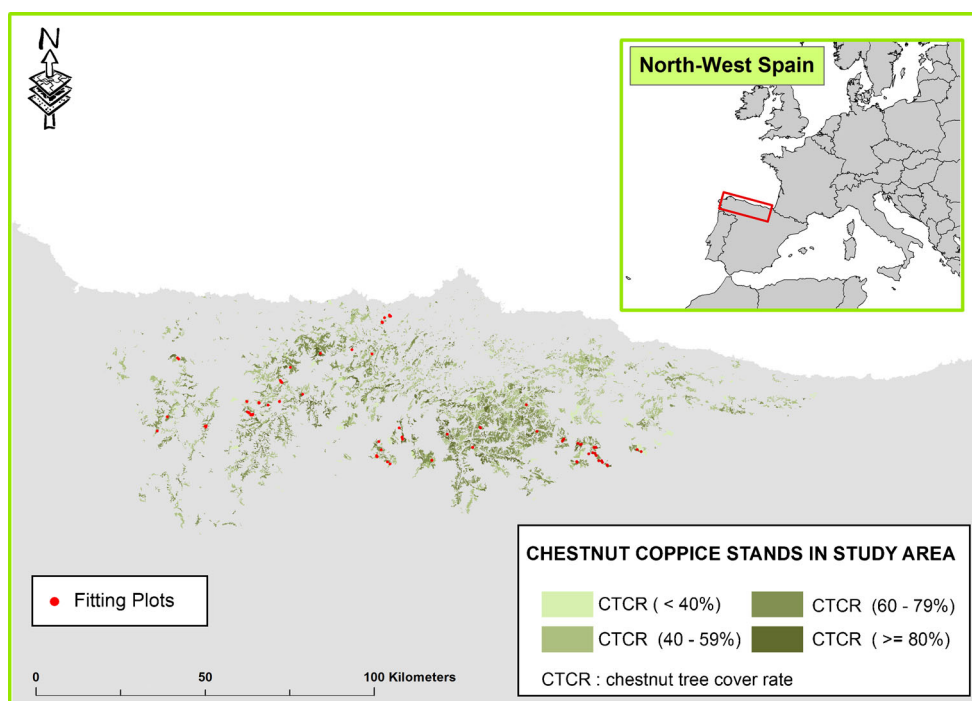
Table 1 Descriptive statistics of tree and stand data sets used in the analysis for fitting and validation

	Fitting					Validation				
	n	Mean	Min	Max	Std. dev.	n	Mean	Min	Max	Std. dev.
Individual tree										
Nº logs	3,188	16.30	6	25	3.89	719	11.83	3	20	4.07
D	3,188	25.69	10.20	44.50	7.94	719	19.65	6.90	49.15	8.12
H	3,188	19.26	9.54	31.02	3.37	719	16.68	8.35	26.08	3.84
n_{tree_a}	3,188	1.87	1	8	1.39	719	2.43	1	7	1.33
h_{stump}	3,188	0.12	0	0.60	0.10	719	0.14	0.020	0.68	0.14
V	3,188	0.35	0.011	1.85	0.30	719	0.19	0.020	1.01	0.19
Stand										
n_{ree_b}	63	2.19	1	26	2.25	30	2.73	1	24	2.42
N	63	1,692.52	410.26	4,753.42	942.07	30	2,025.51	222.82	8,244.23	1,762.97
G	63	40.01	16.33	65.33	11.30	30	10.29	3.41	35.40	6.79
H_m	63	16.93	10.63	24.16	3.09	30	–	–	–	–
H_0	63	19.91	12.15	28.17	3.09	30	16.52	10.75	23.37	3.27

D diameter at breast height (cm), H height (m), n_{tree_a} number of stems from the stool where the sampled tree was cut, h_{stump} stump height (m), V total over bark stem volume (m^3), n_{tree_b} number of stems in the stool, N number of stems per hectare (stems ha^{-1}), G basal area (m^2 ha^{-1}), H_m mean height (m), H_0 average height of the 100 thickest trees per hectare (m)

2 Material and methods

2.1 Data

The data used in this study were collected in 70 coppice stands covering the existing range of ages, stand densities, and sites of this species in the region. Figure 1 shows the map with the locations of the stands used for the fitting data.

A total of 203 trees were felled and destructively sampled. Trees had to be healthy and of a standard shape (i.e., not forked nor excessively branched) and were selected in order to ensure a representative distribution of diameter and height classes (Table 1).

Before felling, diameter at breast height D (diameter at 1.3 m above the top of the stool, in cm) was measured to the nearest 0.1 cm for each tree. The trees were then felled and total bole length, that is, total height H, (in m) measured to the nearest 0.1 m. The trees were cut into 1-m logs, up to a top diameter of 7 cm, and measured to the nearest centimeter. Two perpendicular over bark diameters (d, cm) and two perpendicular bark thicknesses were measured to the nearest 0.1 cm in each cross section (at height h, in m, above the top of the stool). Over bark log volumes were calculated in cubic meters using Smalian's formula, and the top section was treated as a cone. Over bark total stem volume was obtained by summing the over bark log volumes and the volume of the top section. Finally, 3,282 pairs of diameter (d) at a certain height (h) measurements were used for the original fitting data set.

Data from an independent network of plots (established by the Atlantic Forest Systems Research Group (GIS-Forest), Department of Organisms and Systems Biology, University of Oviedo) was used for validation purposes. The height-diameter distributions for the fitting and validation samples are very similar (Fig. 2), indicating that robust conclusions can be reached from the validation analysis.

The scatterplot of relative diameter (d/D) against relative height (h/H) was examined visually to detect possible anomalies in the data. This first analysis detected a number of outliers (many of them corresponded to trees with abnormalities) which were removed. A second analysis was carried out with the systematic procedure proposed by Bi (2000) to detect and remove other possible outliers, whereby local adjustment
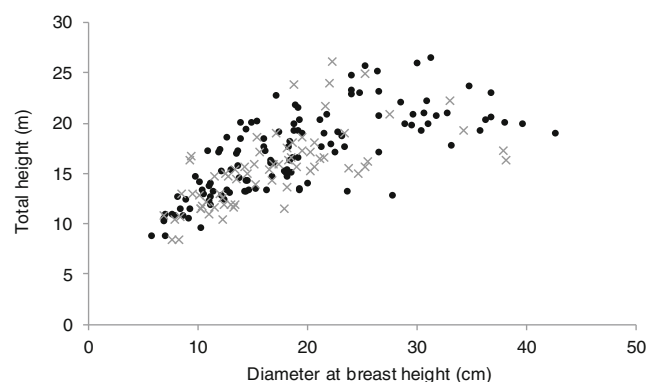

Fig. 2 Plot of diameter at breast height against total height of sampled trees (*black dot*) and validation sample (*multiplication sign*)
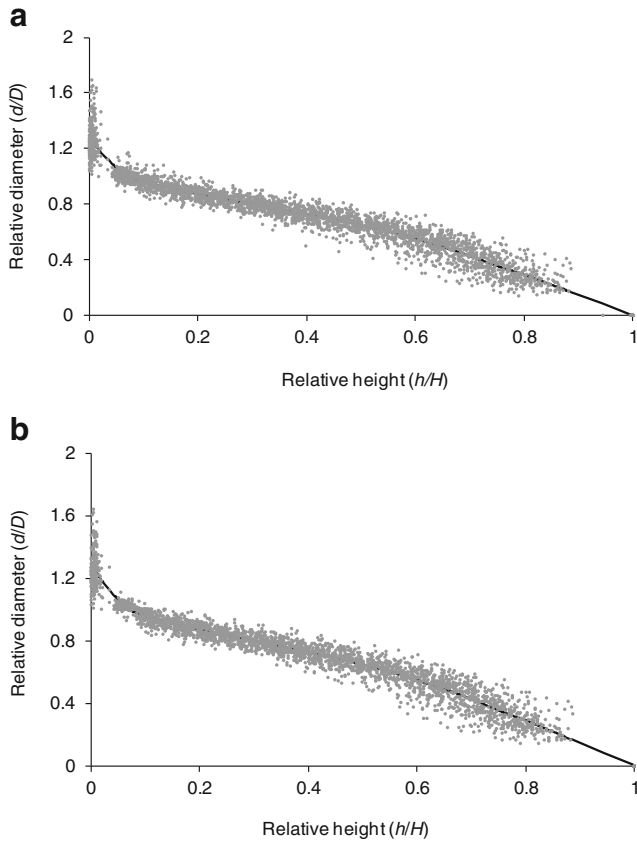
Fig. 3 Data points of relative diameter and relative height plotted with a local regression LOESS smoothing curve (smoothing factor=0.25) for all data (**a**) and after the elimination of outliers (**b**)

was performed by the LOESS procedure of SAS/STAT® (SAS Institute Inc. 2004a) with a smoothing factor of 0.25.

Using this approach, the number of extreme values accounted for 2.83 % of total taper measurements. A small percentage of the extreme data points were the result of errors in measuring bole sections or in the transcription of field notes, but most were the result of measurements in sections where the tree was deformed due to abnormal growth or damage caused by the presence of chancre (*Cryphonectria parasitica* (Murr.) Barr.). Since taper functions are not intended for deformed stems, these data points were excluded from further analysis, resulting in a final total number of observations of 3,188, from 190 trees.

Figure 3a, b shows relative height against relative diameter together with the LOESS regression curve, the upper graphic showing all the collected data and that below, the data excluding outliers, respectively. Summary statistics of the final data used in this study for tree and stand variables, together with model validation data, are shown in Table 1.

2.2 Equations tested

We analyzed a total of five models, which are described below and whose expressions are shown in Table 2:

– Fang et al. (2000). Compatible system formed by a taper function, a total volume equation, and a merchantable volume equation. The taper equation is segmented with two attachment points and three form factors, one for each segment.
– Bi (2000). Non-compatible variable-exponent taper function.

Table 2 Fitted taper equations and their corresponding mathematical expression

Model	Expression
Fang et al. (2000)	$d = c_1 \sqrt{H^{(k-b_1)/b_1}(1-q)^{(k-\beta)/\beta} \alpha_1^{I_1+I_2} \alpha_2^{I_2}}$ $c_1 = \sqrt{\dfrac{a_0 D^{a_1} H^{a_2-k/b_1}}{b_1(r_0-r_1)+b_2(r_1-\alpha_1 r_2)+b_3\alpha_1 r_2}}$ $r_0 = (1-h_{\text{stump}}/H)^{k/b_1}$ $\qquad$ $r_1 = (1-p_1)^{k/b_1}$ $\qquad$ $r_2 = (1-p_2)^{k/b_2}$ $\beta = b_1^{1-(I_1+I_2)} b_2^{I_1} b_3^{I_2}$ $\qquad$ $\alpha_1 = (1-p_1)^{\frac{(b_2-b_1)k}{b_1 b_2}}$ $\qquad$ $\alpha_2 = (1-p_2)^{\frac{(b_3-b_2)k}{b_2 b_3}}$ $I_1=1$ if $p_1 \leq q \leq p_2$, 0 in all other cases $I_2=1$ if $p_2 < q \leq 1$, 0 in all other cases $p_1 = h_1/H$ y $p_2 = h_2/H$
Bi (2000)	$d = D\left[\dfrac{\ln\sin\left(\frac{\pi}{2}q\right)}{\ln\sin\left(\frac{1.3\pi}{2H}\right)}\right]^{a_1+a_2\sin\left(\frac{\pi}{2}q\right)+a_3\cos\left(\frac{3\pi}{2}q\right)+\frac{a_4\sin\left(\frac{\pi}{2}q\right)}{q}+a_5 D+a_6 q\sqrt{D}+a_7 q\sqrt{H}}$
Kozak (2004)	$d = a_0 D^{a_1} H^{a_2} X^{b_1 q^4 + b_2\left(1/e^{D/H}\right) + b_3 x^{0.1} + b_5 H^w + b_6 x}$
Demaerschalk (1972)	$d_i = b_0 d^{b_1}(h-h_i)^{b_2} h^{b_3}$
Thomas and Parresol (1991)	$\left(\dfrac{d_i}{d}\right)^2 = b_1(q-1) + b_2\sin(b_4\pi q) + b_3\cot\left(\frac{\pi q}{2}\right)$

D over bark diameter at breast height (at 1.30 m above the top of the stool, cm), *d* over bark diameter at height *h* (cm), *H* total tree height (m), *h* height from top of the stool to top diameter *d* (m), h_{stump} stump height (m), *V* total over bark stem volume (m³), a_0, ..., a_2, ..., b_0, ..., b_6, p_1, p_2 parameters to be estimated, *k* equal to $\pi/40000$, *q* is equal to h/H

- Kozak (2004). Non-compatible variable-exponent taper function.
- Demaerschalk (1972). Power function whose main advantage is that the volume equations obtained by integrating are algebraically compatible with classic taper functions.
- Thomas and Parresol (1991). Trigonometric compatible model.

2.3 Model fitting and selection

The models tested were fitted by non-linear regression with the MODEL procedure of SAS/ETS® (SAS Institute Inc. 2004b) using generalized least squares for non-linear models.

Of the different options to estimate the parameters in the systems where the taper equation includes a total volume equation (Fang and Bailey 1999; Fang et al. 2000; Goulding and Murray 1976), in this study we prioritized the taper function, setting this first and subsequently performing the predicted volume calculation from the estimation parameters obtained.

To avoid problems in the estimation of the parameters, a value of 0.001 was assigned to the final diameter of the top section. Similarly, a value of 0.001 was also subtracted from the heights equal to the total height, that is $h = H - 0.001$; values which are lower than the appreciation limit are used in the data collection. This approach allows the use of the entire data set for fitting and does not significantly change parameter estimates (Diéguez-Aranda et al. 2006).

There are several problems associated with stem taper and volume equation analyses that violate the fundamental least squares assumption of independence and equal distribution of errors with zero mean and constant variance. One of the most common is the presence of autocorrelation in the data as a result of working with multiple observations on each tree. To resolve this problem, the error term was modeled using a continuous autoregressive error structure (CAR(x)), which allows the model to be applied to irregularly spaced, unbalanced data (Zimmerman and Nuñez-Antón 2001).

Another problem in taper functions is multicollinearity, which refers to the existence of high intercorrelations among the independent variables in multiple linear or non-linear regression analyses. To evaluate the presence of multicollinearity, we used the condition number (CN). According to Belsey (1991), if the condition number is between 5 and 10, collinearity is not a major problem; if it is in the range of 30–100, then there are problems associated with collinearity; and if it is in the range of 1,000–3,000, the problems are severe.

The criteria used for the comparison of the models were based on the residual plot analysis and statistical analysis of the goodness-of-fit statistics: adjusted coefficient of determination (R^2_{adj}), root mean square error (RMSE), and Akaike's information criterion in differences (AICd).

Table 4 Goodness-of-fit statistics and condition number of the taper functions evaluated

Model	RMSE	AICd	R^2_{adj}	CN
Fang et al. (2000)	1.188	0	0.9838	62.40
Kozak (2004)	1.223	184.24	0.9828	62.57
Demaerschalk (1972)	1.537	1638.71	0.9728	62.81
Thomas and Parresol (1991)	2.055	3490.69	0.9515	3.74

RMSE root mean square error, AICd Akaike's information criterion in differences, R^2_{adj} adjusted coefficient of determination, CN condition number

Table 3 Parameter estimates (approximated standard error in parentheses) for the models analyzed

| Parameter	Fang et al. (2000)	Bi (2000)	Kozak (2004)	Demaerschalk (1972)	Thomas and Parresol (1991)
a_0	5.542×10^{-5} (3.7×10^{-6})		0.8600 (0.029)		
a_1	1.914 (0.015)		0.9781 (0.0084)		
a_2	0.936 (0.027)		0.08444 (0.015)		
b_0				1.520 (0.061)	
b_1	9.869×10^{-6} (1.61×10^{-7})	Convergence was not achieved	0.7844 (0.029)	0.9567 (0.0098)	−1.00228 (0.0065)
b_2	3.362×10^{-5} (2.99×10^{-7})		−0.3341 (0.040)	0.8898 (0.0070)	0.1107 (0.0036)
b_3	2.667×10^{-5} (3.79×10^{-7})		0.4866 (0.016)	−0.9378 (0.019)	3.851×10^{-7} (2.93×10^{-8})
b_5			−0.01954 (0.0021)		
b_6			0.2318 (0.026)		
p_1	0.07191 (0.0014)			0.7258 (0.025)	0.8367 (0.024)
p_2	0.5590 (0.0101)			0.3936 (0.020)	0.3639 (0.019)
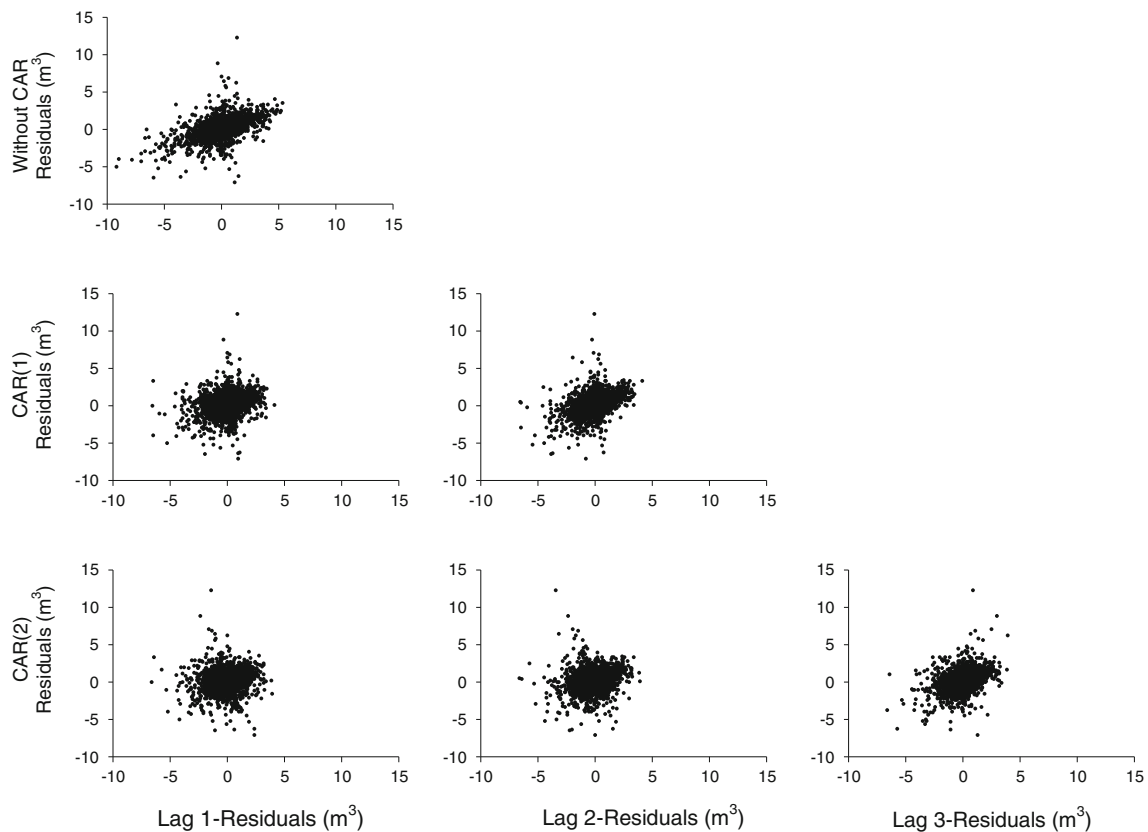
Fig. 4 Residuals against: Lag1-residuals (*left column*), Lag2-residuals (*middle column*), and Lag3-residuals (*right column*) for the model of Fang et al. (2000) fitted without considering the autocorrelation parameters (*first row*) and using continuous time autoregressive error structures of first and second order (*second and third rows*, respectively)

Although the goodness-of-fit statistics reflect the behavior of the data for the different models evaluated, they may not indicate which model is the best for practical purposes (Diéguez-Aranda et al. 2006); hence, this decision should be made after analyzing each model's behavior according to the different stem sections. To evaluate this, the bias and the root mean square error were calculated and plotted for diameter estimation by relative height classes (intervals of 15 %) and for height estimation by diameter classes (intervals of 5 cm). To estimate the height at which the different diameters are achieved, the iterative bisection method was used.

2.4 Model validation

Quality of fit does not necessarily reflect the quality of future prediction (Myers 1990). Only validation with an independent data set enables the accuracy of the selected model to be known (Huang et al. 2003; Kozak and Kozak 2003). In this study, the validation process was carried out with an independent data set consisting of 70 trees (from a network of plots established by the Atlantic Forest Systems Research Group (GIS-Forest), Department of Organism and Systems Biology, University of Oviedo), which produced a total of 719 height/diameter data pairs. Trees

Fig. 5 Bias (**a**) and root mean square error (RMSE) evolution (**b**) for diameter (cm) by relative height class (percentages) for the compatible system of Fang et al. (2000) and the non-compatible taper function of Kozak (2004)
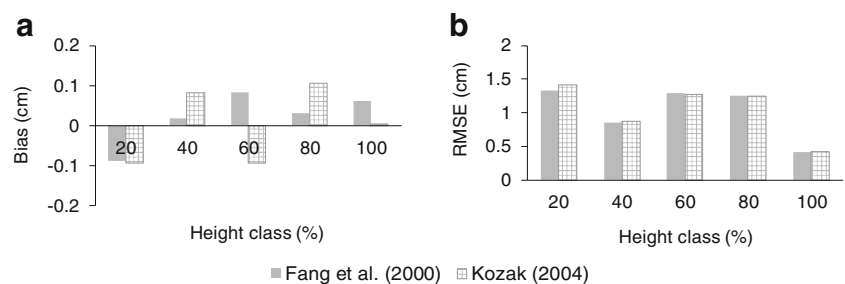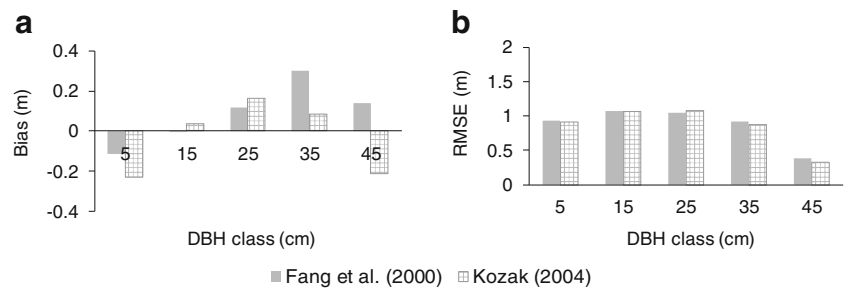
Fig. 6 Bias (a) and root mean square error (RMSE) evolution (b) for height (m) by relative diameter class (DBH class) for the compatible system of Fang et al. (2000) and the non-compatible taper function of Kozak (2004)

were felled and destructively sampled following the same methodology as used for the fitting data set. Two validation statistics were calculated to assess the overall prediction performance of the fitted equations on this validation data set: (a) an estimate of the average prediction error (APE) (Eq. 1) (Weisberg 1985) and (b) mean bias (Eq. 2) estimated as an overall average and summarized by diameter class, similar to that used by Zhang (1997). Both statistics present errors in the same units as the variable used, in this case centimeter for diameters and cubic meter for volumes. The APE statistic in the validation process is similar to the RMSE in the fitting.

$$\mathrm{APE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{n}} \qquad (1)$$

$$\overline{\mathrm{Bias}} = \frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)}{n} \qquad (2)$$

where Y_i is the observed or real value, $\widehat{Y}_i$ is the estimated value with the model, and n is the sample size of the validation data.
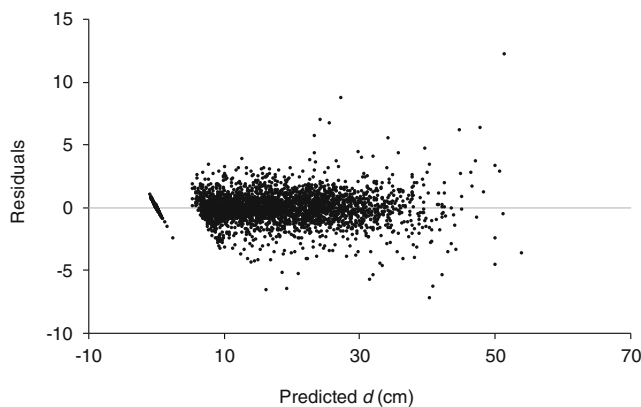
To examine the performance of the models in greater detail, the values of $\overline{\mathrm{Bias}}$ were plotted against diameter and total volume. These graphs are of interest since they illustrate areas in which the adjusted models provide poor or good predictions according to the diameter class of the evaluated trees.

3 Results

Table 3 shows the parameters for the taper functions fitted, all of which were significant at the 5 % level, except for the Bi (2000) model, where convergence was not achieved. The model of Kozak (2004) was modified by removing the b_4 parameter in order to adapt it to local and species conditions (Kozak 2004).

All models performed well, each explaining more than 95 % of the total variability, with mean error below 2.05 cm (Table 4). Comparison of goodness-of-fit statistics indicates

Fig. 7 Plot of residuals against predicted diameter from the taper function proposed by Fang et al. (2000)
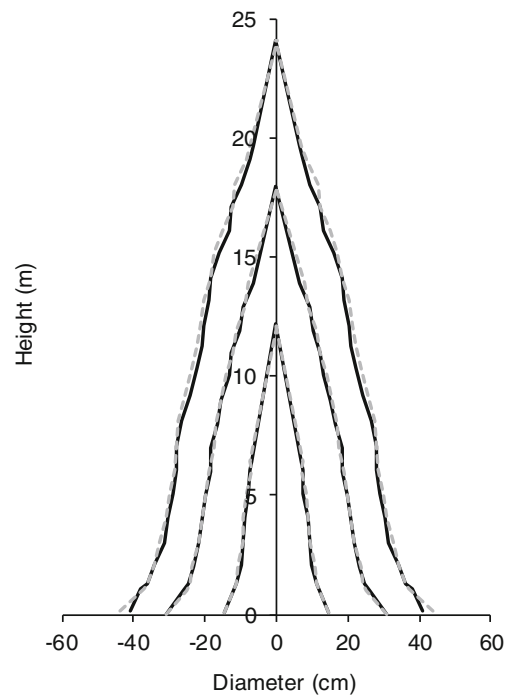

Fig. 8 Observed (solid line) and predicted (dashed line) profiles of three trees (as examples) using the taper function of Fang et al. (2000)

that the best-fitting models are those of Kozak (2004) and Fang et al. (2000), which each explaining more than 98 % of the total variability. In both cases, the presence of multicollinearity was observed (CN around 62) but it was considered to be within acceptable limits.

A trend in the residuals depending on the distance and the relative position of the measurement along the stem was found in the model fitting. Therefore, autocorrelation was corrected applying a second-order autoregressive structure (because using a first-order structure proved to be insufficient) with the aim of obtaining unbiased and efficient estimates, which did not invalidate statistical tests. Following this correction, the trends in residuals virtually disappeared. Figure 4 provides an example using the model of Fang et al. (2000).

Statistics are good indicators of the global performance of the taper function, but alone, they do not allow the best model to be selected. To do this, the evolution of bias and mean square root error in diameter estimation by relative height classes at intervals of 20 % (Fig. 5a, b) and in height estimation by diameter class (Fig. 6a, b) was analyzed for the two best-fit models, Fang et al. (2000) and Kozak (2004).

Graphical analysis of the bias in predicting diameters (Fig. 5a) confirmed the good performance of both models (with bias under ±0.1), with a certain advantage seen for the model of Fang et al. (2000), which showed lower bias at different heights, especially in the lower part of the stem (that with the highest merchantable value). In relation to the evolution of RMSE in predicting diameters (Fig. 5b), both models were very similar, although the model of Fang et al. (2000) was slightly better.

With regard to the evaluation of bias in predicting heights (Fig. 6a), the model of Fang et al. (2000) showed lower bias until diameter class 35, at which point the model of Kozak (2004) performed better, although the model of Fang et al. (2000) again showed the best fit at class 45. For both models, however, there was bias, although up to and including diameter class 25 it was always below 0.3 cm, and for the classes
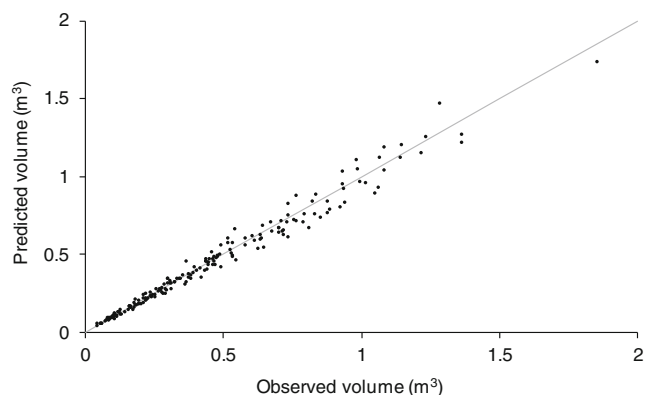
Table 5 APE and $\overline{\text{Bias}}$ statistics generated from the assessment of prediction error for the taper function fitted in the validation process

	APE	$\overline{\text{Bias}}$
Diameter (d)	2.14	−0.41
Total volume (V)	0.059	−0.019

APE average prediction error, $\overline{\text{Bias}}$ mean bias

above this, always less than 0.4 cm. The behavior of both models in terms of RMSE was very similar (Fig. 6b).

Taking into account the results and in particular the practical utility of the compatibility between the classic two inputs volume equation and the taper function, the model of Fang et al. (2000) was selected as the most appropriate for chestnut coppice stands in northwest Spain.

The plotting of values from predicted diameter in the selected taper function against the residuals is shown in Fig. 7, where no systematic trend in the distribution of residuals was observed. Figure 8 shows, as an example, the profile of three trees—one small (d=11.15 cm and h=12.19 m), one medium sized (d=24.5 cm and h=18 m), and one large (d= 36 cm and h=24.12 m) generated from the observed values (solid lines) and predicted values (dashed lines) for the model of Fang et al. (2000). Figure 9 shows predicted values of total volume for the selected taper function against the observed volume values, verifying the accuracy of the estimates (accounting for 98.38 % of the total variability).

3.1 Model validation

Table 5 shows the statistics used in model validation, calculated for different diameter classes. APE generally increased with diameter class in the trees evaluated for the variable diameter and volume and provided good results (average prediction error of 2.14 cm for diameter and 0.059 m^3 for volume).

The graphs of mean prediction bias are shown in Fig. 10. All values obtained, in the case of both diameter and volume, were similar and close to zero, indicating that the selected equation fits well with the real profile of the tree. Up to diameter class 25, the statistics were very close to zero, although, it is important to note that from diameter class 35, the $\overline{\text{Bias}}$ values in predicting diameter were far from zero. $\overline{\text{Bias}}$ values indicate that both diameter and total tree volume are overestimated (negative values).

4 Discussion

Currently, detailed information is available as regards the different functions and methodologies for the correct estimation of diameters at different heights and total or merchantable

Fig. 9 Plot of predicted values against observed values for total tree volume from the taper function proposed by Fang et al. (2000)

Fig. 10 Plot of DBH class against mean prediction bias for diameter (*left*) and total volume (*right*)
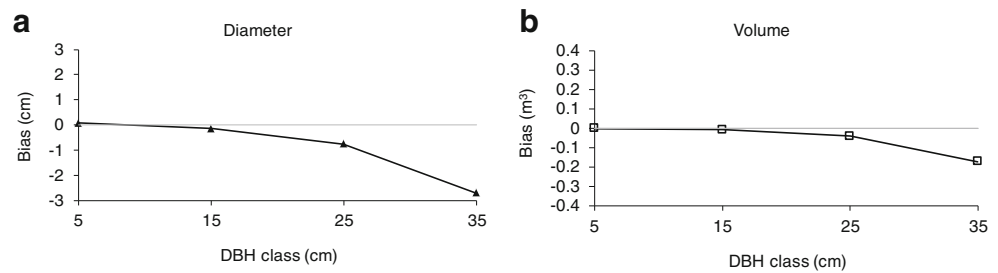
stem volume for different species (e.g., Barrio et al. 2007; Diéguez-Aranda et al. 2006). However, no such tools are yet available for chestnut, neither for high forest nor for coppice stands, hence the relevance of this work, which facilitates a better understanding and management of the species.

The final selected model explained more than 98.4 % of total variability and had mean errors below 1.20 cm. The estimates obtained in the models analyzed were similar to those obtained for other species. The model of Fang et al. (2000) has shown good performance, as much for broadleaf species as for conifers (e.g., Barrio-Anta et al. 2007; Diéguez-Aranda et al. 2006; Pompa-García et al. 2009).

Significant variability in chestnut stem profiles occurs in this study due to the high number of stems (up to eight) which were growing from each stool. Previous studies (e.g., Muhairwe 1994) have already demonstrated that factors such as site index, size and position of the crown, and stand density affect the profile of the tree. Modeling the profile of chestnut, in particular in coppice stands, presents an additional difficulty. Due to the fact that often many stems come from the same stool, it seems logical that stool density (number of stems per stool) as well as stand density might also be a key factor because internal competition affects the profile of the tree. Despite this, the selected model explained over 98 % of total variability, above the values obtained in previous broadleaf studies (Barrio-Anta et al. 2007; Pompa-García et al. 2009). Moreover, as the bias values in predicting diameters show, the results perform well in relation to the basal part of the tree, thereby solving one of the main problems associated with the use of taper functions in trees with prominent basal zones.

Validation with an independent data set confirmed the applicability of the selected taper function and the compatible volume equation for chestnut coppice stands in northwest Spain. Both statistics, APE and $\overline{\text{Bias}}$, increased with diameter class in the trees evaluated. $\overline{\text{Bias}}$ values did not vary greatly until diameter class 35, after which range slightly increased. This can be attributed in part to a relatively lower number of sampled trees in this diameter class, that is, 4 trees from a total of 70 in the whole validation data set.

5 Conclusions

A taper function for chestnut coppice stands in northwest Spain was developed to estimate diameter at any point along the stem, along with a total volume equation compatible with the fitted taper function. A total of five models were evaluated: the segmented model of Fang et al. (2000), the variable exponent functions proposed by Bi (2000) and Kozak (2004), the power function proposed by Demaerschalk (1972), and the trigonometric compatible model proposed by Thomas and Parresol (1991). In the end, the Bi (2000) model was not compared to the other models because convergence was not achieved in this case. All the other functions analyzed had good performance in estimating diameter along the stem, all of them appropriately describing the stem profile for chestnut coppice stands.

The compatible system to estimate volume proposed by Fang et al. (2000) was finally selected as the best taper function to explain the profile of chestnut coppice, as much for its goodness-of-fit statistics (R^2_{adj} of 0.98 and mean error of 1.19 cm) as for its prediction ability for diameter and height along the stem. This system has the advantage of being formed by a taper function, a total volume equation, and a merchantable volume equation, all of which are compatible between themselves.

Validation using an independent data set reflected the quality of predictions and confirmed the ability of the selected taper function to describe the stem profile in chestnut coppice stands in northwest Spain.

The taper function finally selected could be used for coppice stands in the rest of the country or elsewhere in the first instance, until new adjusted taper functions are developed to ensure the most accurate estimations possible for specific areas.

References

Álvarez P, Barrio M, Castedo F, Díaz R, Fernández JL, Mansilla P, Pérez R, Pintos C, Riesgo G, Rodríguez R, Salinero MC (2000) Manual de selvicultura del castaño en Galicia. University of Santiago Press, Santiago, España

Barrio-Anta M, Diéguez-Aranda U, Castedo-Dorado F, Álvarez-González JG, Von Gadow K (2007) Merchantable volume system for pedunculate oak in northwestern Spain. Ann For Sci 64:511–520. doi:10.1051/forest:2007028

Belsey DA (1991) Conditioning diagnostics, collinearity and weak data in regression. Wiley, New York

Bi H (2000) Trigonometric variable-form taper equations for Australian eucalypts. For Sci 46:397–409

Brink C, Gadow KV (1986) On the use of growth and decay functions for modeling stem profiles. EDV in Medizin u Biologie 17:20–27. doi:10.1080/02827580410019562

Burkhart H (1977) Cubic-foot volume of loblolly pine to any merchantable top limit. Sout J Appl For 1:7–9

Clutter JL (1980) Development of taper functions from variable-top merchantable volume equations. For Sci 26:117–120

Conedera M, Manetti MC, Giudici F, Amorini E (2004) Distribution and economic potential of the sweet chestnut (*Castanea sativa* Mill.) in Europe. Ecologia Mediterranea 30:179–193

Crecente-Campo F, Rojo A, Diéguez-Aranda U (2009) A merchantable volume system for *Pinus sylvestris* L. in the major mountain ranges of Spain. Ann For Sci 66:808. doi:10.1051/forest/2009078

Demaerschalk J (1972) Converting volume equations to compatible taper equations. For Sci 18:241–245

DGCONA (2013). III Mapa Forestal de España. MFE50. 1:50000. Ministerio de Medio Ambiente, Madrid, España

Díaz Varela RA, Calvo Iglesias MS, Díaz Varela ER, Ramil Rego P, Crecente Maseda R (2009) *Castanea sativa* forests: a threatened cultural landscape in Galicia NW Spain. In: O'Connell M, Küster H (eds) Cultural Landscapes of Europe. Fields of Demeter Haunts of Pan (Krzywinski K. Aschembeck Media UG, Bremen, pp 94–95

Diéguez-Aranda U, Castedo-Dorado F, Álvarez-González JG, Rojo A (2006) Compatible taper function for Scots pine plantations in northwestern Spain. Can J For Res-Rev Can Rech For 36:1190–1205. doi:10.1139/X06-008

Diéguez-Aranda U, Rojo Alboreca A, Castedo-Dorado F, Álvarez-González JG, Barrio-Anta M, Crecente-Campo F, González-González JM, Pérez-Cruzado C, Rodríguez-Soalleiro R, López-Sánchez CA, Balboa-Murias MA, Gorgoso Varela JJ, Sánchez Rodríguez F (2009). Consellería do Medio Rural, Xunta de Galicia. 268 pp+CD-Rom.

Fang Z, Bailey RL (1999) Compatible volume and taper models with coefficients for tropical species on Hainan Island in Southern China. For Sci 45:85–100

Fang Z, Borders BE, Bailey RL (2000) Compatible volume-taper models for loblolly and slash pine based on a system with segmented-stem form factors. For Sci 46:1–12

Gallardo JF, Rico M, González MI (2000) Some ecological aspects of a chestnut coppice located at the Sierra de Gata mountains (Western Spain) and its relationships with a sustainable management. Ecologia Mediterranea 26:53–69

Giudici F, Amorini E, Manetti MC, Chatziphilippidis G, Pividori M, Sevrin E, Zingg A (2000) Sustainable management of sweet chestnut (*Castanea sativa* Mill.) coppice forest by means of the production of quality timber. Ecologia Mediterranea 26:8–26

Goulding CJ, Murray J (1976) Polynomial taper equations that are compatible with three volume equations. N Z J For Sci 5:313–322

Huang S, Yang Y, Wang Y (2003) A critical look at procedures for validating growth and yield models. In: Amaro A, Reed D, Soares P (eds) Modelling forest systems. CAB International, Wallingford, Oxfordshire, UK

Kerr G, Evans J (1993). Growing broadleaves for timber. Forestry Commission Handbook 9, London.

Kozak A (1988) A variable-exponent taper equation. Can J For Res-Rev Can Rech For 18:1363–1368. doi:10.1139/X88-213

Kozak A (2004) My last words on taper equations. For Chron 80:507–515. doi:10.5558/tfc80507-4

Kozak A, Kozak R (2003) Does cross validation provide additional information in the evaluation of regression models? Can J For Res 33:976–987. doi:10.1139/X03-022

Lindstrom MJ, Bates DM (1990) Nonlinear mixed effects models for repeated measures data. Biometrics 46:673–687. doi:10.2307/2532087

MARM (2011) Avance del Anuario de Estadística Forestal. Área de Medio Ambiente, Madrid, España

Muhairwe CK (1994) Tree form and taper variation over time for interior lodgepole pine. Can J For Res 24:1904–1913. doi:10.1139/X94-245

Myers RH (1990) Classical and modern regression with applications, 2nd edn. Duxbury Press, Belmont, CA

Pompa-García M, Corral-Rivas JJ, Hernández-Díaz JC, Álvarez-González JG (2009) A system for calculating the merchantable volume of oak trees in the northwest of the state of Chihuahua, Mexico. J For Res 20:293–300. doi:10.1007/s11676-009-0051-x

Reed DD, Green E (1984) Compatible stem taper and volume ratio equations. For Sci 30:977–990

Riemer T, Gadow KV, Sloboda B (1995) Eim modell zur Beschreibung von Baumschäften. Allg Forst Jagdztg 166:144–147

Rodríguez F, Broto M, Lizarralde I (2008). CubiFOR: herramienta para cubicar, clasificar productos, calcular biomasa y CO_2 en masas forestales de Castilla y León. Revista Montes: 33–3

SAS Institute Inc (2004a) SAS/STAT®, 0.1. User's Guide. SAS Institute Inc., Cary, NC

SAS SAS Institute Inc (2004b) SAS/ETS®, 0.1. User's Guide. SAS Institute Inc., Cary, NC

Thomas CE, Parresol BR (1991) Simple, flexible, trigonometric taper equations. Can J For Res 21:1132–1137. doi:10.1139/X91-157

Weisberg S (1985) Applied linear regression, 2nd edn. Wiley, Inc

Zhang D (1997) Cross-validation of non-linear growth functions for modeling tree height-diameter relationships. Ann Bot 79:251–257. doi:10.1006/anbo.1996.0334

Zimmerman DL, Núñez-Antón V (2001) Parametric modeling of growth curve data: an overview (with discussion). Test 10:1–73. doi:10.1007/BF02595823