Forest Systems
24(3), e034, 13 pages (2015)
eISSN: 2171-9845
http://dx.doi.org/10.5424/fs/2015243-06229
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)

RESEARCH ARTICLE

OPEN ACCESS

Comparison of stem taper equations for eight major tree species in the Spanish Plateau

Francisco Rodríguez[1]*, Iñigo Lizarralde[1] and Felipe Bravo[2]

[1] *föra forest technologies, Soria, Spain.* [2] *Sustainable Forest Management Research Institute University of Valladolid-INIA, Spain*

Abstract

Aim of study: A stem taper function and a compatible merchantable volume system are compared to evaluate which provides a better description of the stem profile for the main species in central Spain.

Area of study: This research was carried out in the region of Castile-Leon, located in Central Spain.

Material and Methods: A total of 6,357 trees were selected for destructive sampling. All models were fitted using a first-order continuous autoregressive error structure to address the problem of autocorrelation.

Main results: In terms of accuracy, the root mean square error (RMSE) in both models ranged from 0.75 to 2.72 depending on the species analyzed, presenting values similar to those reported in other studies. Small differences in the goodness-of-fit for both procedures were also found, and the Stud model provided better accuracy for 6 of the 8 species studied, with RMSE reductions of 0.5% to 8.6%. The RMSE obtained in the cross-validation phase was on average 1.22 times higher than what was obtained in the fitting phase.

Research highlights: The non-linear extra sum of squares method indicated that the stem taper differs among the five softwood species and three hardwood species. In hardwoods, the first inflection point is lower than in softwoods (at around 5%) and the second inflection point is higher (at around 85%) than those of softwoods.

Keywords: taper function; volume system; Central Spain; softwoods; hardwoods.

Citation: Rodríguez, F., Lizarralde, I., Bravo, F. (2015). Comparison of stem taper equations for eight major tree species in the Spanish Plateau. Forest Systems, Volume 24, Issue 3, e034, 13 pages. http://dx.doi.org/10.5424/fs/2015243-06229.

Received: 12 May 2014. **Accepted:** 20 Jul 2015

Funding: The Forest Services (*Consejería de Medio Ambiente de Castilla y León*) and Fundación Cesefor provided support for this study.

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Francisco Rodríguez: paco.rodriguez@fora.es

Introduction

Accurate predictions for wood products classified by merchantable size are a matter of interest for forest managers and forestry companies, in order to estimate the monetary value of some of the many commodities and services that forests provide to society. The accuracy of this estimate is directly related to its final use. Thus, while forest managers need it for planning and quantification of land use, forestry companies need it to assess the profitability of a harvest.

Nowadays, society demands multifunctionality from our forests, which increases the need for more detailed knowledge of the different products and services a forest can provide. Product classification thus becomes an important tool for assessing the role of forest and wood products, especially in light of climate changes due to carbon fixation and sequestration. Volume predictions by different merchantable sizes facilitates further assessment of the life cycle of wood products by making it possible to precisely estimate the carbon stored in every wood product and subsequently evaluate entire forests (Skog & Nicholson, 1998).

Volume prediction to any merchantable limit can be achieved by several methods, most of which involve the use of volume-ratio or stem taper equations (Crecente-Campo *et al.*, 2009). Volume-ratio equations predict merchantable volume as a percentage of total tree volume, while taper equations are mathematical formulae that describe the stem shape (Burkhart & Tomé, 2012). Integration of the taper equation from the ground to any height provides an estimate of the merchantable volume to that height (Bravo *et al.*, 2011). Stem form and taper variation have been widely studied using taper functions developed by forest researchers throughout the world (e.g., Newnham, 1992). Although merchantable volume equations derived from volume-ratio equations are very easy to use and develop, those obtained from taper functions are preferred nowadays; perhaps because they allow for estimation of diameter at a given height (Diéguez-Aranda *et al.*, 2006).

Compatibility means that an integrated model can be obtained through summation of the differential model. Thus, for a given merchantable volume equation, there is an intrinsically defined compatible taper function (Clutter, 1980). This implies that integration of the taper function from the ground to the top of the tree would provide the appropriate volume, and subsequently the merchantable volume equation (Fang *et al.*, 2000). Existing research indicates that segmented models appear to be more accurate than other model formulations (e.g., Corral-Rivas *et al.*, 2007) for creating compatible systems.

To develop a taper function, diameter/height data pairs are required along the stem. Most taper functions can be included in the following groups: single and segmented taper models, trigonometric equations, and variable-form taper models. Stem taper functions are usually based on the diameter at breast height (*dbh*), and predict diameter over bark (Rojo *et al.*, 2005; Crecente-Campo *et al.*, 2009); though they can also predict diameter inside bark (Garber & Maguire, 2003; Calama & Montero, 2006) and bark thickness (Laasasenaho *et al.*, 2005).

Forestry researchers in Spain have been developing taper equations since the 1970s. Recently, a considerable number of taper equations have been developed for particular regions and species: some for hardwoods but most for softwoods (Bravo *et al.*, 2011). Due to their complicated formulations, most taper functions are implemented with specific software in order to estimate total and merchantable volume from inventory data (e.g. SiManFor; www.simanfor.es).

In central Spain, most harvested tree species are Scots pine (*Pinus sylvestris* L.), Mediterranean maritime pine (*Pinus pinaster* Ait.), Pyrenean oak (*Quercus pyrenaica* Willd.), poplar (*Populus x euramericana* (Dode) Guinier), stone pine (*Pinus pinea* L.), Spanish juniper (*Juniperus thurifera* L.), black pine (*Pinus nigra* Arnold.) and beech (*Fagus sylvatica* L.). According to the Spanish National Forest Inventory (SNFI) in Central Spain these species occupy around 1,742,975 ha, have a standing volume of approximately 113,437,012 m^3 and produce an annual harvest of about 1,274,594 m^3.

The objective of this study was to compare a stem taper function and a compatible merchantable volume system to ascertain which provides a better description of the stem profile, in order to obtain accurate partial or total stem volume estimates for the main species in the Spanish plateau (central Spain).

Materials and Methods

Study area and data collection

This research was carried out in the region of Castile-Leon, located in Central Spain. The region covers approximately 9.4 million ha and is one of the most important areas for timber production in Spain. Predominant oak stands cover more than half of the area, while pine stands cover one third of the area. Altitude fluctuates between 110 and 2650 m above sea level. The climate in Castile-Leon is both continental and Mediterranean: average winter temperatures range between 4 and 7 °C, while average summer temperatures range from 19 to 22 °C; with three or four dry summer months that are typical of a Mediterranean climate. Average annual rainfall is only 450 to 500 mm, mostly in lower altitudes.

The data used in this study were collected in 242 public and private forests in Central Spain. Data from 1,844 Scots pines, 456 stone pines, 533 black pines, 1,715 Mediterranean maritime pines, 326 Spanish junipers, 302 Pyrenean oaks, 992 poplars and 189 beeches were used to test the statistical performance of the taper functions. Thus, a total of 6,357 trees were selected for destructive sampling using the protocol of Garber and Maguire (2003). The trees were felled from thinned and unthinned stands, which were subjectively selected to represent the existing range of site qualities. Before felling, two attributes were recorded for each sample tree: (i) diameter at breast height, D (to the nearest 0.1 cm); (ii) total tree height, H (to the nearest 0.01 m). Each sample tree was felled in a manner that minimized stem breakage. A tape measure was stretched along the bole and total height was recorded from the base of the stump to the top of the tree. The tree was divided into sections, and thin disks were removed for diameter measurements at the base of the stump; at 80 cm above the ground level; at breast height; and at a height midway between every other whorl pair above breast height (1–3 m intervals). For each disk, height (h, to the nearest 0.01 m) and diameter over bark (d, to the nearest 0.1 cm) were measured along two perpendicular axes. Log volumes in cubic meters were calculated using Smalian's formula. The topmost log was considered conical in shape. Total volume (V in m^3) was obtained by summing the log volumes. Between 3 and 40 disks were cut per tree, for a total of 87,568 disks. Summary statistics including the number of observations, arithmetic mean, standard deviation, and minimum/maximum values of the main variables are presented in Table 1.

Table 1. Summary statistics of the tree data set

Species [Sp]	Variable	No. of observations	Mean	Standard deviation	Maximum	Minimum
Scots pine [21]	No. of sections	23319	8.09	6.22	35.00	2.00
	D [cm]	1844	30.48	12.07	82.00	9.50
	H [m]	1844	16.76	5.24	35.80	5.05
	V [m^3]	1844	0.72	0.84	9.13	0.00
Stone pine [23]	No. of sections	4534	9.94	2.98	20.00	2.00
	D [cm]	456	36.49	14.30	98.00	12.10
	H [m]	456	10.67	3.85	23.40	3.57
	V [m^3]	456	0.73	0.83	5.00	0.01
Black pine [25]	No. of sections	5691	9.44	4.29	25.00	2.00
	D [cm]	533	24.86	8.06	51.85	9.85
	H [m]	533	15.35	4.70	29.80	5.00
	V [m^3]	533	0.44	0.38	2.41	0.01
Mediterranean maritime pine [26]	No. of sections	17576	6.63	4.36	23.00	2.00
	D [cm]	1685	30.75	8.84	77.00	10.25
	H [m]	1685	13.04	3.43	26.87	4.60
	V [m^3]	1685	0.52	0.43	4.54	0.02
Spanish juniper [38]	No. of sections	2593	7.95	2.49	15.00	2.00
	D [cm]	326	21.41	7.38	45.87	7.99
	H [m]	326	6.33	1.72	12.60	3.00
	V [m^3]	326	0.13	0.11	0.69	0.01
Pyrenean oak [43]	No. of sections	3046	10.09	3.23	20.00	3.00
	D [cm]	302	19.24	9.85	63.34	5.00
	H [m]	302	11.67	4.01	24.50	3.90
	V [m^3]	302	0.23	0.29	2.20	0.00
Poplar [58]	No. of sections	28980	29.24	4.56	40.00	17.00
	D [cm]	992	29.94	6.28	47.90	11.10
	H [m]	992	24.17	4.70	36.20	13.20
	V [m^3]	992	0.75	0.42	2.13	0.05
Beech [71]	No. of sections	1829	9.68	2.78	22.00	3.00
	D [cm]	189	25.93	10.29	72.45	9.51
	H [m]	189	18.29	4.15	31.80	8.10
	V [m^3]	189	0.49	0.55	5.31	0.04

Note: D = diameter at breast height over bark; H = total tree height (m); V = total volume (m^3).

Figure 1 shows relative height against relative diameter for the data set used. The range of the data reflects the magnitude of variation in stem form among the sample trees. To detect possible anomalies in the data and increase the efficiency of the process, the systematic approach proposed by Bi (2000) was applied. This approach is flexible because no assumptions about the parametric form of the regression model are needed. In this case, a local quadratic fitting with a smoothing parameter of 0.3 was used for all components. The maximum number of extreme values corresponded to Spanish juniper (around 20%). At the opposite end of the spectrum, extreme values for black pine and poplar were only 4% and 1%, respectively. In the remaining species, extreme values represented approximately 10% of the total. Most of these data points corresponded to stem deformations, large knots and other physical damage. Since taper functions are not intended for deformed stems, these data points (but not the whole tree) were excluded from further analysis. This approach corresponds to the LOESS procedure, using SAS/STAT software (SAS Institute Inc., 2010a).

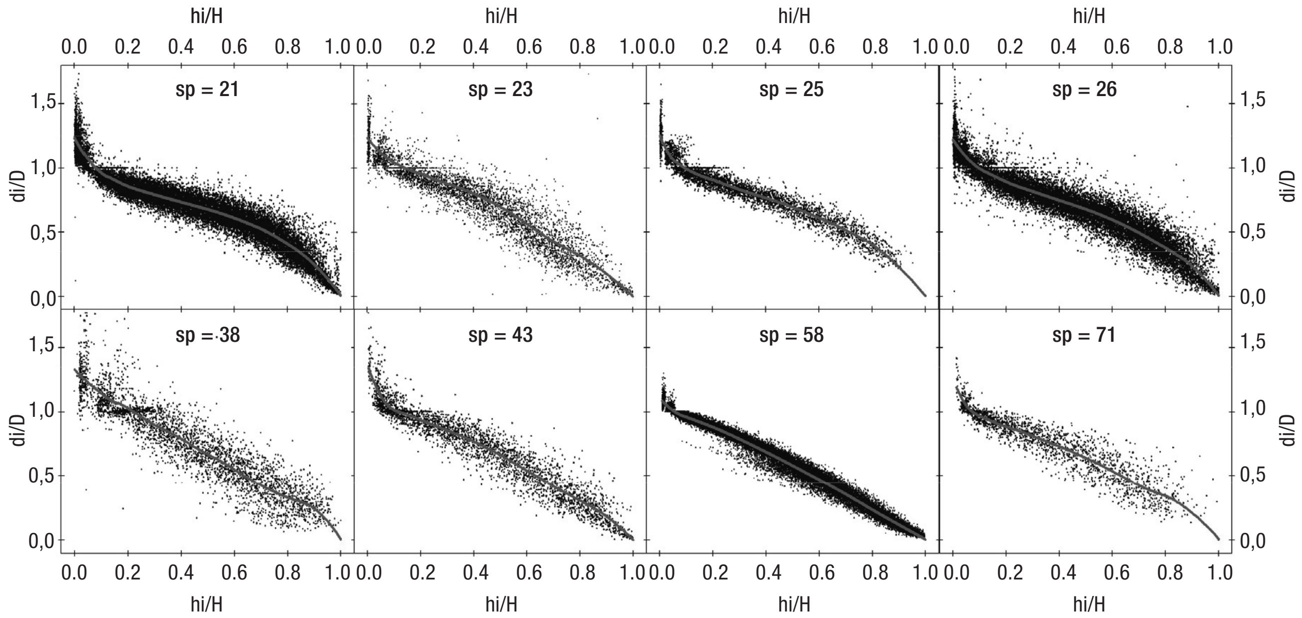

Figure 1. Data points of relative diameter and relative height plotted with a local regression Loess smoothing curve (smoothing factor = 0.3) for each species studied.

Functions selected for comparison

Numerous taper functions have been developed and many describe the diameter along the stem quite well. Among them, the segmented function of Fang et al. (2000), which we will refer to as the *Fang system*, and the *Stud model*, or Stud variable exponent function developed by Daquitaine *et al.* (1999), have shown very good results in many studies of several species in Spain (Rojo et al., 2005; Diéguez-Aranda *et al.*, 2006; Barrio-Anta *et al.*, 2007a; Rodriguez *et al.*, 2010). These two outperformed other functions in preliminary analyses and were therefore selected for further analysis.

Variable-exponent taper equations describe the stem shape with a changing exponent from ground to top. This approach is based on the assumption that the stem form varies continuously along the length of a tree. The Stud model is basically an allometric function of the form $d=u\cdot(h/H)^q$, where *u* is an exponential function that describes the butt region and *q* is the exponent term describing the tree form. The parameters are dendrometrically and biologically interpretable: θ_1 and θ_2 describe the upper and middle stem, respectively. The other parameters pertain to the function *u*, which describes the butt region, where θ_3 refers to width, θ_4 refers to length and θ_5 to height. The expression of this model is:

$$d=\left(1+\theta_3.\mathrm{e}^{-\theta_4\cdot\left(\frac{h}{H}\right)}\right)\cdot\theta_{50}+\theta_{51}\cdot D\cdot\left(1-\left(\frac{h}{H}\right)\right)^{\theta_1+\theta_{11}\cdot\left(\frac{100\cdot\mathrm{H}}{\mathrm{D}}\right)+\theta_2\left(1-\left(\frac{h}{H}\right)\right)} \quad [1]$$

The Fang system assumes three sections with a variable-form constant factor for each one. The expression of this model is as follows:

$$d=\mathrm{c}_1\sqrt{\mathrm{H}^{\left(\frac{\mathrm{k\text{-}b}_1}{\mathrm{b}_1}\right)}\left(1-\left(\frac{h}{H}\right)\right)^{\left(\frac{k-\beta}{\beta}\right)}\alpha_1{}^{I_1+I_2}\alpha_2{}^{I_2}} \quad [2]$$

where:

$$\left\{\begin{array}{l} I_1=1 \quad si\ p_1\le\frac{h}{H}\le p_2;\ 0 \text{ otherwise} \\ I_2=1 \quad si\ p_2\le\frac{h}{H}\le 1;\ 0 \text{ otherwise}\end{array}\right\}$$

$$k=\pi/40000$$

$$\beta=b_1^{1-(I_1-I_2)}b_2^{I_1}b_3^{I_2}$$

$$\alpha_1=(1-p_1)^{(b_2-b_1)\frac{k}{b_1b_2}}$$

$$\alpha_2=(1-p_2)^{(b_3-b_2)\frac{k}{b_2b_3}}$$

$$\mathrm{r}_0=\left(\frac{1-h_{st}}{H}\right)^{\frac{k}{b_1}};$$

$$r_1=(1-p_1)^{\frac{k}{b_1}}$$

$$\mathrm{r}_2=(1-p_2)^{\frac{k}{b_2}}$$

$$ci = \sqrt{\frac{a_0 D^{a_1} H^{a_2 - \frac{k}{b_1}}}{b_1(r_0 - r_1) + b_2(r_1 - \alpha_1 r_2) + b_3 \alpha_1 r_2}}$$

and a_i, b_i and p_i are the parameters to be estimated. Fang *et al.* (2000) also derived a compatible model for merchantable volume (v, in m^3) and total volume (V, in m^3) by direct integration of the taper model. Their expressions are:

$$v = c_1^2 H^{\frac{k}{b_1}}(b_1 r_0 + (I_1 + I_2)(b_2 - b_1) r_1 + I_2(b_3 - b_2)\alpha_1 r_2 - \beta\left(1 - \left(\frac{h}{H}\right)\right)^{\frac{k}{\beta}} \alpha_1^{I_1 + I_2} \alpha_2^{I_2} \qquad [3]$$

$$V = a_0 \cdot D^{a_1} \cdot H^{a_2} \qquad [4]$$

Although Eq. (4) was used to develop the compatible Fang system, any other volume equation can be used as input into the system.

The Stud model has been widely used with excellent results in radiata pine (Rodríguez *et al.*, 2004), Scots pine (Fonweban *et al.*, 2011) and poplar plantations (Rodríguez *et al.*, 2010). The Fang system has also been used extensively, mainly in loblolly and Ellioti pines (Fang et al., 2000), Scots pine (Diéguez-Aranda *et al.*, 2006), maritime pine (Rojo *et al.*, 2005) and different pines in Mexico (Corral-Rivas *et al.*, 2007).

Model fitting

The models were fitted using least squares techniques. However, there are several problems associated with stem taper function analysis that violate the fundamental least squares assumptions of independence of errors: the two most important are multicollinearity and autocorrelation (Kozak, 1997). The condition number (CN) was used to evaluate the presence of multicollinearity among variables in the models analyzed. Since the database contains multiple observations for each tree, it is realistic to expect that the observations within each tree are spatially correlated. The error terms were modeled using a continuous autoregressive error structure (CAR(x)), which makes it possible to apply the model to irregularly spaced, unbalanced data and overcome possible autocorrelation from the longitudinal data sets used for model fitting (Gregoire *et al.*, 1995). Autocorrelation plots are commonly used to check randomness in a data set by computing autocorrelations for data values at varying time lags. The presence of autocorrelation and the order of the CAR(x) were also assessed. Models were fitted using the SAS/ETS MODEL procedure (SAS Institute Inc., 2010b).

We followed the indications of Fang et al. (2000), for avoiding problems in the estimation of their system parameters when h = H; i.e., when d = 0. A small value, lower than the appreciation limit used in the data collection, was reassigned to diameters equal to zero. This approach does not significantly change the parameter estimates (Diéguez-Aranda *et al.*, 2006; Corral-Rivas et al., 2007; Crecente-Campo *et al.*, 2009).

Among the different options to estimate the parameters in the Fang systems, where the taper equation includes a total volume equation, in this study we prioritized the taper function, fitting it first and subsequently performing the predicted volume calculation from the estimation parameters obtained (Menéndez-Miguélez *et al.*, 2014). In this way, we achieved a more accurate comparison between the estimations of the Fang system and the Stud model where the same variable (d) was optimized during the fitting process in both cases.

Model comparison and validation

Estimates of the different fitted models were compared by numerical and graphical analyses. Four goodness-of-fit statistics were used: the adjusted coefficient of determination (R^2_{adj}), the mean bias error (BE), the root mean square error (RMSE), and the Bayesian Information Criterion (BIC).

The taper functions were also assessed using box plots for diameter residuals (*d*) by relative height along the stem (5%, 15%, 25%, and so on up to 95%). The same was done for volume residuals by diameter classes. These graphs, assessed by class for relative height or diameter at breast height, are very important for visualizing areas or tree size classes for which the functions provide especially good or especially poor predictions (Kozak & Smith, 1993).

An n-way cross-validation of each species was carried out, estimating the residual for one tree by excluding that tree every time. The RMSE and the adjusted model efficiency (MEF_{adj}), equivalent to the R^2_{adj} of the fitting phase) were calculated from the residuals.

Species differences in the taper equations

To evaluate whether the taper equations vary among the different species, the nonlinear extra sum of squares method was used (Bates & Watts, 1988). This is frequently applied for comparing different geographic regions (Crecente-Campo *et al.*, 2009) or analyzing differences among species (Corral-Rivas *et al.*, 2007; Rodríguez *et al.*, 2010). The nonlinear extra sum of squares method is based on the likelihood-ratio test for detecting simultaneous homogeneity among parameters and requires the fitting of full and reduced models. The reduced model consists of the same set of parameters for all species and the full model incorporates the different sets of parameters for each species. The full model was obtained by expanding each global parameter to include an associated parameter and a dummy variable to differentiate species. If the F-test (Eq. 9 above) results revealed that there were no differences (p > 0.05) for different species, only a composite model fitted to the combined data was needed. The nonlinear extra sum of squares follows an F-distribution and uses the expression:

$$F = \frac{\left(SSE_R - SSE_F\right)/\left(df_R - df_F\right)}{SSE_F / df_F} \qquad [5]$$

where SSE_R is the error sum of squares of the reduced model, SSE_F is the error sum of squares of the full model, and df_R and df_F are the degrees of freedom of the reduced and full models, respectively.

Results

The models were first fitted without expanding the error terms to account for autocorrelation, thus a strong autocorrelation among all models was observed. A first-

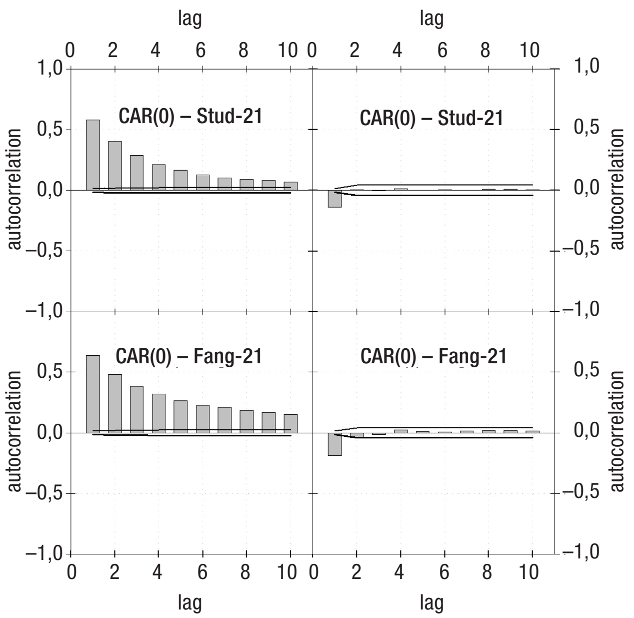

Figure 2. Autocorrelation plots for Scots pine. First and second rows show the Stud model and Fang system, respectively. Left and right columns show autocorrelation plots fitted without the autocorrelation parameters and using continuous autoregressive error structures of first order, respectively. Continuous lines represent the 95% confidence region.

order continuous autoregressive error structure was required to model the inherent autocorrelation of the hierarchical data. An exception was found in the case of Spanish juniper in the Fang system, where the model did not converge. The autocorrelation may be explained by the effect of stand conditions (e.g., stand density, thinning effects, etc.) on stem form (Calama & Montero, 2006), because stand conditions have a great impact on crown length and stem diameter (Burkhart & Tomé, 2012). Figure 2 shows an example of the autocorrelation plots of the residuals obtained for Scots pine (sp=21).

All the parameters were significant at $p<0.05$ (Table 2), except in the case of Spanish juniper in the Stud model, where a_{11} and a_4 were not significant. The detailed error analysis of diameter predictions showed a similar trend in all the fits evaluated (sixteen combinations of the residuals, two models for each species). Although the poplar analysis rendered poorer graphics than the other species, all cases indicated a random pattern of residuals around zero (Figure 3). All species obtained good data fits in both models and both phases. The root mean square error obtained in the cross-validation phase ranged from 1.01 to 1.49 times higher and was 1.22 times higher on average than those obtained in the fitting phase (Table 3). More than 99% of the total variance of the diameter was explained for poplar plantations and black pine; more than 96% of total variance was explained for Spanish juniper. As expected, taper functions for stone pine gave worse predictions than equations obtained for other pines, primarily due to its high variability in bark thickness. The greatest RMSE was found in stone pine ($RMSE_{STUD}$ = 2.7241; $RMSE_{FANG}$ = 2.6939), while the largest bias occurred in beech (BE_{STUD} = –0.1057; BE_{FANG} = –0.1882). By contrast, smaller RMSE were found in poplar plantations ($RMSE_{STUD}$ = 0.7531; $RMSE_{FANG}$ = 0.7574) and smaller BE in Scots pine (BE_{STUD} = -0.0689; BE_{FANG} = -0.0314). In all species, the Stud model and Fang system showed a similar behavior pattern for the root mean square error against classes of diameter at breast height: accuracy decreased with increasing size class (Table 4). Multicollinearity in both models was moderate, as was inferred from the condition number (Table 3), with values ranging from 9 to 30 for the Stud model and 34 to 110 for the Fang system.

The box plots of diameter residuals against relative height classes (Figure 4) did not show any clear systematic trend that could describe deficient behavior in the models or any clear differences between the Stud model and the Fang system.

Results of the fitting process for full and reduced forms of the model are shown in Table 5. All of the 10 possible paired comparisons in softwoods and the 3 possible paired comparisons in hardwoods produced

Table 2. Parameter estimates for Stud model (Daquitaine et al., 1999) for each species. Standard errors of the estimated coefficients are shown in parentheses

	θ10	**θ11**	**θ2**	**θ3**	**θ4**	**θ5**	**ρ1**
21	0.643484 (0.009760)	0.001226 (0.000173)	0.592286 (0.029500)	0.471799 (0.013900)	10.631420 (0.233300)	0.830364 (0.008120)	0.872243 (0.004780)
23	1.176657 (0.025800)	0.006518 (0.000621)	0.886843 (0.081800)	0.214083 (0.021900)	14.671320 (1.577500)	0.979925 (0.018200)	0.575957 (0.015600)
25	0.740926 (0.028800)	0.001542 (0.000431)	0.782074 (0.101400)	0.453832 (0.051100)	9.669382 (0.789200)	0.817718 (0.029700)	0.799161 (0.021100)
26	0.699124 (0.011500)	0.001529 (0.000255)	0.272968 (0.027400)	0.273056 (0.009810)	12.810610 (0.440600)	0.959094 (0.007630)	0.785089 (0.006720)
38	0.805341 (0.023800)	–0.00047 * (0.000734)	-0.178400 (0.062700)	0.046853 (0.016900)	26.461670 * (20.849200)	1.216238 (0.014000)	0.281314 (0.022400)
43	1.061312 (0.025600)	0.002895 (0.000384)	0.328551 (0.043100)	0.264640 (0.013200)	38.122070 (2.885000)	1.065794 (0.008030)	0.687558 (0.017200)
58	1.406493 (0.017800)	0.005139 (0.000214)	0.290212 (0.012200)	0.091016 (0.002350)	45.150030 (2.572900)	1.032982 (0.002190)	0.719496 (0.004370)
71	0.933074 (0.043100)	0.002620 (0.000583)	0.151976 (0.065300)	0.166771 (0.018700)	23.029430 (4.358500)	1.026560 (0.013700)	0.936709 (0.010200)

Note: Sp = species as defined in Table 1. Asterisk (*) indicates a not significant parameter estimate at a probability level of 5%.

Table 3. Parameter estimates for the system of Fang *et al.* (2000) for each species. Standard errors of the estimated coefficients are shown in parentheses

	a0	**a1**	**a2**	**b1**	**b2**	**b3**	**p1**	**p2**	**ρ1**
21	0.000051 (0.000001)	1.845867 (0.006210)	1.045022 (0.007610)	0.000011 (0.000000)	0.000038 (0.000000)	0.000030 (0.000000)	0.093625 (0.000527)	0.763750 (0.003160)	0.998258 (0.000025)
23	0.000067 (0.000002)	1.698754 (0.010100)	1.210604 (0.012100)	0.000006 (0.000000)	0.000033 (0.000000)	0.000026 (0.000000)	0.021072 (0.000947)	0.475953 (0.007690)	0.996810 (0.000204)
25	0.000049 (0.000002)	1.982808 (0.012500)	0.905147 (0.014900)	0.000014 (0.000000)	0.000036 (0.000000)	0.000029 (0.000001)	0.091275 (0.001940)	0.781990 (0.009650)	0.996393 (0.000472)
26	0.000048 (0.000001)	1.929098 (0.004900)	0.976356 (0.005450)	0.000010 (0.000000)	0.000035 (0.000000)	0.000033 (0.000000)	0.064157 (0.000578)	0.681476 (0.015700)	0.996435 (0.000144)
38	0.000074 (0.000002)	1.86289 (0.011800)	0.901233 (0.013500)	0.000001 (0.000000)	0.000028 (0.000000)	0.000037 (0.000002)	0.008578 (0.001330)	0.711639 (0.013500)	0.980337 (0.009110)
43	0.000051 (0.000002)	1.867810 (0.009070)	0.989625 (0.014600)	0.000007 (0.000000)	0.000030 (0.000000)	0.000032 (0.000002)	0.047757 (0.000978)	0.825279 (0.038500)	0.998535 (0.000056)
58	0.000044 (0.000000)	1.872438 (0.005160)	1.023328 (0.005870)	0.000013 (0.000000)	0.000028 (0.000000)	0.000026 (0.000000)	0.032326 (0.000563)	0.645012 (0.005420)	0.996334 (0.000087)
71	0.000120 (0.000042)	2.036193 (0.008810)	0.799343 (0.021700)	0.000015 (0.000001)	0.000033 (0.000000)	0.005194 (0.002920)	0.074439 (0.002720)	0.873445 (0.007560)	0.882674 (0.005150)

Note: Sp = species as defined in Table 1. Asterisk (*) indicates a not significant parameter estimate at a probability level of 5%.

significant F-values, suggesting that significantly different equations are needed for different species. The greatest differences (as inferred from the F-values) were found between hybrid poplar and beech ($F_{58\text{-}71}$=505.73) and the smallest differences occurred between black pine and Mediterranean maritime pine ($F_{25\text{-}26}$=4.63).

Discussion

Numerous taper equations have been developed for all but two of the species analyzed. Stem taper functions for Spanish juniper and Pyrenean oak have not yet been developed. This paper provides a good starting point for fitting these species, for which only growth (Adame *et al.*, 2008) and biomass (Carvalho & Parresol, 2005) have been modeled so far. In terms of accuracy, the proposed models show that RMSE ranged from 0.75 to 2.72 depending on the species analyzed; a range that is similar to what has been reported in other softwood and hardwood studies. Taper functions for Scots pine have been widely studied in Spain (Diéguez-Aranda *et al.*, 2006; Crecente-Campo *et al.*, 2009) and Europe (Lappi, 1986; Petersson, 1999; Karlsson *et al.*, 2002). All report small errors, similar to those obtained in this work (RMSE around 1.55). The stem form of stone pine has only been characterized in Spain by Calama and Montero (2006), who obtained a 60% lower RMSE (1.71 compared to 2.72 obtained in this study) because they analyzed the diameter inside the bark. There are few taper models for black pine, (Meridieu, 1998), and the results are similar to those observed in Central Spain (RMSE = 0.84). Maritime pine is a widely modeled species (Rojo et al., 2005). Accuracy in the Mediterranean variety (RMSE: 1.479), which is usu-

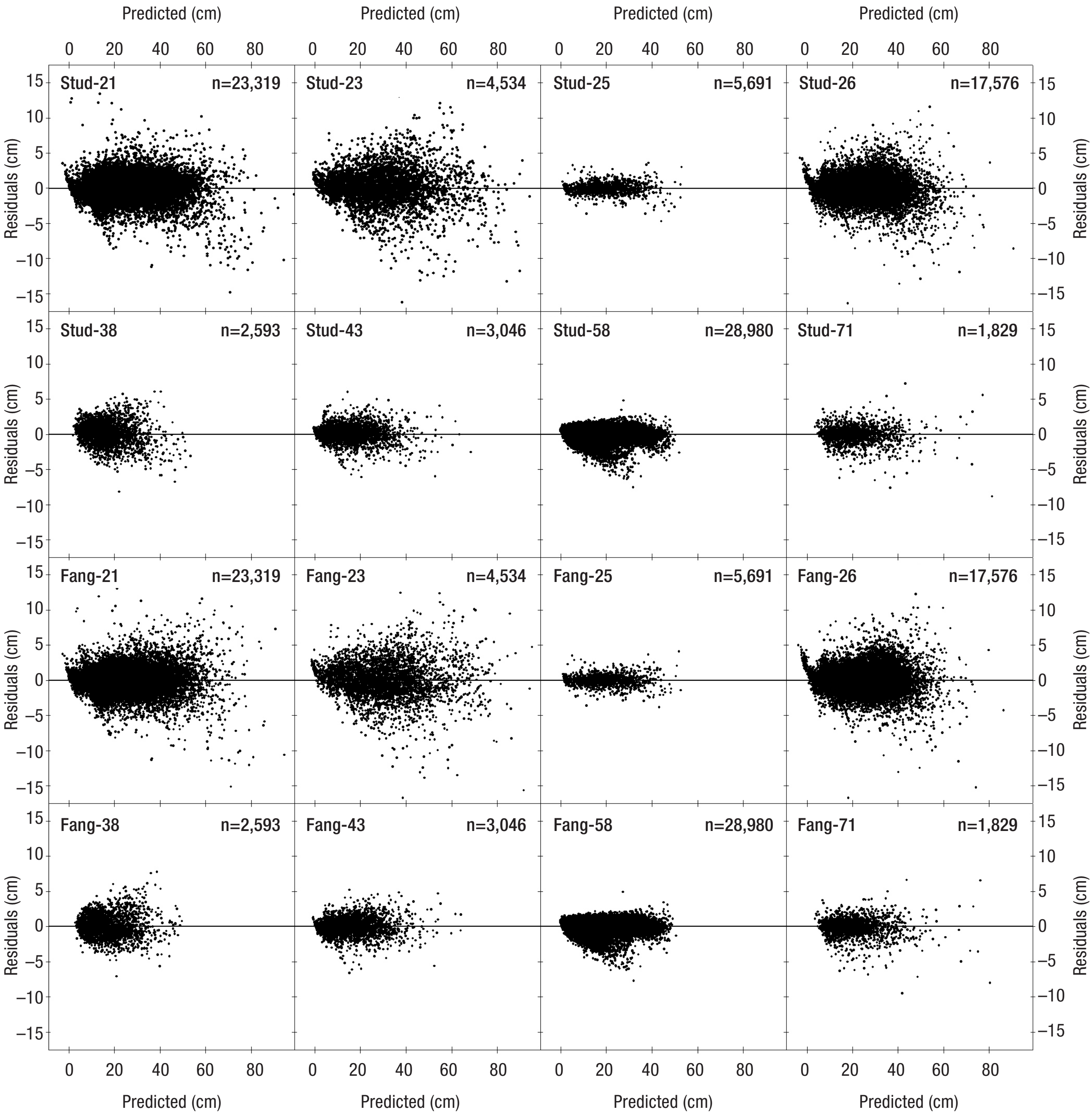

Figure 3. Scatterplots of residuals versus predicted residuals. The Stud model is shown in the first and second rows; the Fang system in the third and fourth rows. The number next to the model indicates the species evaluated (i.e. Stud-21 represents the scatterplot of the Stud model in Scots pine); *n* indicates the number of observations.

Table 4. Goodness-of-fit for the species and models evaluated in the fitting phase and in the cross-validation phase.

		Fitting phase					Cross-validation phase	
Sp	**Model**	**R^2**	**BE**	**RMSE**	**BIC**	**CN**	**MEF**	**RMSE**
21	Stud	0.9845	–0.0689*	1.5489	7272.9	23.1	0.9648	1.9981
	Fang	0.9829	–0.0314*	1.6285	8082.4	42.9	0.9632	2.1171
23	Stud	0.9714	0.0836	2.7241	3551.8	21.4	0.9612	3.1695
	Fang	0.972	0.0598	2.6939	3520.5	42.6	0.9569	3.3384
25	Stud	0.9931	–0.0459	0.8408	–151.1	30.5	0.9865	1.1744
	Fang	0.993	–0.0142	0.8486	–135.2	39.3	0.9843	1.2637
26	Stud	0.9775	–0.0294*	1.6426	6875.6	18.0	0.9580	2.1847
	Fang	0.9756	0.0049	1.7121	7456.6	43.7	0.9569	2.2788
38	Stud	0.9602	0.1519*	1.6892	1053.9	14.5	0.9594	1.7047
	Fang	0.9651	0.0115	1.5808	930.1	34.2	0.9642	1.6003
43	Stud	0.9855	0.0316	1.1855	426.1	9.6	0.9805	1.3712
	Fang	0.9853	–0.0036	1.1909	444.8	35.8	0.9783	1.4475
58	Stud	0.9949	0.0336*	0.7531	–6873.7	26.8	0.9850	0.8736
	Fang	0.9948	0.033*	0.7574	–6726.0	109.9	0.9897	0.8772
71	Stud	0.9836	–0.1057*	1.2901	381.3	10.3	0.9769	1.5267
	Fang	0.9806	–0.1882*	1.4016	507.0	–	0.9776	1.5019

Note: Sp = species as defined in Table 1. Asterisk (*) indicates a not significant BE at a probability level of 5%.

Table 5. Root mean square error against diameter at breast height classes (DC) for the different models and species.

Sp	DC	5	15	25	35	45	55	65	75
21	**n**	18	2906	5025	5607	3159	1532	605	79
	Fang	0.489	1.160	1.780	2.499	3.634	5.092	6.381	11.523
	Stud	0.546	1.147	1.585	2.142	3.369	4.810	5.964	9.738
23	**n**	–	282	802	1204	1018	329	255	169
	Fang	–	1.837	3.453	5.692	7.580	10.935	13.907	26.139
	Stud	–	2.100	3.471	5.658	7.553	10.965	14.310	29.203
25	**n**	–	233	434	309	128	–	–	–
	Fang	–	0.321	0.709	0.840	1.166	–	–	–
	Stud	–	0.363	0.662	0.865	1.063	–	–	–
26	**n**	–	1267	4588	6817	2554	518	138	16
	Fang	–	1.514	2.469	2.884	3.896	4.427	9.716	5.941
	Stud	–	1.262	2.152	2.654	3.766	4.196	9.704	11.557
38	**n**	56	897	1025	245	52	–	–	–
	Fang	0.324	1.303	2.844	5.345	5.077	–	–	–
	Stud	0.822	1.566	3.014	6.255	7.789	–	–	–
43	**n**	367	1053	891	351	68	22	12	–
	Fang	0.350	0.755	1.642	2.875	3.742	4.954	12.807	–
	Stud	0.466	0.758	1.608	2.749	4.199	4.353	10.532	–
58	**n**	–	858	12134	12737	2279	–	–	–
	Fang	–	0.230	0.420	0.657	1.056	–	–	–
	Stud	–	0.257	0.417	0.642	1.064	–	–	–
71	**n**	13	432	728	326	107	33	6	5
	Fang	0.299	0.577	1.249	3.008	5.952	5.383	23.933	25.698
	Stud	0.260	0.562	1.112	2.658	4.725	3.203	13.867	23.485

Note: Sp = species as defined in Table 1. Asterisk (*) indicates a not significant BE at a probability level of 5%.

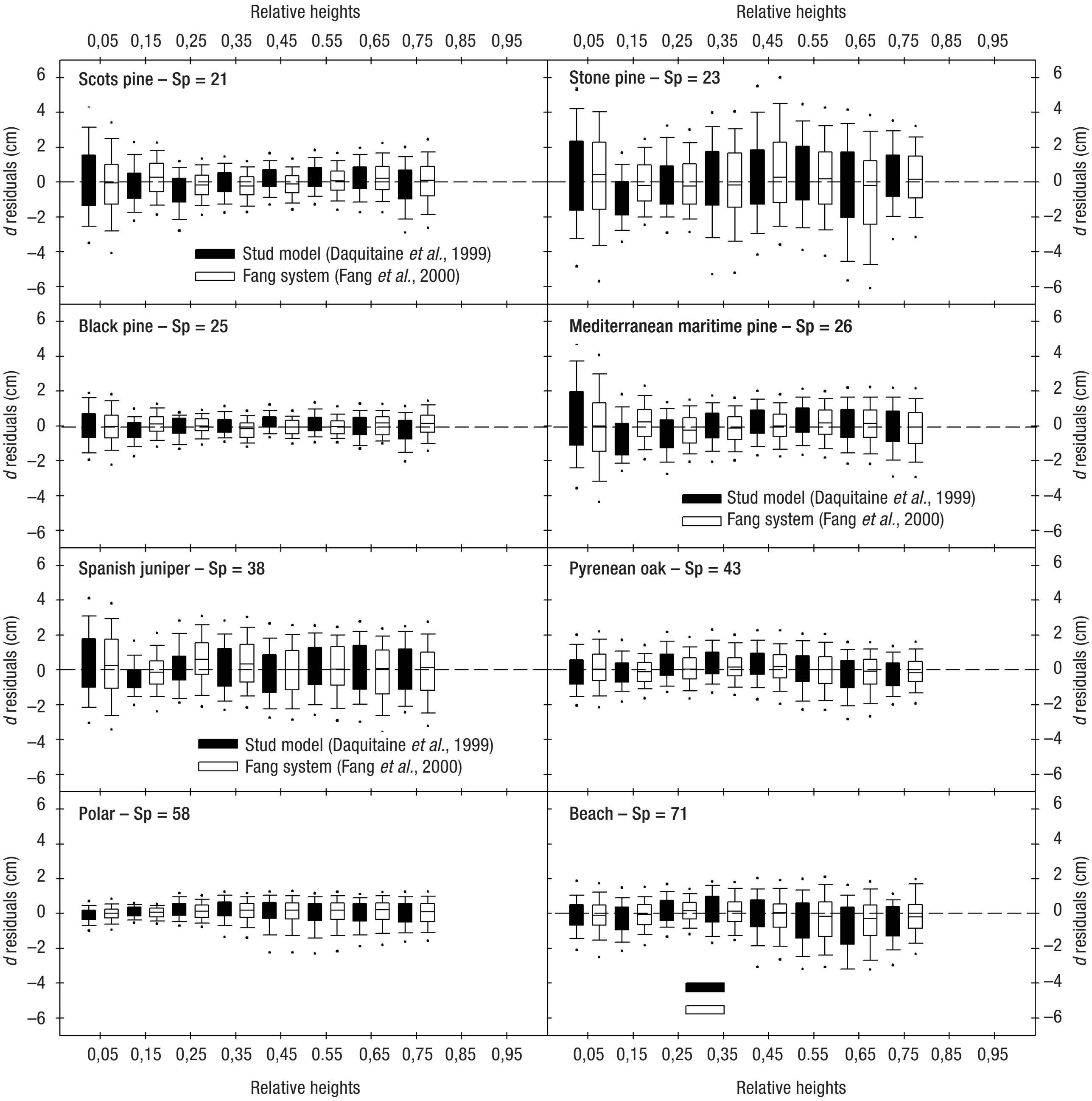

Figure 4. Box plot graph showing error in diameter along different relative stem heights for the eight analyzed species. The boxes represent the interquartile ranges. The prediction errors for maximum and minimum diameter over bark are represented by point; the 5th/95th percentiles are represented by the small top and bottom horizontal lines crossing the vertical lines, respectively.

ally measured in natural stands, is only 10% lower than in the Atlantic variety (RMSE: 1.642), which is usually measured in restocked stands. Hardwood species displayed noteworthy goodness-of-fit. Very high precision was obtained for Pyrenean oak (RMSE<1.2, on average 30% lower than other oaks, (Tarp-Johansen *et al.*, 1997; Barrio-Anta *et al.*, 2007a)), which explained more than 98% of the total variance of the dependent variables. There are many published papers on poplar plantations, due to their high commercial value (Roda, 2001; Barrio-Anta *et al.*, 2007b; Rodriguez *et al.*, 2010). In this work, the poplar was the most successful species, probably due to minimal variability in its management and the fact that all data corresponded to the same clone. The results of the present work are slightly better than those reported in other published papers (about 10% more accurate than Barrio-Anta et al., 2007b). Finally, results for beech were similar to those observed in other European locations (Trincado *et al.*, 1997; Stoltze, 2000).

Table 6. F-test of the differences between species obtained with the Stud model (Daquitaine *et al.*, 1999)

	Sp	n	Full			Reduced			Extra sum of squares		
			SSE_F	nº $parms_F$	dfF=n-nº parmsF	SSE_R	nº $parms_R$	df_R=n-nº $parms_R$	df_R-df_F	F-value	Prob>F
Soft-woods	21-23	22,990	75,971.0	14	22,976	83,615.7	7	22,983	7	330.29	0.0000
	21-25	20,035	46,174.0	14	20,021	46,250.7	7	20,028	7	4.75	0.0002
	21-26	36,294	50,283.3	14	36,280	50,168.6	7	36,287	7	11.82	0.0000
	21-38	21,206	60,988.8	14	21,192	56,103.2	7	21,199	7	242.52	0.0000
	23-25	5,163	30,844.4	14	5,149	31,767.9	7	5,156	7	22.02	0.0000
	23-26	19,957	72,947.2	14	19,943	77,440.2	7	19,950	7	175.48	0.0000
	23-38	6,334	40,654.9	14	6,320	37,416.6	7	6,327	7	71.92	0.0000
	25-26	17,002	43,653.8	14	16,988	43,737.0	7	16,995	7	4.63	0.0000
	25-38	3,379	1,344,361.0	14	3,365	7,751.6	7	3,372	7	477.94	0.0003
	26-38	18,173	50,129.8	14	18,159	51,560.2	7	18,166	7	74.02	0.0000
Hard-woods	43-58	30,772	19,756.8	14	30,758	21,561.1	7	30,765	7	401.28	0.0000
	43-71	4,414	6,608.8	14	4,400	7,049.6	7	4,407	7	41.93	0.0000
	58-71	29,658	18,616.6	14	29,644	20,839.8	7	29,651	7	505.73	0.0000

Note: SSEF, dfF, SSER and dfR are the sum of squared errors and the degrees of freedom associated with the full and reduced models, respectively.

The results for Scots pine were similar to other findings (Crecente-Campo *et al.*, 2009), where small differences in the goodness-of-fit for both procedures were also found. The Stud model provided lower errors for 6 out of the 8 species, with reductions of 0.004 to 0.111 cm in RMSE (or 0.5% to 8.6%). Both models estimated diameters along the tree reasonably well across the diameter classes that were sampled. It should be noted that the bias tends to increase with the diameter class (Diéguez-Aranda *et al.*, 2006; Corral-Rivas *et al.*, 2007).

Cao et al., (1980), found that the variable exponent model represented stem shape quite accurately, especially in the high-volume butt region. The pattern of d residual plots against relative height classes (Figure 4) was similar to corresponding plots for other species (Garber & Maguire, 2003; Crecente-Campo *et al.*, 2009). Both models tended to underestimate the diameters in the lower and upper sections, whereas the mid-section diameters were overestimated. For relative heights of 0–10%, both models showed larger standard errors of the estimates than at other height intervals, except in poplar plantations. In the remaining sections, the standard errors of the estimates were more or less constant, depending on the species analyzed. Because stem analysis was usually stopped at a diameter of 7.5 cm, few measurements exist in the top sections and these results should be considered carefully. However, as the latter part of the stem accumulates least volume and is the least valuable, these results do not have a great impact on the overall performance and applied use of the models. For sections closer to the ground level, both models provided good estimates. Accurate diameter predictions in these sections are more relevant because the basal log is particularly important from a commercial point of view. According to Crecente-Campo *et al.*, (2009), this behavior is common in stem taper predictions. All these statistics and plots revealed no clear advantage of the Stud model or the Fang system. However, Diéguez-Aranda *et al.*, (2006) consider the Fang system to be advantageous because it is compatible with both a merchantable and a total volume equation. Additionally, a new or pre-existing volume equation can be used as input into the system, making its application more flexible.

The non-linear extra sum of squares method indicated that the stem taper differs among the five softwood species and three hardwood species. All of the 10 possible paired comparisons in softwoods and the 3 possible paired comparisons in hardwoods produced significant F-values. In softwoods, the greatest differences (as inferred from the F-values) occurred between black pine and Spanish juniper (F-value = 477.94) while the smallest differences were found between black pine and Mediterranean maritime pine (F-value = 4.63). In hardwoods, the greatest differences appeared between beech and hybrid poplar (F-value = 505.73) while the smallest differences were observed between Pyrenean oak and beech (F-value=41.93). These results are probably due to the strong apical dominance of the hybrid poplar compared to the Pyrenean oak and the beech, and their systematic sylviculture. All the pines studied were found to have the first inflection point at around 10% of total height, and the second inflection point at around 70% of total height (Table 2), with the exception of stone pine. Similar inflection points have been obtained for other pines in different locations around Spain (e.g., Crecente-Campo *et al.*, 2009). In hardwoods, the first inflection point is lower than in softwoods (around 5% of total height) and the second is higher (around 85%), except in poplar plantations, where the second inflection point is similar to that of pines.

Acknowledgements

We are grateful to the Forest Services (*Consejería de Medio Ambiente de Castilla y León*) and Fundación Cesefor for providing support for this study and access to study sites. Additionally, many people assisted in the field work. We thank each of them, particularly Alfredo Fernández, Iñigo García, Beatriz Detorre and Luis Carlos Becerril.

References

Adame P, Hynynen J, Cañellas I, del Río M, 2008. Individual-tree diameter growth model for rebollo oak (*Quercus pyrenaica* Willd.) coppices. For Ecol Manage 255: 1011–1022.

Barrio-Anta M, Diéguez-Aranda U, Castedo-Dorado F, Álvarez-González JG, Gadow Kv, 2007a. Merchantable volume system for pedunculate oak in northwestern Spain Ann For Sci 64(5): 511-520. http://dx.doi.org/10.1051/forest:2007028

Barrio-Anta M, Sixto H, Cañellas I, González-Antoñanzas F, 2007b. Sistema de cubicación con clasificación de productos para plantaciones de Populus x euramericana (Dode) Guinier cv. 'I-214' en la meseta norte y centro de España. Invest Agrar: Sist Recur For 16(1): 65-75.

Bates DM, Watts DG, 1988. Nonlinear Regression Analysis and Its Applications.

Bi H, 2000. Trigonometric variable-form taper equations for Australian eucalyptus. For Sci 46: 397–409.

Bravo F, Álvarez-González JG, Del Río M, Barrio-Anta M, Bonet JA, Bravo-Oviedo A, Calama R, Castedo-Dorado F, Crecente-Campo F, Condés, *et al.*, 2011. Growth And Yield Models In Spain: Historical Overview, Contemporary Examples And Perspectives. Forest Systems 20(2): 315-328. http://dx.doi.org/10.5424/fs/2011202-11512

Burkhart H, Tomé M, 2012. Modeling Forest Trees and Stands. Springer Ed. 1st Edition., 547 pp. http://dx.doi.org/10.1007/978-90-481-3170-9

Calama R, Montero G, 2006. Stand and tree level variability on stem form and tree volume in *Pinus pinea* L: a multilevel random components approach. Invest Agr Sist Recur For 15(1): 24-41 http://dx.doi.org/10.5424/srf/2006151-00951

Cao QV, Burkhart HE, Max TA, 1980. Evaluations of two methods for cubic-foot volume prediction of loblolly pine to any merchantable limit. For Sci 26: 71–80.

Carvalho JP, Parresol BR, 2005. A site model for Pyrenean oak (*Quercus pyrenaica*) stands using a dynamic difference equation. Can J For Res 35: 93-99. http://dx.doi.org/10.1139/x04-155

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

Corral-Rivas JJ, Diéguez-Aranda U, Corral S, Castedo-Dorado F, 2007. A merchantable volume system for major pine species in El Salto, Durango (Mexico). For Ecol Manage 238(1-3): 118-129.

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: 808-820. http://dx.doi.org/10.1051/forest/2009078

Daquitaine R, Saint-Andre L, Leban JM, 1999. Product properties prediction - improved utilisation in the forestry-wood chain applied on spruce sawnwood: Modelling stem properties distribution. Final Report sub-task A2.1. Nancy.

Diéguez-Aranda U, Castedo F, Álvarez JG, Rojo A, 2006. Compatible taper function for Scots pine plantations in northwestern Spain. Can J For Res 36: 1190–1205. http://dx.doi.org/10.1139/x06-008

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

Fonweban J, Gardiner B, Macdonald E, Auty D, 2011. Taper functions for Scots pine (*Pinus sylvestris*, L.) and Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Northern Britain. Forestry, 84(1): 49-60. http://dx.doi.org/10.1093/forestry/cpq043

Garber SM, Maguire DA, 2003. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. For Ecol Manage 179: 507–522.

Gregoire TG, Schabenberger O, Barrett JP, 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can J For Res 25: 137–156. http://dx.doi.org/10.1139/x95-017

Karlsson A, Albrektson A, Elfving B, Fries C, 2002. Development of *Pinus sylvestris* main stems following three different precommercial thinning methods in a mixed stand. Scand J Forest Res 17(3): 256–262. http://dx.doi.org/10.1080/028275802753742927

Kozak A, Smith JHG, 1993. Standards for evaluating taper estimating systems. For Chron 69: 438–444.

Kozak A, 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. Can J For Res 27: 619–629. http://dx.doi.org/10.1139/x97-011

Laasasenaho J, Melkas T, Alden S, 2005. Modeling bark thickness of Picea abies with taper curves. For Ecol Manage 206: 35-47.

Lappi J, 1986. Mixed linear models for analyzing and predicting stem form variation of Scots pine. Communicationes Instituto Forestalis Fenniae 134: 1–69.

Menéndez-Miguélez M, Canga E, Alvarez-Alvarez P, Majada J, 2014. Stem taper function for sweet chestnut (*Castanea sativa Mill.*) coppice stands in northwest Spain.

Ann For Sci 71(7): 761-770. http://dx.doi.org/10.1007/s13595-014-0372-6
Meredieu C, 1998. Croissance et branchaison du Pin laricio (*Pinus nigra* Arnold ssp. Laricio (Poiret) Maire). Elaboration et évaluation d'un système de modèles pour la prévision de caractéristiques des arbres et du bois. Thèse de Doctorat, Université Claude.
Newnham RM, 1992. Variable-Form Taper Functions for Four Alberta Tree Species. Can J For Res 22: 210-223. http://dx.doi.org/10.1139/x92-028
Petersson H, 1999. Biomassafunktioner för trädfunktioner av tall, gran och björk i Sverige. Arbetsrapport 59. Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Science, Umea.
Roda JM, 2001. Form function for the 'I-214' poplar merchantable stem (Populus x euramericana (Dode) Guinier cv cultivar 'I-214'). Ann For Sci 58: 77-87. http://dx.doi.org/10.1051/forest:2001108
Rodríguez F, Broto M, Cantero A, 2004 Aplicación del programa "Cubica" a distintos regímenes de gestión del pino radiata en el País Vasco. Cuadernos Sociedad Española Ciencias Forestales (18):193-198.
Rodríguez F, Pemán J, Aunós A, 2010. A reduced growth model based on stand basal area. A case for hybrid poplar plantations in northeast Spain. For Eco Man 259: 2093-2102.
Rojo A, Perales X, Sánchez-Rodríguez F, Álvarez-González JG, Von Gadow K, 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). Eur J Forest Res 124: 177–186. http://dx.doi.org/10.1007/s10342-005-0066-6
SAS Institute Inc., 2010a. SAS/STAT® 9.2 User's Guide, Second Edition. Available at: http://support.sas.com/ [last revised on February 3, 2010].
SAS Institute Inc., 2010b. SAS/ETS®9.2 User's Guide. Available at: http://support.sas.com/ [last revised on April 16, 2010].
Skog KE, Nicholson GA, 1998. Carbon cycling through wood products: the role of wood and paper products in carbon sequestration. Forest Products Journal 48:75-86.
Stoltze P, 2000. Overenstemmende stammemasseog stammesidefunktioner for bøg i Danmark. Ministy of Environment and Energy. Danish Forest and Landscape Research Institute. Working report.
Tarp-Johansen MJ, Skovsgaard JP, Madsen SF, Johannsen VK, Skovgaard I, 1997. Compatible stem taper and stem volume functions for oak (*Quercus robur* L. and *Q. petrae* (Matt.) Liebl) in Denmark. Ann For Sci 54(7): 577–595. http://dx.doi.org/10.1051/forest:19970701
Trincado G, Gadow Kv, Sandoval V, 1997. Estimación de volumen comercial en latifoliadas. Bosque 18(1): 39-44.