Dynamic dominant height growth model for Spanish and Tunisian cork oak (*Quercus suber* L.) forests

M. Sánchez-González[1]*, B. Stiti[2], H. Chaar[3] and I. Cañellas[1]

[1] Centro de Investigación Forestal-INIA. Instituto Universitario de Investigación en Gestión Forestal Sostenible. Ctra. A Coruña, km 7,5. 28040 Madrid. Spain
[2] Institut Nacional de Recherches en Génie Rural. Eaux et forêts, INRGREF. BP. 10, Ariana. Tunisia
[3] Institut National d'Agronomy de Tunisie (INAT). 43, rue Charles Nicole. 1082 Tunis-Nahrajène

Abstract

Seven simple and advanced dynamic polymorphic functions were considered to develop a dominant height growth model for Spanish and Tunisian cork oak forests. Data from 115 stem analyses performed in two regions in each country were used to fit the equations. Parameter estimates were obtained using the Dummy variable method. Both numerical, graphical and biological consistency were used to compare alternative models. The dynamic equation finally selected was derived from the Hossfeld model by considering the shape parameter to be related to site productivity. An analysis of the dominant height growth patterns between the two countries indicated that the same dominant height growth model was valid for both countries. This dominant height growth model allows estimation of dominant height with a level of reliability of at least 83% from an age of 15 years for a prediction interval of less than 40 years.

Key words: dummy variable method; base-age invariant; cork oak; Tunisia; Spain.

Resumen

Modelo dinámico de crecimiento en altura para monte alcornocal (*Quercus suber* L.) españoles y tunecinos

Para desarrollar un modelo de crecimiento en altura dominante para monte alcornocal en España y Túnez se han probado siete funciones polimórficas simples y dinámicas. Para ello se han utilizado datos procedentes de 115 análisis de troncos realizados en dos regiones distintas de cada país. Para estimar los parámetros se utilizó el método de las variables ficticias. Los modelos fueron comparados de forma numérica y gráfica. La ecuación dinámica finalmente seleccionada fue derivada a partir de la ecuación de crecimiento de Hossfeld considerando el parámetro de forma como relacionado con la productividad de la estación. El análisis del patrón que sigue el crecimiento en altura dominante en ambos países indica que el mismo modelo de crecimiento en altura dominante es valido para ambos países. Este modelo de crecimiento en altura dominante nos permite estimar la altura dominante con una fiabilidad de más de un 83% a partir de los 15 años para un periodo de predicción menor de 40 años.

Palabras clave: método de las variables ficticias; invariante respecto de la edad; alcornoque; Túnez; España.

Introduction

The natural distribution area of cork oak (*Quercus suber* L.) extends across in the countries of the western Mediterranean basin, covering a total area of over 2 million hectares between latitudes 33°N and 45°N (Pereira, 2007). Cork oak forests and woodlands currently occupy considerable areas in Spain (475,000 ha) and Portugal (713,000 ha) as well as in Morocco (348,000 ha), Algeria (230,000 ha) and Tunisia (90,000 ha)

(Pereira, 2007). They can also be found at a smaller scale in areas of France (68,000 ha) and Italy (65,000 ha) (Fig. 1).

In addition, cork oak plays a key role in the economies of the regions in which it is found, not only due to the direct economic benefits from cork production, but because of indirect benefits derived from activities associated with this type of forests such as hunting and cattle grazing or other forest products, for example, firewood, acorns, etc. In addition, the cork oak forests are also valued for their ecological role in mitigating desertification and soil erosion, as well as for their contribution to the preservation of biodiversity and carbon storage (Cañellas *et al.*, 2008).
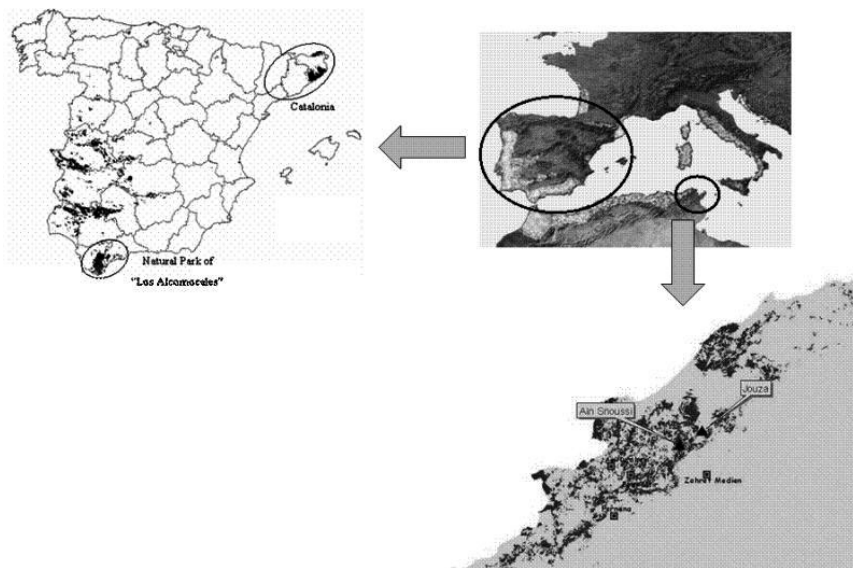
Figure 1. Natural distribution of *Quercus suber* L. in the Mediterranean Basin and location of the studied regions in Tunisia and Spain.

Cork oak stands can be differentiated into open cork oak woodlands (low tree density, called «dehesas» in Spain and «montados» in Portugal) and cork oak forests (higher tree density) according to ecological, silvicultural and productive characteristics (Montero and Cañellas, 1999; San Miguel *et al*., 1992; Pereira, 2007). Although cork is the main product of open cork oak woodlands, they also provide grazing for domestic and wild livestock. These different uses are regulated by reducing the number of trees per hectare. Cork oak woodlands present an open structure with a density of 20-100 trees per hectare, 10-50% canopy cover and a well developed understory of annual grasses. Cork oak forests, on the other hand, have a higher density (over 200 trees per hectare) and a substantial understory of shrubs such as *Arbutus unedo, Phillyrea angustifolia, Erica* sp., *Cistus* sp., etc. The main product is also cork, so the management of these stands is oriented towards optimizing this production.

In order to facilitate inventory updates and growth projections it is very important to adequately characterize site quality in these forests. This information allows applying management plans and ensures sustainable forest management. In even-aged management systems site quality can be assessed developing site index models. A wide variety of methods have been employed in the construction of these models. According to Clutter *et al*. (1983), three general methods can be used for developing site index curves: the guide curve approach, the parameter prediction method and the algebraic difference approach. The latter permits the development of models based on dynamic equations, with the advantage, among others, of greater flexibility as well as being more parsimonious (*e.g.* Cieszewski and Bailey, 2000; Cieszewski, 2001, 2003; Krumland and Eng, 2005).

Dynamic equations can be derived by using base-age invariant methods such as the algebraic difference approach (ADA) or the generalized algebraic difference approach (GADA). The ADA method is founded on the replacement of a base-model site-specific parameter with its initial-condition solution, the dynamic polymorphic site models derived by this method had single asymptotes for all productivity sites and can be also defined as simple polymorphic models (Cieszewski *et al*., 2006). The GADA method is a generalization of the ADA and can be used to derive the same models as those derived from the ADA but with the advantage that the base equations can be expanded according to various theories about growth characteristics (*e.g.* asymptote, growth rate), allowing more than one parameter to be site-specific, and permitting the derivation of more flexible dynamic equations (Cieszewski and Bailey, 2000; Cieszewski, 2001, 2002, 2003). It includes the ability to simulate concurrent polymorphism and multiple asymptotes, so the models derived by GADA method can also be defined as advanced polymorphism (Cieszewski *et al.*, 2006).

The implementation of a base-age invariant dominant height growth model in the sense of Bailey

and Clutter (1974) involves the use of dynamic equations along with the estimation of the model parameters using a base-age invariant method, which does not depend on the arbitrary base age and allows the use of all the data available. An effective system for parameter estimation needs to be based on identifying individual trends represented in the data (Bailey and Clutter, 1974; García, 1983; Cieszewski *et al*., 2000; Cieszewski, 2003). These trends can be modeled by considering local and global parameters. Local parameters vary among individuals and global parameters are common to all individuals. In practice, both base age specific methods and base-age invariant methods can be used, but the latter assume that the data measurements always contain environmental as well as measurement errors (both on the left and the right sides of the model) which must be modeled (Cieszewski, 2003).

The dominant height growth models available for defining site quality in the distribution area of cork oak are very scarce (Chaar *et al*., 2005; Sánchez-González *et al*., 2005; Tomé *et al*., 2001). The first and only attempt at modeling height growth for dominant trees in Spanish cork oak forests was carried out by Sánchez-González *et al*. (2005) using the dynamic and polymorphic equation derived by McDill-Amateis (1992) from the Hossfeld (1822) model by considering only one parameter to be site-specific. In Tunisia the dominant height model available for cork oak was developed by Chaar *et al*. (2005) using the algebraic difference approach through Amaro's *et al*. (1998) modification of Lundqvist-Korf model (Korf, 1939; Lundqvist, 1957). This model was not accurate enough and had the inconvenience of overestimating dominant height at earlier and advanced ages for which there were insufficient data. In addition, this model is anamorphic meaning that all the dominant height curves have the same parameters, excepting the asymptote. Among the desirable attributes of site equations (Bailey and Clutter, 1974; Cieszewski and Bailey, 2000; Cieszewski, 2002, 2003) one of the most frequently listed criteria is polymorphism meaning that the shape of the curves varies from site to site.

The aims of this study are to developed an improved dominant height growth model for Tunisia and to obtain a common model for Mediterranean cork oak forests including data from Spain and Tunisia by using the difference equation method. In addition, the dominant height growth patterns and the differences between the two countries has been analyzed.

Material

Study area and data

Stem analysis data from Spanish cork forests were collected in two of the country's most important cork producing areas and can be considered representative of higher density Spanish cork oak forests (Sánchez-González *et al*., 2005): the Natural Park of «Los Alcornocales» in the South, and Catalonia in the North-East. 85 sample trees were selected, 45 in the Natural Park of «Los Alcornocales» and 40 in Catalonia.

In Tunisia, stem analysis data were collected in the Snoussi and Jouza regions, located in the Kroumirie-Mogods, an area in the northwest of the country. This area is the most important one in terms of forest production. The area of the Ain Snoussi forest is 3,230 ha (Direction Générale des Forêts, 1983) and that of the Jouza forest is 5,867 ha (Direction Générale des Forêts, 1986). 82 healthy sample trees, were selected in 99 plots.

A total of 57 trees from the Spanish sample and 58 from Tunisia were used to fit the model. The rest of the trees were discarded for initial model selection and parameter estimation since only stump and breast height data was available.

The following variables were measured for each sample tree in both countries: diameter at breast height (cm), crown projection diameter (m) measured in two perpendicular directions, and bole and tree heights (m) measured with a tape-measure on the felled tree. Table 1 shows the characteristics of the sample trees in each country.

In each country, sample trees deemed to be dominant, healthy and rot free, were selected in even-aged stands with different site conditions. Trees were felled as close to the ground as possible. Sectioning was carried out by cutting disks at the base of the tree, at a height of 50 cm, at breast height (1.30 m), and at 50 cm intervals along the stem. In the Tunisian sample, sectioning trees was only carried out up to bole height. The rings were counted on each disk and tree age was determined as the number of rings on the base disks. The age at each height level was calculated as the difference between tree age and the number of rings at that level.

Tree height at a given age (obtained from stem analysis data) (Carmean, 1972; Fabbio *et al*., 1994) was not corrected because it was considered that since height growth in cork oaks is very slow, the possible error

Table 1. Mean, standard deviation and range of the main characteristics of the sample trees in the studied areas

Area	n_f	n_d	Variable	Fitting data set			
				Min	Max	Mean	Std Dev
AS	34	25	d	11.78	54.11	24.97	9.41
			h	3.00	13.00	7.33	2.82
			Crown	2.00	11.00	5.46	2.20
			Age	43	131	90	20.88
			N	125.00	1,625.00	586.54	314.89
			G	3.05	109.89	32.27	21.24
JOU	24		d	5.41	55.07	20.24	11.94
			h	3.40	12.50	7.67	2.93
			hf	1.15	6.90	3.82	1.76
			Crown	1.25	7.90	3.80	1.83
			Age	44	107	81	21.55
			N	67.55	261.66	148.45	72.38
			G	7.19	67.58	23.14	19.75
CAT	27	13	d	11.14	60.48	24.61	12.98
			h	4.70	12.00	7.78	1.66
			hf	1.60	3.60	2.40	0.59
			Crown	1.79	9.40	4.72	2.32
			Age	30	158	60	31.89
			N	102.3	500.2	260	30.1
			G	3.9	31.2	13	25.3
PNLA	30	15	d	10.82	26.10	17.83	4.74
			h	4.30	9.70	6.59	1.38
			hf	1.50	5.73	2.88	1.12
			Crown	1.55	4.95	3.08	1.03
			Age	34	65	44	7.80
			N	72	350	220	35.2
			G	4.2	32.6	14	38.4

AS: Ain Snoussi. JOU: Jouza. CAT: Catalonia. PNLA: Natural Park of «Los Alcornocales». n_f: number of sample trees in the fitting data set. n_d: number of sample trees not included in the fitting data set (trees with only two height-age pairs). d: diameter at breast height (cm). h: tree height (m). hf: bole height (m). Crown: mean crown diameter (m). Age: number of rings at stump dominant height (years). N: number of trees per hectare (stems $\cdot h^{-1}$.; G: basal area ($m^2 \cdot ha^{-1}$).

produced as a consequence of considering the crosscut height as the maximum reached by the tree that year would be imperceptible.

Methods

Dominant height growth modeling

Candidate functions

A total of seven simple and advanced polymorphic dynamic models were formulated. Table 2 shows the different dynamic functions used as candidate dominant height growth models in the parameter estimation process. Three base equations previously used for cork oak dominant height-age modeling in Spain and Tunisia were considered: the Lundqvist-Korf (Lundqvist, 1957; Korf, 1939) equation, the Richards (1959) function and the Hossfeld model, in particular the base equation was Hossfeld IV (Kiviste, 1998).

Model C1 was formulated considering only one parameter to be site-specific from the Lundqvist-Korf (Lundqvist, 1957; Korf, 1939) base equation, being polymorphic with a common asymptote. Model C2 was derived by Cieszewski (2004) from the same base equation considering the asymptote as an exponential function of the unobserved site variable X and the shape parameter as a linear function of the inverse X. The resulting model is polymorphic with variable asymptotes changing with site productivity.

Table 2. Candidate equations for dominant height growth modelling

Citation source and base size-site-time equation	Solution for X with initital values (t_0, Y_0)	Dynamic equation	Function code
Bailey & Clutter, 1974 $Y = b_1 e^{-X \cdot t^{-b_3}}$	$X_0 = -\ln\left(Y_0 b_1^{-1}\right) \cdot t_0^{b_3}$	$Y = b_1 \left(Y_0 b_1^{-1}\right)^{\left(\frac{t_0}{t}\right)^{b_3}}$	C1
Cieszewski, 2004 $Y = e^X e^{-\left(b_1 + b_2 X^{-1}\right) \cdot t^{-b_3}}$	$X_0 = \frac{1}{2} t_0^{-b_3} \left(b_1 + t_0^{b_3} \ln\left(Y_0\right) \pm \sqrt{4 b_2 t_0^{b_3} + \left(-b_1 - t_0^{b_3} \ln\left(Y_0\right)\right)^2} \right)$	$Y = e^{X_0} e^{-\left(b_1 + b_2 X^{-1}\right) t^{-b_3}}$	C2
Clutter *et al.,* 1983 $Y = b_1 \cdot \left(1 - e^{-Xt}\right)^{b_3}$	$X_0 = -\ln\left(1 - \left(Y_0 b_1^{-1}\right)^{(1/b_3)}\right) t_0^{-1}$	$Y = b_1 \cdot \left(1 - \left(1 - \left(Y_0 \cdot b_1^{-1}\right)^{(1/b_3)}\right)^{\frac{t}{t_0}}\right)^{b3}$	C3
Krumland & Eng, 2005 $Y = e^X \cdot \left(1 - e^{b_1 \cdot t}\right)^{\frac{b_2 + b_3}{X}}$	$X_0 = \frac{1}{2}\left(\left(\ln\left(Y_0\right) - b_2 L_0\right) \pm \sqrt{\left(\ln\left(Y_0\right) - b_2 L_0\right)^2 - 4 b_3 L_0}\right)$ with $L_0 = \ln\left(1 - e^{-b_1 t_0}\right)$	$Y = Y_0 \cdot \left(\left(1 - e^{-b_1 \cdot t}\right) \cdot \left(1 - e^{-b_1 \cdot t_0}\right)^{-1}\right)^{\frac{b_2 + b_3}{X_0}}$	C4
McDill & Amateis, 1992 $Y = \dfrac{b_1 \cdot t^{b_3}}{t^{b_3} + X}$	$X_0 = t_0^{-b_3}\left(b_1 Y_0^{-1} - 1\right)$	$Y = b_1 \cdot \left(1 - \left(1 - \left(b_1 \cdot Y_0^{-1}\right)\right) \cdot \left(t_0 \cdot t^{-1}\right)^{b_3}\right)^{-1}$	C5
Cieszweski, 2001 $Y = \dfrac{\left(b_1 + X\right) \cdot t^{b_3}}{t^{b_3} + b_2 / X}$	$X_0 = \left(Y_0 - b_1\right) + \sqrt{\left(Y_0 - b_1\right)^2 + 2 b_2 Y_0 t_0^{-b_3}}$	$Y = Y_0 \cdot \dfrac{t^{b_3} \cdot \left(t_0^{b_3} \cdot X_0 + b_2\right)}{t_0^{b_3} \cdot \left(t^{b_3} \cdot X_0 + b_2\right)}$	C6
Cieszweski, 2003 $Y^3 = \dfrac{\left(b_1 + X\right) \cdot t^{b_3}}{\left(\frac{1}{2} b_2 / X\right) + t^{b_3 - 1}}$	$X_0 = \frac{1}{2}\left(\left(Y_0^3 t^{-1} - b_1\right) \pm \sqrt{\left(Y_0^3 t^{-1} - b_1\right)^2 + 2 b_2 Y_0^3 t^{-b_3}}\right)$	$Y = Y_0 \cdot \left(\left(\left(b_2 + t_0^{b_3-1} \cdot X_0\right) \cdot t^{b_3}\right) \cdot \left(\left(b_2 + t^{b_3-1} \cdot X_0\right) \cdot t_0^{b_3}\right)^{-1}\right)^{1/3}$ or if $b_3 < 1$ $Y = Y_0 \cdot \left(\left(\left(b_2 \cdot t_0^{1-b_3} + X_0\right) \cdot t\right) \cdot \left(\left(b_2 \cdot t^{1-b_3} + X_0\right) \cdot t_0\right)^{-1}\right)^{1/3}$	C7

Models C3 and C4 were derived from the Richards (1959) base equation. The first one was formulated considering only one parameter to be site-specific. In this case, is a simple polymorphic model. The fourth model, derived by Krumland and Eng (2005), is an advanced polymorphic model which considers more than one parameter as a function of site productivity and more specifically, expresses the asymptote as an exponential function of X and the shape parameter as a linear function of the inverse of the unobserved site variable X.

Models C5 and C6 were derived from the Hossfeld IV equation (1822). C5 was formulated considering only the shape parameter to be site-specific being, therefore, polymorphic with a single asymptote. C6 provides polymorphic curves with variable asymptotes and was derived by Cieszewski (2001) by replacing the asymptote with a constant plus X and the shape parameter with the inverse of X.

The last model evaluated, C7, was derived by Cieszewski (2003) from the cubical transformation of the Hossfeld IV function (1822) with the exponent in the denominator reduced by 1. The resulting advanced polymorphic model was derived by considering the maximum yield parameter as a linear function of X and

the half-saturation parameter as a function of the inverse of X.

Parameter estimation

Among the different base-age invariant parameter estimation techniques available, the Dummy variable method was deemed suitable for our purposes. This method is described in Cieszewski *et al.* (2000) and uses an initial height as a starting value for the site-specific parameter for each tree, assuming all measurements of a particular tree to have the same starting value. The initial height must be substituted by a sum of terms containing a site-specific parameter and a dummy variable for each tree. During the fitting process the initial height is estimated simultaneously along with all the global parameters and the sum of terms containing the site-specific parameters and dummy variables collapses for each tree into a single parameter unique to each one.

In order to address the occurrence of serial correlation resulting from repeated measurements taken from the same tree, which is a situation common to stem analysis, the error term was modeled using an *x*-order continuous autoregressive error structure [CAR(*x*)]. This error structure permits the application of the models to irregularly spaced, unbalanced data (Gregoire *et al.*, 1995; Zimmerman & Núñez-Antón, 2001). The CAR(*x*) expands the error terms as follows:

$$e_{ij} = \sum_{x=1}^{n} \Psi_x \rho^{t_{ij} - t_{ij-x}} e_{ij-x} + \varepsilon_{ij} \qquad [1]$$

where e_{ij} is the jth ordinary residual on the ith individual (*i.e.,* the difference between the observed and the estimated height of tree i at age measurement j), $\Psi_x = 1$ for j > 1 and is zero for j = 1, ρ is the *x*-order autoregressive parameter to be estimated, t_{ij}-t_{ij-x} is the time distance separating the jth from the jth-*x* observations within each individual ($t_{ij} > t_{ij-x}$) and ε_{ij} is now the error term under conditions of independence. To test for the presence of autocorrelation and the order of the CAR(x) to be used, graphs representing residuals *versus* lag-residuals from previous observations for each tree were examined visually. The heterocedasticity was analyzed by plotting residuals against predicted values. If necessary weighting factors should be applied to work under conditions of homocedastic residuals distribution.

The Dummy variable method including the CAR(*x*) error structure was programmed using a SAS macro which uses the SAS/ETS MODEL procedure (SAS Institute Inc., 2004), allowing for dynamic updating of the residuals. The Marquardt algorithm, which is most useful when the parameter estimates are highly correlated (Fang and Bailey, 1998; Parresol, 2001), was used for fitting the models.

Model selection

The evaluation of the models was based on numerical and graphical analyses of the residuals. Four statistics were examined: the bias, which reflects the deviation of the model with respect to observed values; the root mean square error (*RMSE*), which analyses the precision of the estimates; the coefficient of determination for nonlinear regression (*R²*) and Akaike's information criterion differences or delta AIC (Δ_i). The expressions may be summarized as follows:

$$bias = \frac{\sum (y_i - \hat{y}_i)}{n} \qquad [2]$$

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - p}} \qquad [3]$$

$$R^2 = r^2_{y_i \hat{y}_i} \qquad [4]$$

$$\Delta_i = AIC_i - AIC_{min} \qquad [5]$$

where y_i and $\hat{y}_i$ are the measured and estimated values of the dependent variable respectively, n is the total number of observations used to fit the model, p is the number of model parameters, $r_{y_i \hat{y}_i}$ is the correlation coefficient for a linear regression between the observed and predicted values of the dependent variable (Ryan, 1997), AIC_i is the value of Akaike's information criterion (Akaike, 1974) for model *i,* AIC_{min} is the minimum of the different AIC_i values. As a rule of thumb, a $\Delta_i < 2$ suggests substantial evidence for the model, values between 4 and 7 indicate that the model has considerably less support, whereas a $\Delta_i > 10$ indicates that the model is very unlikely (Burnham and Anderson, 2002).

Another important step in evaluating the models was to perform graphical analyses on both the residuals and the appearance of the fitted curves overlaid on the trajectories of the dominant height growth for each individual. The biological sense of the models through

its asymptotic value, which had to be realistic, was also evaluated; as well as the curves shape. Besides, the behavior of the models outside the range of the values of the modeling data was assessed.

The characterization of the model error was analyzed simultaneously using a validation technique. Since an independent validation data set was not available, a self-sufficient resampling type validation method (multiple cross-validation) was used. Thus, the models were fitted n times, leaving out each tree once, so that the number of fittings was equal to the number of trees. The residuals obtained in the multiple cross-validation were used to calculate the bias, $RMSE$, the model efficiency (MEF, equivalent to the R^2 in the fitting phase) and Akaike's information criterion differences or delta AIC (Δ_i). Although this approach does not really constitute a method for model validation (Vanclay and Skovsgaard, 1997; Pretzsch *et al*., 2002; Huang *et al*., 2003), it has been used as an additional criterion for selecting the best model (Myers, 1990).

Comparison of dominant height growth models between countries

First a dominant height growth model for each country was selected. Then, the question of whether significant differences existed between the growth patterns was assessed. This comparison was made using two tests based on the likelihood-ratio method, the non-linear extra sum of squares (Draper and Smith, 1981; Ratkowski, 1983) and the Lakkis-Jones test (Khatree and Naik, 1995). The likelihood ratio methods are the most appropriate in nested methods where a set of parameters are shared by two models (Gregoire *et al*., 1995). Both methods require the fitting of a full model where a different set of parameters is needed for each of the regions considered, along with a reduced model with a common set of parameters for all regions. The full model is obtained by expanding the parameters with n-1 dummy variables, n being the number of regions involved.

The non-linear extra sum of squares method uses the following statistic:

$$F = \frac{\left(\dfrac{SSE_r - SSE_f}{df_r - df_f}\right)}{\left(\dfrac{SSE_f}{df_f}\right)} \quad [6]$$

where SSE_f and SSE_r are the sum of squares error for the full and reduced model respectively; and df_f and df_r are the degrees of freedom for the full and reduced model respectively. The F statistic is compared with an F-distribution for degrees of freedom of $v = df_r - df_f$ and $u = df_f$. If $F > Fcrit$, then a model is required for each region.

The Lakkis-Jones test uses the L statistic:

$$L = \left(\frac{SSE_f}{SSE_r}\right)^{\frac{n}{2}} \quad [7]$$

where n is the total number of observations and SSE_f and SSE_r are the sum of squares error for the full and reduced model respectively. If homogeneity of parameters exists then $-2.ln(L)$ follows a Chi squared distribution with v degrees of freedom.

Results

Dominant height growth modeling

The analysis of residuals versus lag-residuals within the same tree indicated that all models required a first order continuous autoregressive error structure, CAR(1), to account for autocorrelation. The fitted curve shapes were not significantly altered in comparison to the models fitted without considering such correction. The sole purpose of autocorrelation correction was to prevent underestimation of the covariance matrix of the parameters, thereby making it possible to carry out the usual statistical tests (West *et al.,* 1984).

Tables 3 and 4 show the results obtained by fitting the candidate equations. Models C2 and C7 did not meet the convergence criterion, even when the number of maximum iterations was increased, the convergence criteria was decreased, or the initial parameter values were changed, so these models were considered further. Table 4 shows only those models with all parameters significant at the 5% level tested (including the site-specific parameter for each tree). All of them explained more than 95% of the total variance in the fitting and validation phases (Table 4) and provided a band of residuals scattered randomly around zero with homogeneous variance and no trends. The bias and root mean square error of dominant height growth is analyzed by age classes in Figure 2. Models performed in a similar way for all ages. The poorest behavior corresponded to ages beyond 55 due to the reduced number of observations for that age class.

Table 3. Parameter estimates, corresponding standard errors and *P*-values for the models analyzed

Model	Para-meter	Tunisia				Spain				Tunisia and Spain			
		Estimate	Standard Error	*t* value	Approx. *Pr.* > |*t*|	Estimate	Standard Error	*t* value	Approx. *Pr.* > |*t*|	Estimate	Standard Error	*t* value	Approx. *Pr.* > |*t*|
C1	*b1*	3,121.24	7,599.20	0.41	0.6816	314.19	259.40	1.21	0.2265	733.66	681	1.08	0.2817
	b3	0.16	0.06	2.69	0.0077	0.26	0.05	5.23	<0.0001	0.21	0.04	5.52	<0.0001
	$\rho 1$	0.92	0.02	50.36	<0.0001	0.85	0.02	37.64	<0.0001	0.90	0.01	68.15	<0.0001
C3	*b1*	21.78	5.21	4.18	<0.0001	15.15	1.57	9.62	<0.0001	17.80	1.79	9.94	<0.0001
	b3	1.25	0.09	13.95	<0.0001	1.52	0.08	18.61	<0.0001	1.38	0.06	23.91	<0.0001
	$\rho 1$	0.91	0.02	49.19	<0.0001	0.85	0.02	37.35	<0.0001	0.90	0.01	65.84	<0.0001
C4	*b1*	0.01	0.00	1.66	0.0975	0.01	0.00	3.62	0.0003	0.008	0.003	3.09	0.0021
	b2	1.29	1.07	1.21	0.2281	1.24	1.05	1.18	0.2389	–0.30	0.53	–0.57	0.5676
	b3	3.69	0.78	4.75	<0.0001	3.72	0.47	7.96	<0.0001	4.11	0.18	23.43	<0.0001
	$\rho 1$	0.91	0.02	49.17	<0.0001	0.86	0.02	37.97	<0.0001	0.90	0.01	66.91	<0.0001
C5	*b1*	30.70	8.95	3.43	0.0007	19.16	2.37	8.08	<0.0001	23.49	2.88	8.17	<0.0001
	b3	1.24	0.08	15.75	<0.0001	1.46	0.06	23.67	<0.0001	1.35	0.0472	28.47	<0.0001
	$\rho 1$	0.91	0.02	49.20	<0.0001	0.85	0.02	37.35	<0.0001	0.90	0.0136	65.85	<0.0001
C6	*b1*	30.89	7.20	4.29	<0.0001	20.12	4.49	4.48	<0.0001	25.12	3.79	6.63	<0.0001
	b2	–1947.98	774.90	–2.51	0.0125	–317.79	1,116.90	–0.28	0.7761	–901.66	823.4	–1.1	0.2739
	b3	1.28	0.07	17.14	<0.0001	1.45	0.06	22.53	<0.0001	1.35	0.05	28.66	<0.0001
	$\rho 1$	0.91	0.02	47.90	<0.0001	0.85	0.02	37.22	<0.0001	0.90	0.01	65.84	<0.0001

Figures 3 and 4 show the graphic appearance of the curves of models with every parameter statistically significant for Tunisian and Spanish data respectively. In Figure 3 it can be appreciated that models C3, C5 and C6 behaved similarly and described that data adequately. Figure 4 shows the graphic behavior of models C3 and C5 overlaid on the trajectories of observed Spanish values over time. It can be appreciated that model C3 showed a rapid increase in growth for the best quality while the growth culmination in this curve was achieved at a relatively young age, which is not according to the empirical knowledge on cork oak (Montero and Cañellas, 1999). These models were therefore rejected.

When fitting the equations for the two countries combined, equation C5 proposed by McDill and Amateis (1992), again presented the best performance of all the equations.

Comparison of dominant height growth models among countries

The values of the non-linear extra sum of squares F test and the Lakkis Jones test are presented in Table 5. The results revealed that, at a significance level of 5%, the null hypothesis of parameter homogeneity cannot be rejected for comparison between countries or bet-

Table 4. Values of the goodness-of-fit statistics for fitting and cross-validation phases for the analyzed models

	Model	Fitting phase				Cross-validation phase			
		Bias	RMSE	R^2	ΔAIC	Bias	RMSE	R^2	ΔAIC
Tunisia	C3	0.10	0.48	0.95	5.38	0.10	0.48	0.95	0.004
	C5	0.10	0.48	0.95	5.35	0.10	0.48	0.95	0.000
	C6	0.10	0.47	0.95	0.00	0.10	0.47	0.95	1.006
Spain	C3	0.04	0.35	0.97	0.58	0.04	0.35	0.97	0.063
	C5	0.04	0.35	0.97	0.00	0.04	0.35	0.97	0.000
Tunisia and Spain	C3	0.07	0.41	0.96	0.43	0.07	0.41	0.96	0.42
	C5	0.07	0.41	0.96	0.00	0.07	0.41	0.96	0.00
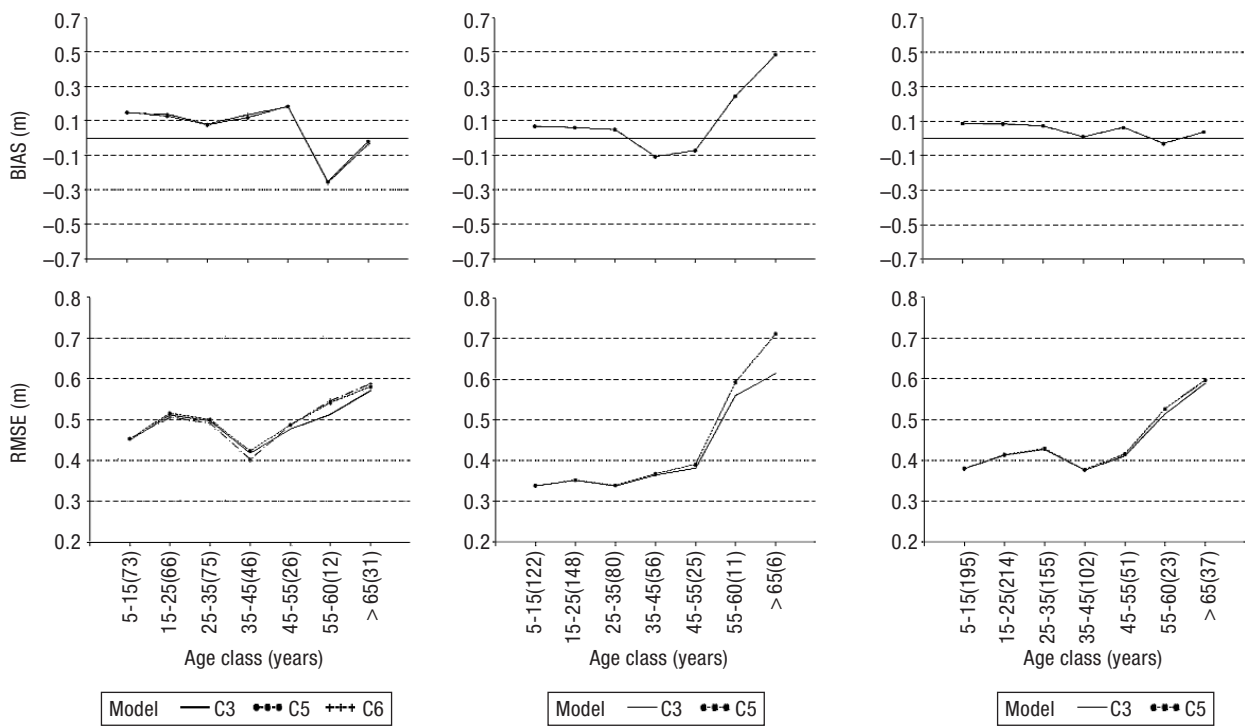
Figure 2. Bias and root mean square error (RMSE) in domiant height estimations by 10-year age classes for models with all parametrs significance at a 5% level in each country: Tunisia (a), Spain (b) and both countries (c).

ween the regions of each one, indicating that a single model could be used for both countries.

The analysis of the variability of the modeling efficiency and the relative error against age and against prediction interval for separate countries (full model) and combined countries (reduced model) are shown in Figure 5. The results confirm that a single dominant height growth model could be used for Tunisian and Spanish cork oak forests.

Discussion

In this study, base age invariant dominant height growth models for Tunisian and Spanish cork oak forests were developed. The methodology applied lead to models that complied with the two conditions required to be considered as base age invariant: (1) the models specifications are algebraic since they make use of dynamic equations derived by the ADA (Bailey and Clutter, 1974) and GADA approach (Cieszewski, 1994; Cieszewski and Bailey, 2000), and (2) the models parameters were estimated regardless of the choice of the base age by using the dummy variable approach (Cieszewski, 2004). Other base-age invariant growth models have been proposed for other Mediterranean oaks such as *Quercus ilex* (Gea-Izquierdo *et al.*, 2008) and *Quercus pyrenaica* (Adame *et al.*, 2006), as well as for modeling cork growth (Sánchez-González *et al.*, 2008).

Table 5. Comparison between countries and regions using the nonlinear sum of squares (F) and the Lakkis-Jones statistics (L)

| Region compared	Number of trees	SSEf	SSEr	F	P > F	L	P > L
Both countries	93	15.25	15.53	0.40	0.81	1.71	0.79
Tunisia	44	8.58	9.27	1.62	0.21	3.42	0.18
Spain	49	5.50	5.83	1.31	0.28	2.77	0.25

SSEf: error sum of squares for the full model. SSEr: error sum of squares for the reduced model. P > F: probability for F to be under a F distribution with dfr-dff degrees of freedom. P > L: probability for L to be under a χ^2 distribution with dfr-dff degrees of freedom.

C3

C5

C6
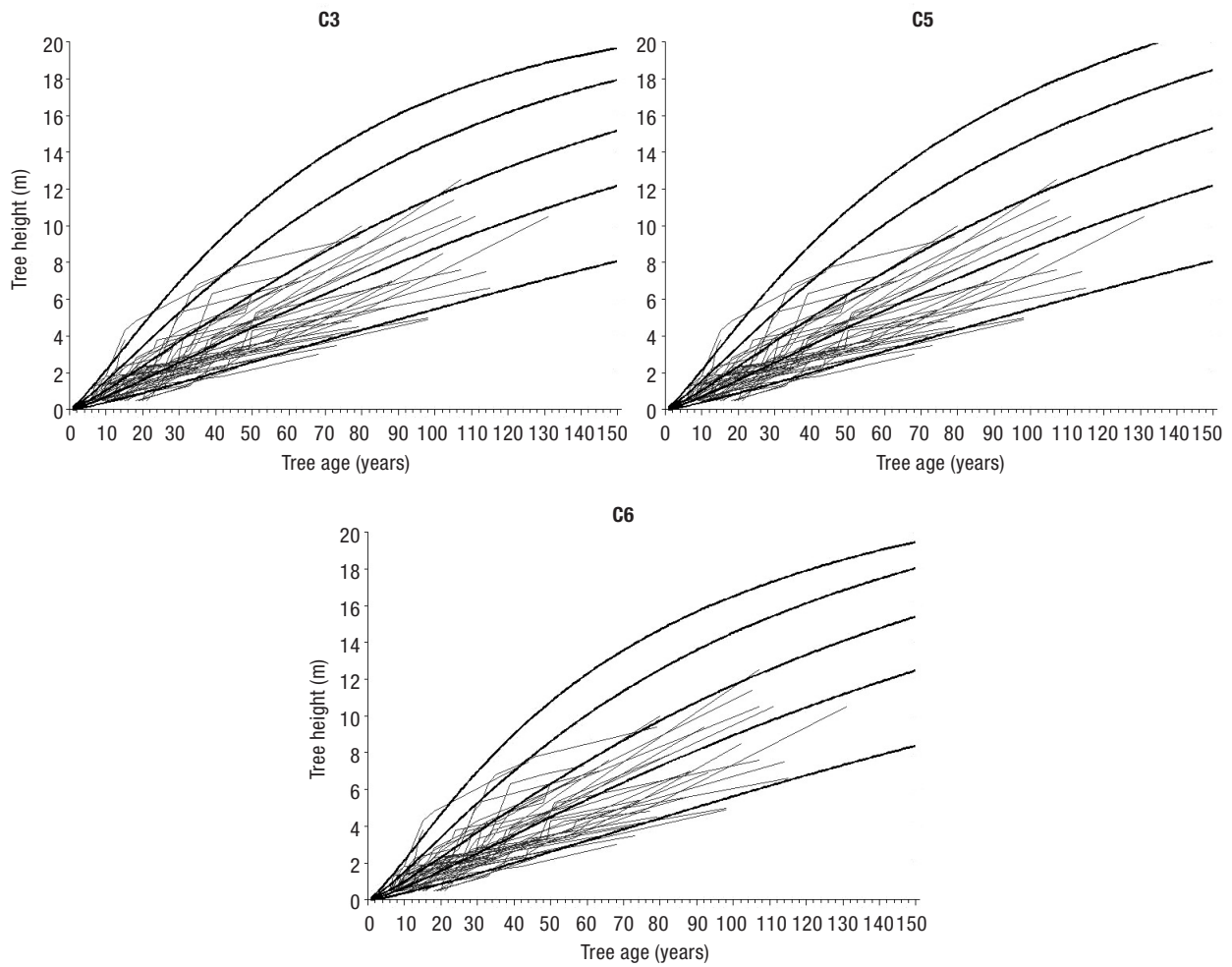
Figure 3. Fitted curves obtained using Tunisian data for site indices of 9, 7, 5, 3.5 and 2 m at a reference age of 40 years overlaid on the trayectories of the observed dominant heights over time for models C3, C5 and C6.

C3

C5
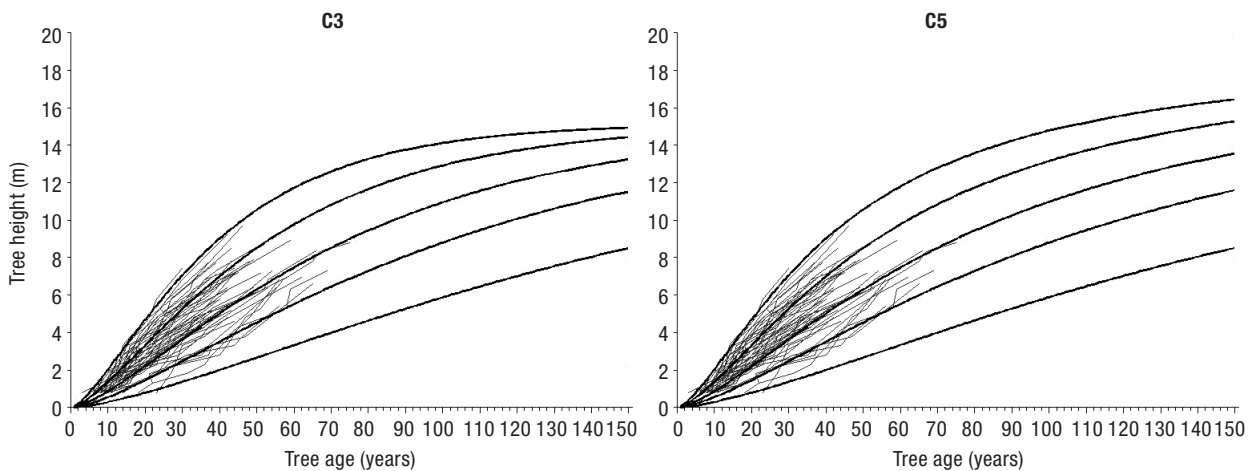
Figure 4. Fitted curves obtained using Spanish data for site indices of 9, 7, 5, 3.5 and 2 m at a reference age of 40 years overlaid on the trajectories of the observed dominant heights over time for models C3 and C5.
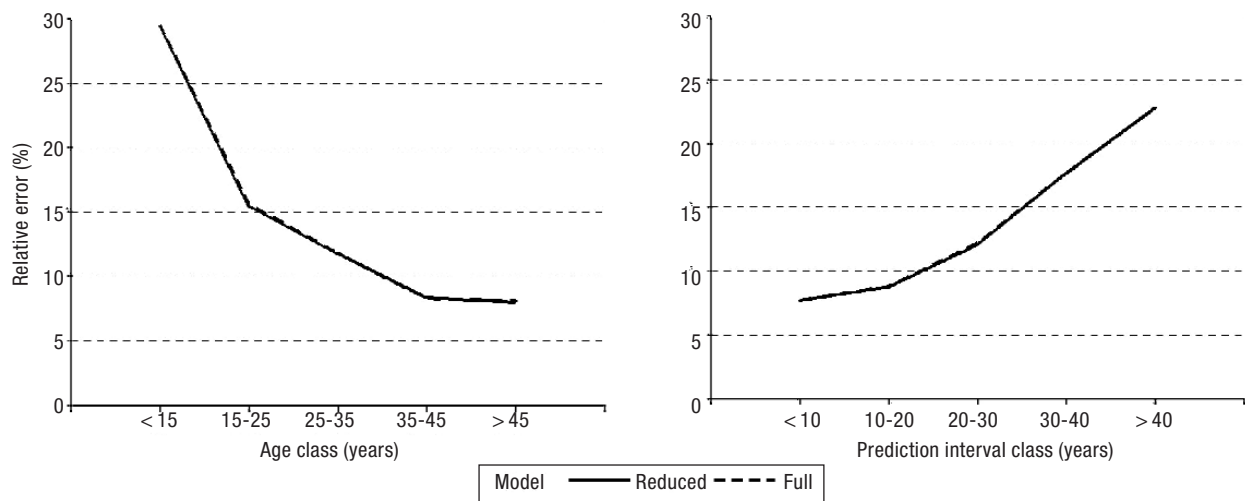
Figure 5. Relative error of dominant height predictins by age classes (above plot) and by prediction interval classes (below plot) for the full and reduced model using the selected model C5.

The simple polymorphic models (models C1, C3 and C5) described the data as well as advanced polymorphic ones (models C2, C4 and C6) (Table 4 and Fig. 3 and 4). A similar result was reported by Diéguez-Aranda *et al*. (2006) when developing a site index model for Loblolly pine plantations in USA. Although polymorphism and multiple asymptotes are two desirable characteristics of site equations (Cieszewski, 2002), the second one has importance when the behaviour of the curves is not suitable for the range of ages that would be used in practice, and if the single asymptote is achieved within the rotation period of the species. In cork oak stands, today's management plans usually consider rotation periods between 100 and 150 years (Pereira, 2007), which is considered the upper limit for quality cork production. As shown in Figures 3 and 4, the asymptotes are achieved after 150 years except for model C3 fitted to Spanish data, which was rejected. GADA models are more flexible and better described a wide variety of height-age trends, but in practice convergence of GADA models is hard to obtain as shown in several previous studies (Sánchez-González *et al.,* 2008). This could be due to the fact that GADA are more sophisticated model forms that probably require a more exhaustive description of growth processes by the data.

Among the seven dynamic equations evaluated for describing the dominant height growth in Spanish cork oak forests, the ADA formulation derived by McDill and Amateis (1992) from the Hossfeld model by considering the shape parameter as related to site productivity (C5) was selected. This model was found to be the best, taking into account both biological and statistical considerations. In a previous work, Sánchez-González *et al*. (2005) recommended the same equation for Spanish cork oak forests. Although the models were built using different approaches, their behavior is quite similar (Fig. 6b).

Figure 6a, shows the comparison between the dominant height growth curves obtained in this study and those obtained by Chaar *et al*. (2005) for Tunisian cork oak forests. Both models were plotted using a reference age of 40 years. The predictions of the model proposed in this study are more realistic than those of the model derived by Amaro *et al*. (1998) considering the asymptote as site-specific and proposed by Chaar *et al*. (2005), which is currently recommended for dominant height growth estimation in Tunisian cork oak forests. This anamorphic and static model was fitted only to trees from the Ain Snoussi region using a base-age specific method and does not describe the data adequately. By contrast, the dynamic model (C5) developed in the present study was fitted using data from both regions (Jouza and Ain Snoussi) and a base-age invariant method for parameter estimation, providing a good description of the trends in the data.

To compare the differences in dominant height growth models between the two countries, two tests were conducted based on the likelihood-ratio method, the non-linear extra sum of squares and the Lakkis-Jones test. This methodology has been proposed to evaluate the interregional variability in the dominant height growth pattern of other Mediterranean species such as *Pinus pinea* (Calama *et al*., 2003), *Pinus pinaster* (Álvarez
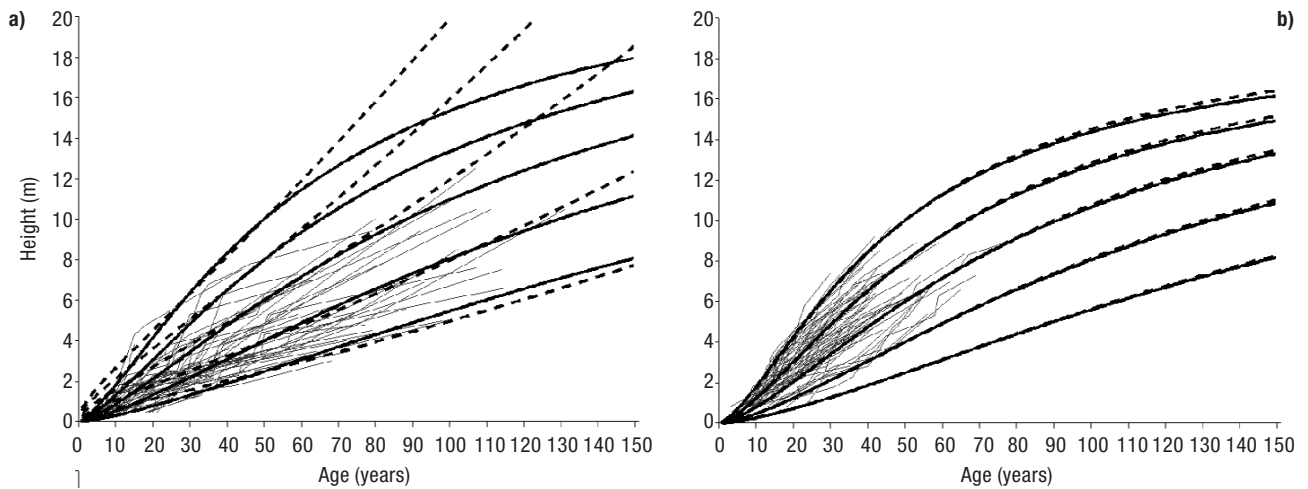
Figure 6. a) Comparison of the Tunisian curves for site indices of 9, 7, 5, 3.5 and 2 m at a reference age of 40 years for the static equation proposed by Chaar *et al.* (2005) —dashed lines—and the dynamic equation proposed by this paper (model C5) —solid lines—. b) Comparison of the Spanish curves for site indices of 9, 7, 5, 3.5 and 2 m at a reference age of 40 years for the equation proposed by Sánchez-González *et al.* (2005) —dashed lines— and the equation proposed by this paper (model C5) —solid lines—. Both plots overlaid on the trajectories of the observed dominant heights over time.

González *et al*., 2005; Bravo-Oviedo *et al*., 2007) and *Quercus pyrenaica* (Adame *et al*., 2006). The results obtained from the intra-country regional comparison (Table 5) point to the suitability of a single model for Spain, as previously reported by Sánchez-González *et al*. (2005) when developing a dominant height growth model for Spanish cork oak forests using a base-age specific method. The same result has been obtained for Tunisian regions. With regard to the inter-country comparison (Table 5), the results reveal that no significant differences exist between the two countries,
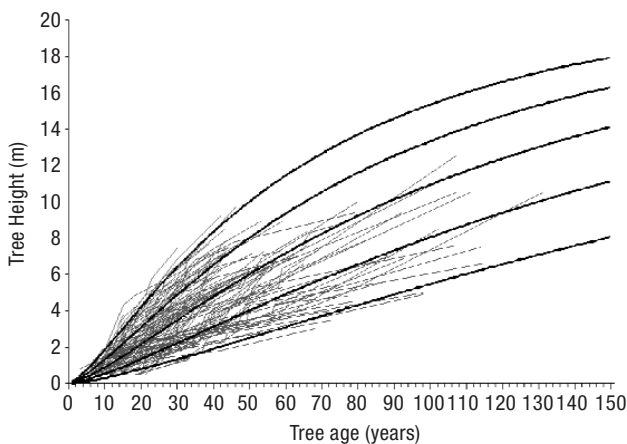
hence the use of a single model for Spain and Tunisia would appear to be appropriate. Figure 5 shows that there are no differences in the reliability of the predictions in the full and the reduced model, being above 85% for ages over 15 years, increasing to nearly 93 % for ages over 35 years.

The development of a single dominant height model for Spanish and Tunisian cork oak forests allow us to classify stands according to site quality. This tool is necessary to schedule the most suitable silvicultural treatments for these forests and forecast their consequences. Besides, this site index model is the first step towards elaborating a complete growth model for Tunisian cork oak forests such as ALCORNOQUE v1.0, an integrated growth and yield model at individual tree level for Spanish cork oak forests (Sánchez-González *et al*., 2007).

Based on the results presented above, the general form of the dominant height growth model proposed for Spanish and Tunisian cork oak forests is:

Figure 7. Site index curves generated by the dynamic equation C5 for site indices of 9, 7, 5, 3.5 and 2 m at a reference age of 40 years overlaid on the trajectories of the observed dominant heights over time from both countries.

$$H = \frac{23.49}{1-\left(1-\frac{23.49}{H_0}\right)\left(\frac{t_0}{t}\right)^{1.35}} \qquad [8]$$

where H_0 is the dominant height (m) at age t_0 (years) and H is the dominant height (m) at age t (years). The mean error made by this model when estimating total dominant height of trees not included in the fitting data set (trees with only two dominant height-age pairs),

was –0.38 m. It should be emphasized that the mean prediction interval was more than 80 years.

The graphical representation of equation [8] is shown in Figure 7. Site index can be calculated from the selected equation, by replacing t by the reference age (index age was taken as 40 years), and H_0 and t_0 by the observed dominant height and age in the subject stand. Site index is defined as the top dominant height reached at 40 years of age and then five quality classes are defined: 9 m for quality I, 7 m for quality II, 5 m for quality II, 3.5 m for quality IV and 2 m for quality V. The selection of base age was made taking into account the reliability in heights predictions from different base ages and their corresponding observed heights (Diéguez-Aranda *et al*., 2006). A base age of 40 was superior for predicting height at other ages with a minimum reliability (plot not shown). The scarcity of data at older ages advise against using a reference age older than 50 years.

Acknowledgements

References

ADAME P., CAÑELLAS I., ROIG S., DEL RÍO M., 2006. Modelling dominant height growth and site index curves for rebollo oak (*Quercus pyrenaica* Willd.) Ann For Sci 63, 929-940.

AKAIKE H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19, 716-723.

ÁLVAREZ-GONZÁLEZ J.G., RUIZ GONZÁLEZ A.D., RODRÍGUEZ-SOALLEIRO R., BARRIO-ANTA M., 2005. Ecoregional-based stand density management diagrams for *Pinus pinaster* L. in Galicia (northwest Spain). Ann For Sci 62, 115-127.

AMARO A., REED D., TOMÉ M., THEMIDO I., 1998. Modelling dominant height growth: eucalyptus plantations in Portugal. For Sci 44, 37-46.

BAILEY R.L., CLUTTER J.L., 1974. Base-age invariant polymorphic site curves. For Sci 20, 155-159.

BRAVO-OVIEDO A., DEL RÍO M., MONTERO G., 2007. Geographic variation and parameter assessment in generalized algebraic difference site index modelling. For Ecol Manage 247, 107-119.

BURNHAM K.P., ANDERSON D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. Springer-Verlag, New York. 488 pp.

CAÑELLAS I., SÁNCHEZ-GONZÁLEZ M., BOVINO S.M., ADAME P., HERRERO C., ROIG S., TOMÉ M., PAULO J.A., BRAVO F., 2008. Silviculture and carbon sequestration in Mediterranean oak forests. In: Managing forest ecosystems: the challenge of climate change (Bravo F., Jandl R., LeMay V., Von Gadow K., eds). Springer, Netherlands. pp. 317-338.

CALAMA R., CAÑADAS N., MONTERO G., 2003. Inter-regional variability in site index models for even-aged stands of stone pine (*Pinus pinea* L.) in Spain. Ann For Sci 60, 259-269.

CARMEAN W.H., 1972. Site index curves for upland oaks in the central states. For Sci 18, 109-120.

CHAAR H., STITI B., MONTERO G., 2005. Production and silviculture of cork oak forest in Ain Snoussi, Tunisia. Proceedings of International Congress Cork Plantations, Factories and Traders: the Past, Present and Future of the Cork Business. Palafrugell, Gerona (Spain).

CIESZEWSKI C.J., 2004. GADA derivation of dynamic site equations with polymorphism and variable asymptotes from Richards, Weibull, and other exponential functions. PMRC Technical Report 2004-5.

CIESZEWSKI C.J., 2003. Developing a well-behaved dynamic site equation using a modified Hossfeld IV function $Y^3 = (ax^m)/(c + x^{m-1})$, a simplified mixed-model and scant subalpine fir data. For Sci 49, 539-554.

CIESZEWSKI C.J., 2002. Comparing fixed and variable base-age site equations having single *versus* multiple asymptote. For Sci 48, 7-23.

CIESZEWSKI C.J., 2001. Three methods of deriving advanced dynamic site equations demonstrated on inland Douglas-fir site curves. Can J For Res 31, 165-173.

CIESZEWSKI C.J., 1994. Development of a variable density height-growth-model through defining multidimensional height growth spaces. PhD thesis. U of Alberta, Edmonton, Ab, Canada. 72 pp.

CIESZEWSKI C.J., BAILEY R.L., 2000. Generalized algebraic difference approach: theory based derivation of dynamic site equations with polymorphism and variables asymptotes. For Sci 46(1), 116-126.

CIESZEWSKI C.J., ZASADA M., STRUB M., 2006. Analysis of different base models and methods of site model derivation for Scots pine. For Sci 52(2), 187-197.

CIESZEWSKI C.J., HARRISON M., MARTIN S.W., 2000. Practical methods for estimating non-biased parameters in self-referencing growth and yield models Plantation Management Research Cooperative Technical Report 2000-7, DB Warnell School of Forest Resources University of Georgia.

CLUTTER J.L., FORTSON J.C., PEINAR L.V., BRISTER G.H., BAILEY R.L., 1983. Timber Management – A quantitative approach. John Wiley and Sons, New York. 333 pp.

DIÉGUEZ-ARANDA U., BURKHART H.E., AMATEIS R.L., 2006. Dynamic site model for loblolly pine (*Pinus*

taeda L.) plantations in the United States. For Sci 52(3), 262-272.

DIRECTION GENERALE DES FORETS, 1986. Plan d'aménagement: 1987-2011. Forêt domaniale de Amdoun. 2ᵉᵐᵉ série. Ministère de l'Agriculture.

DIRECTION GENERALE DES FORETS, 1983. Plan d'aménagement: 1984-2007.Forêt domaniale de Amdoun, 1ᵉʳ série. Ministère de l'Agriculture.

DRAPER N.R., SMITH H., 1981. Applied regression analysis. John Wiley & Sons, New York, 709 pp.

FABBIO G., FRATTEGIANI M., MANETTI M.C., 1994. Height estimation in stem analysis using second differences. For Sci 40, 329-340.

FANG Z., BAILEY R.L., 1998. Height-diameter models for tropical forests on Hainan island in southern China. For Ecol Manage 110, 315-327.

FAO, 2006. World Reference Base for Soil Resources World Soil Resources Reports 103, Rome. 145 pp.

GARCÍA O., 1983. A stochastic differential equation model for the height growth of forest stands. Biometrics 39, 1059-1072.

GEA-IZQUIERDO G., CAÑELLAS I., MONTERO G., 2008. Site index in agroforestry systems: age-dependent and age-independent dynamic diameter growth model for *Quercus ilex* in Iberian open oak woodlands. Can J For Res 38, 101-113.

GREGOIRE T., SCHANBENBERGER O., BARRETT J.P,. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can J For Res 25, 137-156.

HOSSFELD J.W., 1822. Mathematik für Forstmänner, Ökonomen und Cameralisten (Gotha, 4 Bd, S 310).

HUANG S., YANG Y., WANG Y., 2003. A critical look at procedures for validating growth and yield models in Modelling forest systems (Amaro A., Reed D., Soares P., eds). CAB International, Wallingford, Oxfordshire, UK. pp. 271-293.

KHATREE R., NAIK D.N., 1995. Applied Multivariate Statistics with SAS Software. SAS Institute Inc, Cary, NC. 588 pp.

KIVISTE A., 1998. Estimation of Estonian forest growth change in 1951-1994 on the basis of forest inventory data. In: Climate Change Studies in Estonia (Lallista T.T., Kuldna P., eds). Stockholm Environment Institute, Estonian Ministry of Environment, Tallinn, Estonia. pp. 191-196.

KORF V., 1939. Príspevek k matematické definici vzrustového zákona hmot lesních porostu. Lesnicka Prace 18, 339-379.

KRUMLAND B.E., ENG H., 2005. Site index systems for major young-growth forest and woodland species in northern California. Cal Dept Forestry and Fire Protection, Cal For Rep No 4.

LUNDQVIST B., 1957. On the height growth in cultivated stands of pine and spruce in Northern Sweden. Medd Fran Statens Skogfoesk Band 47, pp. 1-64.

MCDILL M.E., AMATEIS R.L., 1992. Measuring forest site quality using the parameters of a dimensionally compatible height growth function. For Sci 38, 409-429.

MONTERO G., CAÑELLAS I., 1999. Manual de forestación del alcornoque (*Quercus suber* L). MAPA-INIA. 103 pp.

MYERS R.H., 1990. Classical and modern regression with applications, 2ⁿᵈ ed. Duxbury Press, Belmont, CA. 488 pp.

PARRESOL B.R., 2001. Additivity of nonlinear biomass equations. Can J For Res 31, 865-878.

PEREIRA H., 2007. Cork: biology, production and uses. Elsevier. 336 pp.

PRETZSCH H., BIBER P., DURSKÝ J., GADOW K.V., HASENAUER H., KÄNDLER G., KENK G., KUBLIN E., NAGEL J., PUKKALA T., SKOVSGAARD J.P., SODTKE R., STERBA H., 2002. Recommendations for standardized documentation and further development of forest growth simulators. Forstw Cbl 121, 138-151.

RATKOWSKI D.A., 1983. Nonlinear regression modelling. A unified practical approach. Marcel Dekker, Inc, New York, Basel. 276 pp.

RICHARDS F.J., 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10, 290-300.

RYAN T.P., 1997. Modern regression methods. John Wiley and Sons, New York.

SÁNCHEZ-GONZÁLEZ M., CAÑELLAS I., MONTERO G., 2008. Base-age invariant cork growth model for Spanish cork oak (*Quercus suber* L.) forests. Eur J Forest Res 127, 173-182.

SÁNCHEZ-GONZÁLEZ M., CALAMA R., CAÑELLAS I. AND MONTERO G., 2007. Management oriented growth models for multifunctional Mediterranean forests: the case of the cork oak (*Quercus suber* L.) In: Scientific tools and research needs for multifunctional Mediterranean forest ecosystem managemen (Palahí M., Birot Y.,Rois M., eds). EFI proceedings 56, 71-84.

SÁNCHEZ-GONZÁLEZ M., TOMÉ M., MONTERO G., 2005. Modelling height and diameter growth of dominant cork oak trees in Spain. Ann For Sci 62, 633-643.

SAN MIGUEL A., ALLUÉ M., CAÑELLAS I., MONTERO G., BENGOA J., TORRES E., 1992. The most important forest ecosystems and their silvopastoral treatments in Spain. IUFRO Centennial, Berlin-Eberswalde (Germany). 11 pp.

SAS INSTITUTE INC, 2004. SAS/ETS 91 user's guide. SAS Institute Inc, Cary, NC. 1028 pp.

TOMÉ M., COELHO M.B., ALMEIDA A., LOPES F., 2001. O modelo SUBER. Estrutura e equações utilizadas, Relatórios técnico-científicos do GIMREF nº 2/2001, Centro de Estudos Florestais, Instituto Superior de Agronomia, Lisboa.

VANCLAY J.K., SKOVSGAARD J.P., 1997. Evaluating forest growth models. Ecol Model 98, 1-12.

WEST P.W., RATKOWSKY D.A., DAVIS A.W., 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. Forest Ecology and Management 7, 207-224.

ZIMMERMAN D.L., NÚÑEZ-ANTÓN V., 2001. Parametric modelling of growth curve data: an overview (with discussion). Test 10, 1-73.