

Modelling diameter distributions of birch (*Betula alba* L.) and pedunculate oak (*Quercus robur* L.) stands in northwest Spain with the beta distribution

J. J. Gorgoso-Varela^{1*}, A. Rojo-Alboreca², E. Afif-Khourri¹ and M. Barrio-Anta¹

¹ Departamento de Biología de Organismos y Sistemas. Universidad de Oviedo. Escuela Universitaria de Ingenierías Técnicas. C/ Gonzalo Gutiérrez Quirós s/n. 33600 Mieres (Asturias), Spain.

² Departamento de Enxeñaría Agroforestal. Universidade de Santiago de Compostela. Escola Politécnica Superior. Campus Universitario, 27002 Lugo (Galicia), Spain.

Resumen

Modelización de las distribuciones diamétricas en masas de abedul (*Betula alba* L.) y de roble (*Quercus robur* L.) en el noroeste de España con la distribución beta

Se han modelizado con la distribución beta las distribuciones diamétricas de 125 y 172 parcelas permanentes instaladas respectivamente en masas puras de abedul (*Betula alba* L.) y de roble pedunculado (*Quercus robur* L.) en Galicia (noroeste de España). Se ha empleado el método basado en los momentos de la distribución para ajustar el modelo a las distribuciones reales de frecuencias relativas de árboles, considerando los siguientes estadísticos en la comparación de resultados: sesgo, error medio absoluto, error medio cuadrático y número de parcelas rechazadas por el test de Kolmogorov-Smirnov (KS). En los ajustes de la función, el porcentaje de rechazos por el test KS fue del 0,8% del número total de casos en las masas de abedul y de un 1,2% en las masas de roble pedunculado, con un nivel de significación del 5%.

A continuación los parámetros de la distribución fueron recuperados con modelos de recuperación de parámetros (PRM) a partir del primero y del segundo momento de la distribución (diámetro medio y varianza, respectivamente). Los extremos de la función (diámetro mínimo y máximo de las distribuciones) fueron modelizados con modelos lineales sencillos, empleando variables de masa fáciles de obtener de las tablas de producción existentes para las especies en el noroeste de España (diámetro medio cuadrático, altura dominante, número de pies por hectárea, edad, área basimétrica e Índice de Hart-Becking). Los modelos recuperados y calibrados en masas de *Quercus robur* L. fueron más precisos en términos del error medio cuadrático (con un valor medio de 0,00055 en frecuencias relativas del número de árboles). Sin embargo, los modelos fueron más precisos en las masas de *Betula alba* L. en términos de rechazos por el test de Kolmogorov-Smirnov: 21 parcelas rechazadas en masas de abedul (16,8% del total de parcelas) y 33 parcelas en masas de roble (19,2% del total).

Palabras clave: Modelo de clases diamétricas, método de los momentos, modelo de recuperación de parámetros, Galicia.

Abstract

The diameter distribution of 125 and 172 permanent plots installed respectively in birch-dominated (*Betula alba* L.) and pedunculate oak-dominated (*Quercus robur* L.) stands in Galicia (northwest Spain) were modelled with the beta distribution. The method based on the moments of the distribution was used to fit the model to real distributions of relative frequencies of trees with the following statistics considered in the comparison of results: bias, mean absolute error (MAE), mean square error (MSE) and number of plots rejected by the Kolmogorov-Smirnov (KS) test. In the fits of the function, the percentage of rejections by the KS test was 0.8% of the total number of cases in birch stands and 1.2% in pedunculate oak stands, at a significance level of 5%.

Then, the parameters of the distribution were recovered with the parameter recovery models (PRM) from the first and the second moments of the distribution (mean diameter and variance, respectively). The extremes of the function (minimum and maximum diameters of the distributions) were modelled with simple linear models, with stand variables easily obtained from yield tables for the species in northwest Spain (quadratic mean diameter, dominant height, number of trees per hectare, age, basal area and the relative spacing index). The models recovered and calibrated in *Quercus robur* L. stands were most accurate in terms of MSE (with a mean value of 0.00055 in frequency of trees per one). However, the models were more accurate in *Betula alba* L. stands in terms of the number of rejections by the Kolmogorov-Smirnov test: 21 plots rejected in birch stands (16.8% of the total of plots) and 33 plots in pedunculate oak stands (19.2% of the total).

Keywords: Diameter class model, method of moments, parameter recovery model, Galicia.

* Corresponding author: gorgoso@uniovi.es

Received: 18-09-08. Accepted: 25-11-08.

1. Introduction

Birch (*Betula alba* L.) and pedunculate oak (*Quercus robur* L.) forests are valuable natural resources in north-west Spain. Birch stands currently cover 49,000 ha in northwest Spain, with 32,000 ha in Galicia (Xunta de Galicia, 2001) and 17,000 ha in the adjoining region of Asturias (INDUROT, 2001). Nowadays, birch is the fifth most abundant species in terms of number of trees in Galicia, after *Eucalyptus globulus*, *Pinus pinaster*, *Quercus robur* and *Quercus pyrenaica*.

Birch appears as the dominant species in these stands, which are mainly derived from natural regeneration, frequently on abandoned land, and occasionally from plantations. The total volume of Galician birch stands is nearly 3 million m³, which represents 2.3 % of the total wood produced in the region (Xunta de Galicia, 2001), with an annual harvest volume of 17,000 m³ in the period 1995-2001.

Stands in which pedunculate oak dominates cover 204,000 ha in NW Spain, with 188,000 ha in Galicia (Xunta de Galicia, 2001) and 16,000 ha in Asturias (INDUROT, 2001), and are mainly derived from natural regeneration. *Quercus robur* is the third most abundant tree species in Galicia (Xunta de Galicia, 2001), with an annual harvest volume of 74,500 m³ in the period 1995-2001.

Development of growth models for the species (Barrio-Anta and Álvarez González, 2005; Barrio-Anta and Diéguez-Aranda, 2005; Rojo *et al.* 2005; Diéguez-Aranda *et al.*, 2006; Gorgoso *et al.* 2007) enables promotion of the productive and protective aspects of both species in northwest Spain. Diameter class models allow planning of various uses and provide data about stand structure. These models are used to estimate stand variables and their structure with a density or distribution function, which is fitted to diameter distributions at breast height (1.3 m) or individual tree volume.

Forest managers are interested, for example, in being able to estimate the number of trees in different diameter classes in a stand, because the size of the diameter determines the industrial use of the wood and thus the price of the different products. Diameter distributions also provide information about stand structure, age structure, stand stability, etc. and enable the planning of silvicultural treatments. Furthermore, tree diameter is an important factor in harvesting because it determines the type of machines used and how they perform during felling and transportation of the wood.

The first mathematical description of a specific form of the diameter function in all-aged forest was provided

by DeLiocourt (1898). He found that plotting the number of stems against equal-diameter classes as a frequency histogram results in a reverse J-shaped curve. Field studies in virgin and old-growth forest have confirmed the utility of the negative exponential model (Meyer and Stevenson, 1943; Meyer, 1952; Lorimer, 1980; Leak, 1996), although there are occasionally small changes in the decreasing curve (Westphal, 2006). Diameter class models also have been applied to even-aged stands. In this case the frequency histogram in a diametric range is similar to the Gauss distribution but with a different shape because of the skewness and the kurtosis of the corresponding curve.

Various distribution functions have been used to describe and predict stem frequency in even-aged and uneven-aged stands, such as the normal (Bliss and Reinker, 1964), gamma (Nelson, 1964), Gram-Charlier (Schnur, 1934), Johnson's S_B (Johnson and Kitchen, 1971; Knoebel and Burkhardt, 1991; Kamziah *et al.*, 1999) and Weibull function (Zhang *et al.*, 2003; Liu *et al.*, 2004). The popularity of the latter is based on its relative simplicity and flexibility (Bailey and Dell, 1973).

The beta density function was used by Zöhrer (1969 and 1970) due it had the considerable advantage of being simple, highly adaptable and practically any form of diameter distribution can be described, and also was used by authors as Lenhart and Clutter (1971), Loetsch *et al.* (1973) and Maltamo *et al.* (1995). Furthermore, this density function was used in the first models in which the function parameters were related to stand variables, in *Pinus elliottii* Engelm stands in Georgia (Clutter and Bennett, 1965).

Li *et al.* (2002) modelled the bivariate generalized beta distribution (GBD-2), which was more flexible and accurate than the bivariate Johnson's S_{BB} distribution for modelling diameters and heights in *Pseudotsuga menziesii* stands in Idaho, United States. Wang and Rennolls (2005) also found the beta function to be more accurate than the Johnson's S_B distribution in *Cunninghamia lanceolata* stands and concluded that this function can cover a wide range of skewness and kurtosis.

In Spain the goodness of fit of the beta distribution was compared with other distributions by Álvarez-González (1997) in *Pinus pinaster* Ait. stands in Galicia (northwest Spain). Gorgoso (2003) compared the accuracy of the beta, Weibull and Johnson's S_B functions fitted to diameter distributions of *Betula alba* L. stands in this region. Palahí *et al.* (2007) compared the beta function with other distributions for fitting diameter distributions in forests stands in Catalonia (northeast Spain).

Two different methods have been developed for modelling parameters with stand variables (Hyink and Moser, 1983): Parameter prediction models (PPM) and parameter recovery models (PRM). Parameter prediction models (Torres-Rojo *et al.*, 2000; Cao, 2004) relate the parameters of a distribution function and stand variables by use of simple linear models. Parameter recovery models (PRM) were developed to improve the estimates obtained with the PPM models, and relate stand variables to percentiles (Bailey *et al.*, 1982; Cao and Burkhart, 1984) or moments (Lindsay *et al.*, 1996; Ríó, 1999) of the diameter distribution.

The purpose of this study was to calibrate a general diameter distribution model based on a beta distribution and a parameter recovery approach by application to 125 *Betula alba* L. and 172 *Quercus robur* L. stands. The parameters of the function were recovered from the first and the second moments of the distribution (mean diameter and variance, respectively) and the extreme superior and inferior values of the distributions were predicted with stand variables that are easy to obtain from existing yield tables for the species in northwest Spain (Barrio-Anta, 2003; Rojo *et al.*, 2005).

Materials and methods

Data set and processing

The data used to develop the diameter distribution models were obtained from a total of 125 research plots in even-aged *Betula alba* L. stands and 172 in even-aged *Quercus robur* L. stands throughout Galicia (northwest Spain). The size of the plots ranged from 200 m² to 1,000 m² in birch stands and from 225 m² to 1,345 m² in pedunculate oak stands (depending on the stand density), in order to achieve a minimum of 30 trees per plot. The plots were established in dominant stands (more than 85% of species standing basal area), and were established to cover a wide variety of combinations of age, number of trees per hectare and site.

All trees in each plot were numbered. Two perpendicular diameters at breast height were measured with callipers, to the nearest 0.1 cm, and the arithmetic average was calculated. The empirical data represent left-truncated distributions, as the smallest diameter measured in the field was 5 cm (trees smaller than this limit

were not considered). A total of 10,386 diameter measurements in birch stands and 10,248 in oak stands were available for analysis. Furthermore, the heights of a random sample of 30 trees in each plot were measured and the mean height calculated. Moreover, dominant trees were measured in each plot and dominant height was calculated from the percentage of the 100 thickest trees per ha (Assmann, 1970). Age was estimated (only in birch stands) by counting growth rings in samples extracted from some stems.

The following stand variables were calculated from the inventory data: quadratic mean diameter, mean diameter, minimum diameter, maximum diameter, number of trees per hectare, basal area, dominant height and the relative spacing index. Age was only estimated in birch stands. Summary statistics including mean, maximum and minimum values and standard deviation of the main stand variables are shown in Table 1.

Fitting methods

The general expression of the beta distribution for a random variable x is the following:

$$f(x) = c \cdot (x - L)^{\alpha} \cdot (U - x)^{\gamma} \quad L \leq x \leq U \quad (1)$$

$$x = 0, \text{ elsewhere}$$

where x is the diameter at breast height and it is assumed to be continuous, $f(x)$ is the density associated with diameter x , U is the upper limit of the beta distribution, L the lower limit of the beta distribution, c is the scaling factor of the function, and α and γ are, respectively, the first and the second exponents that determine the shape of the distribution. The diameter classes were considered of size equal to 1 cm.

Firstly, the parameters α , γ and c were estimated for both species by fitting the diameter distribution in each plot. The upper limit U was considered as the maximum diameter inventoried in each plot, and the lower limit L was equal to the minimum diameter of the plot. Zöhrer (1969, 1970) and Loetsch *et al.* (1973) proposed similar values using truncated distributions.

Fitting method based on the moments of the distribution

The method of fit was firstly used by Loetsch *et al.* (1973) and more recently by other authors such as Mal-tamo *et al.* (1995) and Palahí *et al.* (2007). Parameters of the function are obtained from first and second moments of the distributions, mean diameter (\bar{d}) and variance (s^2):

$$\gamma = \frac{\frac{Z}{s_{rel}^2 \cdot (Z+1)^2} - 1}{Z+1} - 1 \quad (2); \quad \alpha = Z \cdot (\gamma + 1) - 1 \quad (3)$$

where:

$$Z = \frac{x_{rel}}{1 - x_{rel}} \quad (4); \quad x_{rel} = \frac{\bar{d} - L}{U - L} \quad (5); \quad s_{rel}^2 = \frac{s^2}{(U - L)^2} \quad (6);$$

and U and L equal the maximum and minimum diameters inventoried.

The parameter c was estimated as:

$$c = \frac{1}{\int_L^U (x - L)^\alpha \cdot (U - x)^\gamma dx} \quad (7)$$

If the integral is resolved:

$$c = \frac{1}{(-L + U)^{1+\gamma} \cdot \Gamma(1 + \alpha) \cdot \Gamma(1 + \gamma) \left(\frac{1}{-L + U}\right)^\alpha \cdot \Gamma(2 + \alpha + \gamma)} \quad (8)$$

where $\Gamma(i)$ is the Gamma function in the point i .

Method comparison

The fitting method consistency was evaluated by the bias, mean absolute error (MAE), and mean square error (MSE), with the following expressions:

$$Bias = \frac{\sum_{i=1}^N Y_i - \hat{Y}_i}{N} \quad (9); \quad MAE = \frac{\sum_{i=1}^N |Y_i - \hat{Y}_i|}{N} \quad (10);$$

$$MSE = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} \quad (11)$$

where Y_i is the observed value, \hat{Y}_i is the theoretical value predicted by the model and N is the number of data points.

Mean square error (MSE), mean absolute error (MAE) and bias were calculated for each fit in mean relative frequency of trees per one for all diameter classes and plots.

Each fit was also tested with the Kolmogorov-Smirnov (KS) test. This test compares the cumulative estimated frequency with the observed frequency. The most

striking difference between the two distributions was the D_n value of the KS test:

$$D_n = \max |F(x_i) - F_0(x_i)| \quad \forall 0 \leq i \leq n \quad (12)$$

where $F(x_i)$ is the cumulative frequency distribution observed for the sample x_i ($i = 1, 2, \dots, n$) and $F_0(x)$ the probability of the theoretical cumulative frequency distribution. Diameter classes of 1 cm were chosen. The value obtained was compared with the approximation proposed by Miller (1956), and the significance level was established at 5% with the following expression:

$$D_{n,\alpha} = \sqrt{\frac{-Ln\left(\frac{1}{2} \cdot \alpha\right)}{2 \cdot n}} \quad (13)$$

where Ln is the natural logarithm, α the significance level and n the number of trees inventoried in each plot.

Modelling distribution parameters with stand variables

Parameters fitted by the method based on the moments of the distribution were recovered with parameter recovery models (PRM). With this method the mean diameter and the variance must be known (first and second order moments of the distribution, respectively). The variance was estimated with the following expression:

$$\text{var}(d - \bar{d}) = d_g^2 - \bar{d}^2 \quad (14)$$

The mean diameter (\bar{d}) was related to the quadratic mean diameter (d_g), disaggregated from yield tables for birch stands (Rojo *et al.*, 2005) and pedunculate oak stands (Barrio-Anta, 2003) in northwest Spain, and other stand variables that can be obtained from yield tables or from the stand density management diagram for even-aged pedunculate oak stands (Barrio-Anta and Álvarez-González, 2005). Models in which mean diameter estimated is always smaller than quadratic mean diameter were used:

$$\bar{d} = d_g - e^{x\beta} \quad (15)$$

where x is the vector of independent stand variables in a fixed instant and β is the vector of parameters to be estimated.

The upper limit U and the lower limit L of the distributions were modelled with simple linear models that relate the minimum and maximum diameters with stand variables that are easy to obtain from the yield tables of the species. These variables are quadratic mean diameter (d_g), dominant height (H_0), basal area (G), age (t) (only in birch stands), number of trees per hectare (N) and the relative spacing index (RS).

Stepwise regression was done in the linear models to include the independent variables. The accuracy of the models was tested with the adjusted determination coefficient (R^2_{adj}):

$$R^2_{adj} = 1 - \frac{(N-1) \sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{(N-p) \sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (16)$$

where Y_i , \hat{Y}_i and Y , and are, respectively, the observed, predicted and mean values of the dependent variable; N is the number of the data points in the fit and p is the number of parameters to be estimated.

Colinearity of the independent variables was considered and in all cases and the variance inflation factor (VIF) was lower than 10 (Ryan, 1997). Finally, the general diameter distribution model was calibrated for each plot.

Results

Fits of the function

The mean values of bias, mean absolute error (MAE) and mean square error (MSE) in relative frequency of number of trees per one, and the number of plots rejected by the Kolmogorov-Smirnov (KS) test in the fits are shown in Table 2.

The lowest value of MSE was obtained with pedunculate oak stands but the MAE and the bias were higher than for birch stands. The rejections by the KS test were similar in the fits for both species, and were rejected 0.8% of the total of plots in birch stands and 1,2 % in pedunculate oak stands, at a significance level of 5%.

The changes in the mean value of bias and MSE in each diameter class for both species in the fits are shown in Figure 1 (a and b).

The values of bias were close to zero, and a similar mean value was obtained for both species (0.00057 for oak stands and 0.00046 for birch stands). This implies a

tendency towards underestimation of the real distributions. The mean values of MSE for each diameter classes are shown in Figure 1b. The MSE values in the lowest diameter class in birch stands were high, with the maximum in the diameter class equal to 15.5 cm because the number of data was greater for this species in this diameter range (see Figure 2). In the largest diameter class the value of MSE was higher in pedunculate oak stands. This is important because of the greater economic value of the products in the largest diameter class.

Modelling the parameters of the function with stand variables

Mean diameter (\bar{d}) was modelled with stand variables for both species with the following expressions:

$$\textit{Betula alba L.: } \bar{d} = d_g - e^{(-2.125+0.066 \cdot H_0+0.021 \cdot t)}; \quad R^2 = 0.941 \quad (17)$$

$$\textit{Quercus robur L.: } \bar{d} = d_g - e^{(-0.3655+0.0296 \cdot H_0)}; \quad R^2 = 0.990 \quad (18)$$

where d_g is the quadratic mean diameter in cm, H_0 is the dominant height in m and t is the age in years. Number of trees per hectare also was included in the models but it was not significant at a level of 5%.

The variance was calculated with equation [14]. The lower limit L and the upper limit U of the distributions were modelled with the following expressions:

Betula alba L.:

$$L = -4.153 + 0.597 \cdot d_g + 0.101 \cdot RS + 0.001 \cdot N - 0.074 \cdot G; \quad (19)$$

$$R^2 = 0.509; \quad \text{VIF} = 5.12$$

$$U = -1.106 + 0.999 \cdot d_g + 0.170 \cdot t + 0.385 \cdot H_0; \quad (20)$$

$$R^2 = 0.668; \quad \text{VIF} = 2.60$$

Quercus robur L.:

$$L = 1.236 + 0.330 \cdot d_g; \quad R^2 = 0.347 \quad (21)$$

$$U = 10.280 + 1.263 \cdot d_g; \quad R^2 = 0.619 \quad (22)$$

where d_g is the quadratic mean diameter in cm, H_0 is the dominant height in m, t is the age in years, G is the basal area in $\text{m}^2 \cdot \text{ha}^{-1}$, N is the density in $\text{trees} \cdot \text{ha}^{-1}$ and RS is the relative spacing index.

The results obtained for the total mean values of bias, MAE, MSE and number of rejections by the KS test after modelling the parameters of the beta distributions recovering the parameters with the method of the moments are shown in Table 3. The results are worse than in the fits of the function to real distributions in the two species, and the mean values of bias, MAE and MSE were higher. The number of rejections by the KS

Table 1. Summary of main stand variables for the samples used for modelling

	Variable	Mean	Maximum	Minimum	Standard deviation
<i>Betula alba</i> L. (N=125)	Quadratic mean diameter (cm)	14.1	23.2	7.4	3.5
	Mean diameter (cm)	13.5	22.5	7.1	3.5
	Minimum diameter (cm)	6.4	14.0	5.0	2.0
	Maximum diameter (cm)	23.9	48.2	11.9	6.9
	Density (trees·ha ⁻¹)	1793.3	6000.0	390.0	1134.3
	Age (years)	30.1	60.0	12.0	9.4
	Dominant height (m)	15.0	23.8	7.2	3.6
	Basal area (m ² ·ha ⁻¹)	24.6	66.5	3.3	10.7
	Relative spacing index	19.3	59.2	8.3	8.7
<i>Quercus robur</i> L. (N=172)	Quadratic mean diameter (cm)	21.9	40.0	8.6	6.1
	Mean diameter (cm)	20.7	37.8	8.0	6.0
	Minimum diameter (cm)	8.5	25.5	5.5	3.4
	Maximum diameter (cm)	37.9	71.5	16.5	9.8
	Density (trees·ha ⁻¹)	873.9	3,022.2	302.2	481.6
	Dominant height (m)	17.0	25.6	7.2	3.3
	Basal area (m ² ·ha ⁻¹)	28.3	72.9	3.4	9.2
	Relative spacing index	22.2	50.4	12.7	5.3

test increased from 1 plot (0.8%) to 21 plots (16.8%) in birch stands and from 2 plots (1.2%) to 33 plots (19.2%) in pedunculate oak stands.

The mean values of bias and the MSE in each diameter class when the parameters of the beta distribution were recovered with the method of the moments are shown in Figure 3 (a and b). Values of bias were

always close to zero, except in the smallest diameter class in which the real frequencies were underestimated. The MSE decreased in the largest diameter classes in both species. The observed distributions, the beta distribution fitted by the moments approach and the beta distribution modelled with stand variables in some plots of *Betula alba* and *Quercus robur* are shown in Figure 4.

Table 2. Mean values of bias, mean absolute error, mean square error in number of trees per one and number of plots rejected by the KS test ($\alpha = 0.05$) for the fits with the beta distribution

	Bias	MAE	MSE	KS	%
<i>Betula alba</i>	0.00046	0.01372	0.00047	1	0.8
<i>Quercus robur</i>	0.00057	0.01403	0.00042	2	1.2

Discussion

The modelling data used in the study reflect the complexity and heterogeneity of birch and pedunculate oak stands in northwest Spain. All plots contained enough trees to enable describe and predict the diameter distri-

Table 3. Mean values of bias, mean absolute error, mean square error in number of trees per one and number of plots rejected by the KS test ($\alpha = 0.05$) for the modelling of the parameters of the beta distribution with the method of the moments

	Bias	MAE	MSE	KS	%
<i>Betula alba</i> L.	0.00216	0.01527	0.00065	21	16.8
<i>Quercus robur</i> L.	0.00261	0.01495	0.00055	33	19.2

butions by the size of the plots and the number of trees per hectare of the stands.

Pedunculate oak stands derived from natural regeneration and thinnings were not carried out. Nevertheless only a few structures of diameter distributions have the typical J-shaped curve corresponding to natural forest, because the species is shade intolerant in the region (Ruíz de la Torre, 1979). Therefore, even-aged stands with a diameter distribution similar to a Gauss distribution are more abundant.

Birch stands are mainly derived by natural regeneration (66% of plots) and in some cases by planting (33% of plots); this species is also considered as shade intolerant in the area (Rothmaler and De Carvalho e Vasconcellos, 1940), so the stands are mainly even-aged. Low thinnings have been applied in some cases so there are also a few stands with diameter distributions that follow a J-shaped curve.

One method of fitting the beta distribution was used. Fewer methods of fit have been described for this function than for Weibull or Johnson's SB distributions. The lowest value of MSE in the fits was obtained in pedunculate oak stands. This statistic must be considered more closely than the bias because large errors with a different sign can be compensated for (overestimations

and underestimations of the real distributions). In both species the MSE was higher in the diameter class with the largest number of data. The low number of rejections in the fits by the KS test indicates that the beta distribution is a suitable model for describing the diameter structure of both species.

As regards the modelling of the parameters, the models that predicts the lower and the upper limits of the distributions are more accurate in birch stands (equations [19 and 20]), because they include a larger number of variables (d_g, N, G, RS, H_0) and because the range of the data and the variability is lower. In pedunculate oak stands the stepwise regression only included one independent variable: the quadratic mean diameter, and the adjusted determination coefficient (R^2_{adj}) was smaller than in birch stands (equations [21 and 22]). The poorest estimation was parameter L in *Quercus robur* stands because $R^2 = 0.347$.

Maltamo *et al.* (1995) modelled parameters L and U of the beta distribution in *Pinus sylvestris* and *Picea abies* stands in the centre and east of Finland. In *Pinus sylvestris* stands parameter U was related to stand variables with a model that explained 79% of the variability. For *Picea abies* stands the model provided an R^2 value of 0.52 for modelling of this parameter. For modelling parameter L a linear model was obtained with an adjusted determination coefficient of 0.79 in *Pinus sylvestris* stands and 0.83 for *Picea abies*.

The decrease in MSE in the largest diameter classes was due to the smaller number of trees (in some diameter class there were no trees). In many cases the upper limit of the distribution U modelled with stand variables was lower than the real limit of the distribution, and the model therefore predicted zero values (see Figure 4).

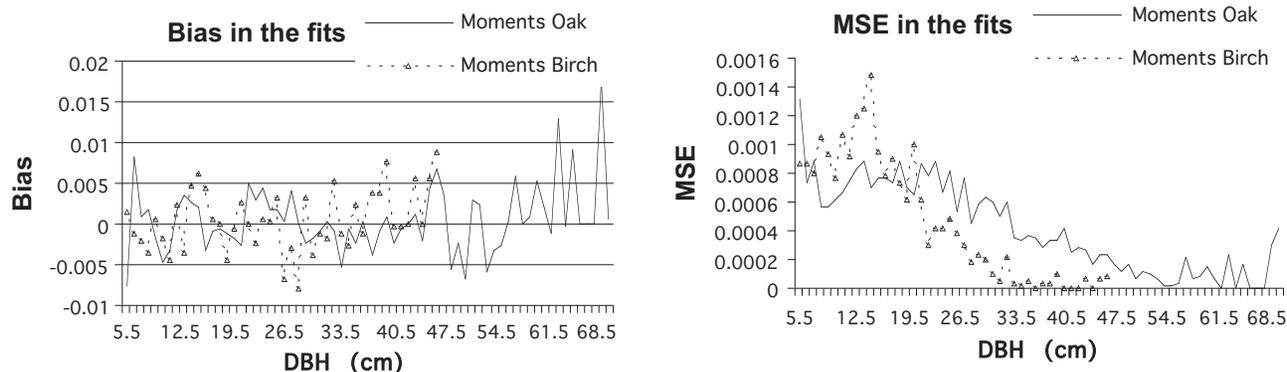


Figure 1 (a and b). Mean values of bias and MSE in number of trees per one in each diameter class, obtained in the fits of the beta distribution.

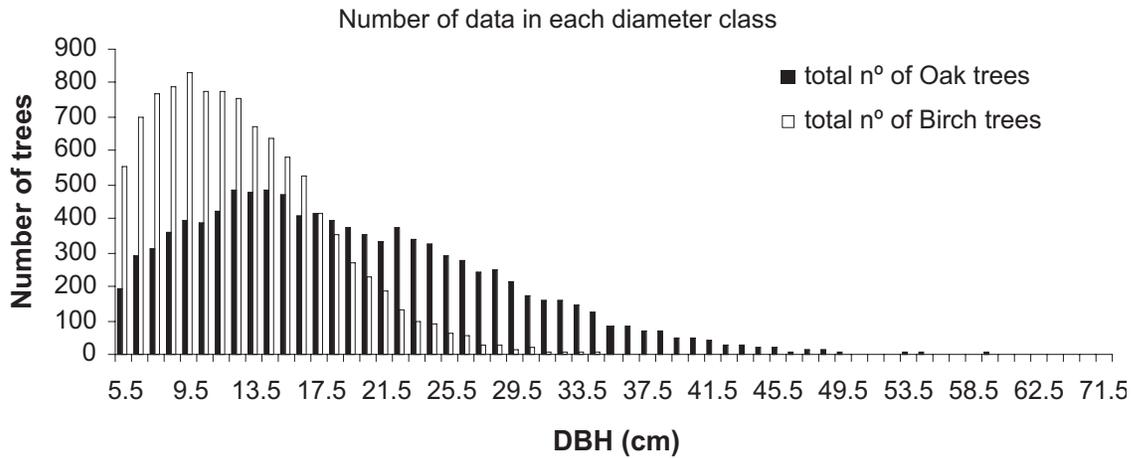


Figure 2. Total number of trees in the sample of *Quercus robur* L. and *Betula alba* L. in each diameter class.

This is more appropriate in stands in which there are a small proportion of trees of other species in the largest diameters of the stands, as with a small percentage of birch stands in which a few specimens of *Quercus robur* and *Castanea sativa* were included in the largest diameter classes.

Predictions for the smallest diameter class were not very accurate, and large underestimates of the real distributions were obtained. This was probably because the lower limits of the distributions are difficult to model. In some stands, better results would perhaps be obtained if the lower limits of the distributions L were fixed as a fraction of the minimum diameter or zero. However, Zöhner (1969, 1970) and Loetsch *et al.* (1973) fixed L in truncated distributions as the lower limit of the smallest diameter class, (for example 7.5 cm in a plot in

Norway spruce and 11 cm in a stand in *Pinus radiata*). In this works, the optimum values for the parameter U were optimized, with or without extensions of the upper limit of the biggest diameter class. Li *et al.* (2002) show diameter distributions fitted with the bivariate generalized beta distribution in *Pseudotsuga menziesii* stands without extensions of the biggest diameter class.

The two-parameter Weibull function (where the situation parameter a equals zero) was used to model the same diameter distributions of *Betula alba* L. stands in northwest Spain (Gorgoso *et al.*, 2007) and the parameters were recovered with the method of the moments. The mean value of the MSE in number of trees per one was 0.0041 utilizing only diameters greater than the minimum diameter inventoried to scale the frequencies, which is higher than the beta distribution in both species, and the

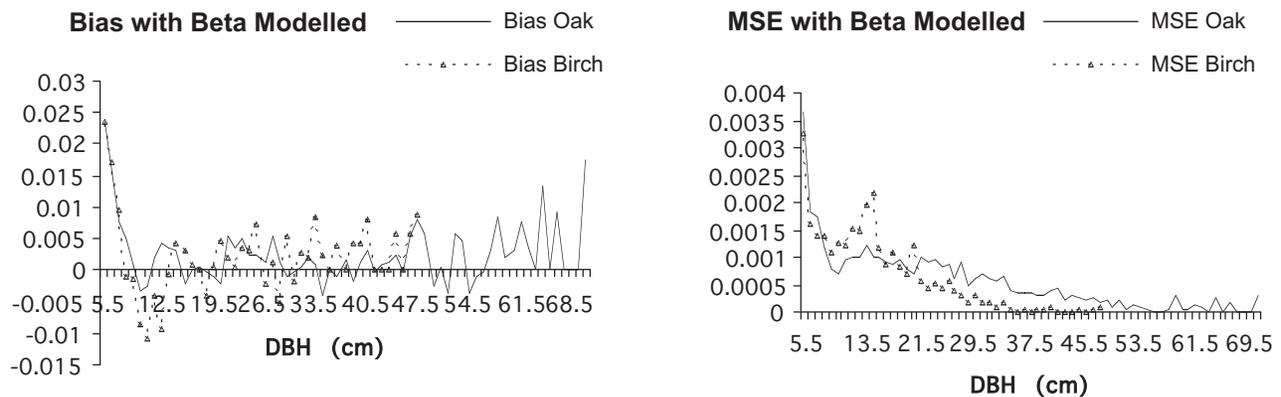


Figure 3 (a and b). Mean values of bias and MSE in number of trees per one in each diameter class obtained with the parameter recovery models.

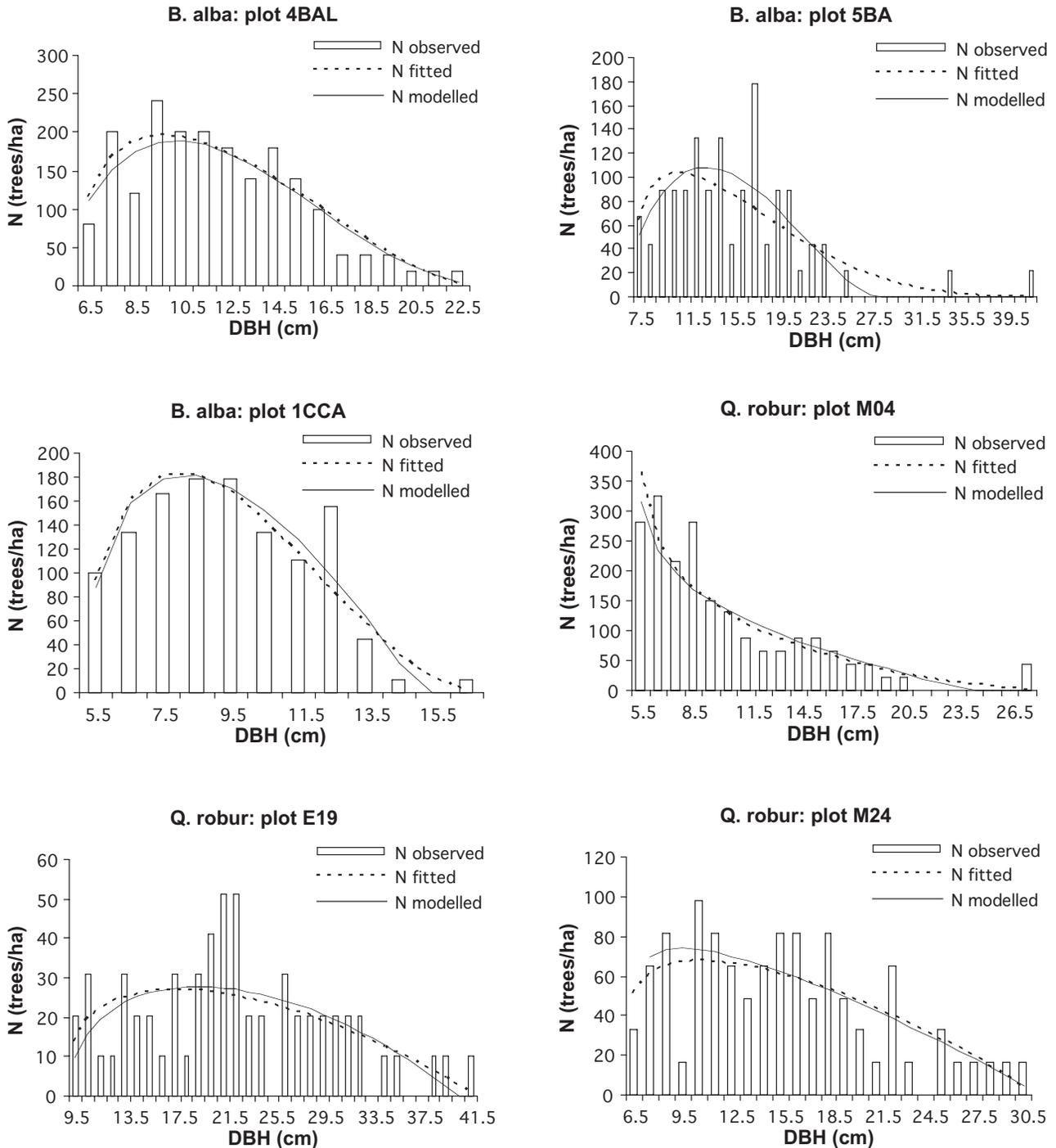


Figure 4. Observed diameter distributions, fitted beta distribution by the moments approach and modelled beta distribution with stand variables in three plots of *Betula alba* stands and three plots of *Quercus robur* stands.

number of rejections by the KS test was 48 plots at significant level of 20% (with this significance level the rejections with the beta distribution would be 29 plots in birch stands). Palahí *et al.* (2007) obtained better goodness of

fit with the truncated Weibull than the beta distribution fitted by the method of the moments in forest stands in northeast Spain. Nevertheless they did not recover the parameters of the beta function.

Li *et al.* (2002) compared the bivariate generalized beta distribution with the bivariate Johnson's S_{BB} distribution for simultaneous modelling of diameter and height distributions of *Pseudotsuga menziesii* in Idaho. The bivariate beta function was more accurate than the bivariate Johnson's S_{BB} distribution. In this study, beta and Johnson's S_B distributions were fitted and the goodness of the fits was tested with the χ^2 test, which in the beta distribution rejected 11.6% of the plots (5 of a total of 43) at a significance level of 5%. The same test rejected a 23.3 % of the plots with the Johnson's S_B distribution.

Conclusions

The method based on the moments of the beta function is suitable for fitting diameter distributions of birch (*Betula alba* L.) and pedunculate oak (*Quercus robur* L.) stands in northwest Spain. The extremes of the distributions were modelled more accurately with stand variables in birch stands. For calibrating the general model, the accuracy was less than in the fits of the function to the real diameter distributions of both species. The results of the calibrating were better in terms of MSE for *Quercus robur* stands, although there were fewer rejections by the KS test for *Betula alba* stands.

Acknowledgements

The present study was financially supported by the Asociación Galega Monte-Industria, project 2000/CG488: *Elaboración de modelos de crecimiento y producción para la gestión sostenible de las masas de abedul*; and by the Xunta de Galicia, project PGIDT99MA29101: *Estudio epidométrico de las masas de Quercus robur L. en Galicia y su influencia sobre la calidad de la madera*.

References

- ÁLVAREZ-GONZÁLEZ J.G., 1997. Análisis y caracterización de las distribuciones diamétricas de *Pinus pinaster* Ait. en Galicia, Ph. D. thesis, Univ. Politécnica de Madrid (Unpublished), 269 pp.
- ASSMANN E., 1970. The Principles of Forest Yield Study. Pergamon Press, Oxford, New York, 506 pp.
- BAILEY R.L., DELL T.R., 1973. Quantifying Diameter Distributions with the Weibull Function, For. Sci. 19: 97-104.
- BAILEY R.L., ABERNETHY N.C., JONES E.P., 1982. Diameter distribution models for repeatedly thinned slash pine plantations. USDA For. Ser. Gen. Tech. Re. SO-34.
- BARRIO-ANTA M., 2003. Crecimiento y producción de las masas naturales de *Quercus robur* L. en Galicia, Ph. D thesis, Univ. Santiago de Compostela (Unpublished), 252 pp.
- BARRIO-ANTA M., ÁLVAREZ-GONZÁLEZ J.G., 2005. Development of a stand density management diagram for even-aged pedunculate oak stands and its use in designing thinning schedules. Forestry, 78(3): 209-216.
- BARRIO-ANTA M., DIÉGUEZ-ARANDA U., 2005. Site quality of pedunculate oak (*Quercus Robur* L.) stands in Galicia (northwest Spain). Eur. J. Forest Res. 124: 19-28.
- BLISS C.I., REINKER K.A., 1964. A log-normal approach to diameter distributions in even-aged stands. Forest Sci. 10: 350-360.
- CAO Q.V., 2004. Predicting parameters of a Weibull function for modelling diameter distributions, For. Sci. 50: 682-685.
- CAO Q.V., BURKHART H.E., 1984. A segmented distribution approach for modeling diameter frequency data, For. Sci. 30: 129-137.
- CLUTTER J.L., BENNETT F.A., 1965. Diameter distributions in old-field slash pine plantations. Ga. For. Res. Council. Rep. 13, 9 pp.
- DeLIOUCOURT F., 1898. De l'aménagement des Sapiniers. Bul. Soc. For. Franche-Comté et Belfort 4, 396-409.
- DIÉGUEZ-ARANDA U., GRANDAS-ARIAS J.A., ÁLVAREZ-GONZÁLEZ J.G., GADOW K.v., 2006. Site quality curves for birch stands in north-western Spain. Silva Fennica 40 (4): 631-644.
- GORGOSO J.J., 2003. Caracterización de las distribuciones diamétricas de *Betula alba* L. en Galicia, Ph. D thesis, Univ. Santiago de Compostela (Unpublished), 176 pp.
- GORGOSO J.J., ÁLVAREZ-GONZÁLEZ J.G., ROJO A., GRANDAS-ARIAS J.A., 2007. Modelling diameter distributions of *Betula alba* L. stands in northwest Spain with the two-parameter Weibull function. Invest. Agrar.: Sist. Recur. For. 16(2): 113-123.
- HYINK D.M., MOSER J.W., 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions, For. Sci. 29: 85-95.
- INDUROT, 2001. Cartografía temática ambiental del Principado de Asturias. Principado de Asturias – Universidad de Oviedo.
- JOHNSON N.L., KITCHEN J.O., 1971. Some notes on tables to facilitate fitting S_B curves. Biometrika 58 (1): 223-226.
- KAMZIAH A.K., AHMAD M.I., LAPONGAN J., 1999. Nonlinear regression approach to estimating Johnson S_B parameters for diameter data. Can. J. For. Res. 29(3): 310-314.
- KNOEBEL B.R., BURKHART H.E., 1991. A bivariate distribution approach to modeling forest diameter distributions at two points in time. Biometrics 47: 241-253.

- LEAK W.A., 1996. Long-term structural change in uneven-aged northern hardwoods. *For. Sci.* 42: 160-165.
- LENHART J.D., CLUTTER J.L., 1971. Cubic-foot yield tables for old field loblolly pine plantations in the Georgia Piedmont. *Ga. For. Res. Counc. Rep.*, 22, series 3, 13 pp.
- LI F., ZHANG L., DAVIS C.J., 2002. Modeling the joint distribution of tree diameters and heights by bivariate generalized Beta distribution. *For. Sci.* 48(1): 47-58.
- LINDSAY S.R., WOOD G.R., WOOLLONS R.C., 1996. Modelling the diameter distribution of forest stands using the Burr distribution. *Journal of Applied Statistics* 23: 609-619.
- LIU C., ZHANG S.Y., LEI Y., NEWTON P.F., ZHANG L., 2004. Evaluation of three methods for predicting diameter distributions of black spruce (*Picea mariana*) plantations in central Canada. *Can. J. For. Res.* 34: 2424-2432.
- LOETSCH F., ZÖHRER F., HALLER K.E., 1973. Forest inventory 2. Verlagsgesellschaft. BLV. Munich. 469 pp.
- LORIMER C.G., 1980. Age structure and disturbance history of a Southern Appalachian virgin forest. *Ecology* 61: 1169-1184.
- MALTAMO M., PUUMALAINEN J., PÄIVINEN R., 1995. Comparison of beta and Weibull functions for modelling basal area diameter distribution in stands of *Pinus sylvestris* and *Picea abies*. *Scan. J. For. Res.* 10: 284-295.
- MEYER H.A., 1952. Structure, growth and drain in balanced uneven-aged forests. *J. For.* 50: 85-92.
- MEYER H.A., STEVENSON D.D., 1943. The structure and growth of virgin beech, birch, maple, hemlock forest in northern Pennsylvania. *J. Agric. Res.* 67: 465-478.
- MILLER L.H., 1956. Tables of percentage points of Kolmogoroff statistic. *Journal of the American Statistical Association* 51: 111-121.
- NELSON T.C., 1964. Diameter distribution and growth of loblolly pine. *Forest Sci.*, 10: 105-115.
- PALAHÍ M., PUKKALA T., BLASCO E., TRASOBARES A., 2007. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). *Eur. J. Forest Res.* 126: 563-571.
- RÍO M. Del, 1999. Régimen de claras y modelo de producción para *Pinus sylvestris* L. en los sistemas Central e Ibérico, Tesis doctorales INIA, Serie: Forestal 2. Madrid, España, 257 pp.
- ROJO A., ÁLVAREZ-GONZÁLEZ J.G., GRANDAS J.A., DIÉGUEZ-ARANDA U., 2005. Tablas de producción de selvicultura media para el abedul (*Betula alba* L.) en Galicia. IV Congreso Forestal Español. Zaragoza, 26-30 septiembre 2005.
- ROTHMALER W., De CARVALHO E VASCONCELLOS J., 1940. *Betula celtiberica* Rothm. et Vasc. Ein Beitrag zur systematik der westeuropaischen Birken. *Bol. Soc. Brotariana XIV H. 2.* Coimbra. 167 pp.
- RUIZ De La TORRE J., 1979. Árboles y arbustos de la España Peninsular. Fundación Conde del Valle de Salazar. Escuela Técnica Superior de Ingenieros de Montes. Madrid, 512 pp.
- RYAN T.P., 1997. Modern regression methods. Ed. Wiley, INC. 515 pp.
- SCHNUR G L (1934) Diameter distributions for old-field loblolly pine stands in Maryland. *J. Agric. Res.* 49: 731-743.
- TORRES-ROJO J.M., MAGAÑA O.S., ACOSTA M., 2000. Metodología para mejorar la predicción de parámetros de distribuciones diamétricas, *Agrociencia* 34: 627-637.
- WANG M., RENNOLLS K., 2005. Tree diameter distribution modelling: introducing the logit-logistic distribution. *Can. J. For. Res.* 35: 1305-1313.
- WESTPHAL C., TREMER N., OHEIMB G.V., HANSEN J., GADOW K.v., HÄRDITL W., 2006. Is the reverse J-shaped diameter distribution universally applicable in European virgin beech forests?. *For. Ecol. Manage.* 223: 75-83.
- XUNTA DE GALICIA, 2001. O monte galego en cifras. Consellería de Medio Ambiente. Dirección Xeral de Montes e Medio Ambiente Natural. 226 pp.
- ZHANG L., PACKARD K.C., LIU C., 2003. A comparison of estimation methods for fitting Weibull and Johnson's SB distributions to mixed spruce-fir stands in northeastern North America. *Can. J. For. Res.* 33: 1340-1347.
- ZÖHRER F., 1969. The application of the beta function for best fit of stem diameter distributions in inventories of tropical forest. *Mitt. Bundesforsch.-anst. Forst- u. Holzwirtsch., Reinbek/Hamburg* 74: 279-293.
- ZÖHRER F., 1970. Das Computerprogram BETKLA zum Ausgleich von Stammzahl-Durchmesser-Verteilungen mit Hilfe der Beta-Verteilung. *Mitt. Bundesforsch.-anst. Forst- u. Holzwirtsch., Reinbek/Hamburg* 76, 50 pp.