Volume 129 • Number 4 • July 2010

European Journal of

Forest Research

176

Springer

Ingrowth model for pyrenean oak stands in north-western Spain using continuous forest inventory data

Patricia Adame · Miren del Rìo · Isabel Cañellas

Abstract In this paper, a two stage ingrowth model is presented for predicting periodic, 10 years ingrowth for pyrenean oak (*Quercus pyrenaica* Willd.) grown in medium to fully stocked coppice stands in north-western Spain. Data from the Spanish National Forest Inventory was used to develop the model, extracting the information from two inventories taken in 222 permanent plots. The first stage of the model predicts the probability of ingrowth occurrence, and in the second stage, the number of recruits is predicted using a conditional model. Both models were biologically realistic and presented logical behaviour. The ingrowth occurrence probability model was dependent on quadratic mean diameter and average height. The recruitment quantification model included stand density and average diameter as explanatory variables. Although the occurrence probability of ingrowth was predicted correctly in 71.7% of cases, the predictions of the number of recruitment are poorer, yielding a coefficient of determination of 0.358. The evaluation criteria included qualitative and quantitative examinations and a testing with independent data from another region. The proposed ingrowth model is the first to be developed for mediterranean oak species in Spain and is an essential feature in any stand growth system.

Keywords National Forest Inventory · Recruitment · *Quercus pyrenaica* Willd · Coppice stand · Two-stage model

P. Adame (✉) · M. d. Rìo · I. Cañellas
Forest Research Centre (CIFOR-INIA), Ctra La Coruña km 7.5,
Madrid 28040, Spain
e-mail: adame.patricia@inia.es

Introduction

A forest is a dynamic biological system that is continuously changing. Management practices influence forest dynamics and successional processes by changing environmental conditions for growth, regeneration and survival (Liu and Ashton 1998). Modelling provides a major tool for synthesis because models can project long-term response to changes in management practises. For dynamic growth and yield models to be useful, the following basic elements must be available (García 1988): (1) a description of the state of the forest at a given point in time; (2) transition functions to define the rate of change of the system depending on the current state of the stand; and (3) control functions to regulate the modifications of the values of the main stand variables caused by instantaneous changes of the state due to silvicultural treatments.

A dynamic growth model simulator consists of different sub-models, including an ingrowth model as one important factor in population dynamics (Grubb 1977). Ingrowth must be taken into account in growth simulations, particularly in long range predictions or where stands structure is very heterogeneous, because of the larger probability of ingrowth. Ingrowth is usually quantified in terms of the number of trees that reach or exceed a specific threshold size over a certain period, as a result of regeneration processes like establishment, growth and mortality of seedlings (Lexerød and Eid 2005). An ingrowth model predicts the recruitments directly from independent variables describing site and stand conditions.

Trees which make up the stands are the primary source of natural recruits, so it is important to determine which variables have the greatest influence on successful ingrowth (Bendzsak 2004). There are two main approaches when modelling ingrowth: static approaches and dynamic

approaches. Static approaches predict a constant amount of ingrowth each year irrespective of stand condition; whereas dynamic approaches respond to stand condition, predicting ingrowth as a function of stand density, composition or others parameters.

In the past, the most commonly used approach involved a single equation (Moser 1972; Adams and Ek 1974; Buongiorno and Michie 1980; Hyink and Moser 1983; Vanclay 1989), but currently, the stochastic nature of recruitment is also modelled by defining two separate states (Hamilton and Brickell 1983; Vanclay 1992; Qin 1998; Kariuki 2004; Lexerød and Eid 2005). According to Hamilton and Brickell (1983), a 'two-state' system generally implies that (1) observations cannot be placed with certainty in either of the two states and (2) the relationship between the dependent variable and potential independent variable is different for each of the two states. In a two-stage approach, the first equation estimates the probability that some ingrowth will occur, usually with a logistic function with presence (or absence) of ingrowth as the response variable. The second stage involves a conditional function to predict the amount of ingrowth, given that ingrowth is known to occur, and can be estimated using ordinary linear regression (Vanclay 1992) or nonlinear regression. Burk (1978) reported that linear forms are not appropriate to predict the amount of ingrowth since the predicted number of ingrowth trees may be negative. To avoid this problem, exponential models are commonly used (Vanclay 1992; Qin 1998; Lexerød and Eid 2005).

Ingrowth estimation can be carried out using continuous forest inventory data from National Forest Inventories or local inventories. When a data set is taken from National Forest Inventories, large amount of data are available, but the sampling methodology was not specifically designed to develop growth and yield models (Trasobares et al. 2004). Usually, this kind of inventories does not measure some necessary variables as age of the trees. On the other hand, some measured variables from rather small inventory plots are also characterized by larger errors when they are compared to measurements from large research plots. Nevertheless, assuming that measurement errors are random, the large sample should compensate for this deficiency (Monserud and Sterba 1996).

The pyrenean oak (*Quercus pyrenaica* Willd.) is distributed throughout Spain covering a total surface area of 1,208,768 ha (DGCN 2006) (Fig. 1). As a result of ecological variability, different silvicultural treatments (usually coppice management), forest fires and the state of neglect into which many stands have fallen in recent years, stand conditions vary greatly. Stands are often of low density or take the form of open-woodlands with large diameter trees, although well-stocked coppices represent around 43% of stands. The increasing interest in using the latter for either commodity production (such as wine barrels and silvopastoral uses) or non-commodity production (such as recreation and environmental preservation) justifies the need to guarantee the sustainable management of pyrenean oak stands, and therefore, the interest in developing a dynamic growth model (Cañellas et al. 2004). Nowadays, a site index model (Adame et al. 2006), the diameter growth model (Adame et al. 2008b) and the height–diameter model (Adame et al. 2008a) are available for pyrenean oak.

Fig. 1 Map of sample locations and biogeoclimatic strata (Elena Roselló 1997) of the Spanish National Forest Inventory plots from which the data was drawn to fit the occurrence probability and amount of ingrowth models
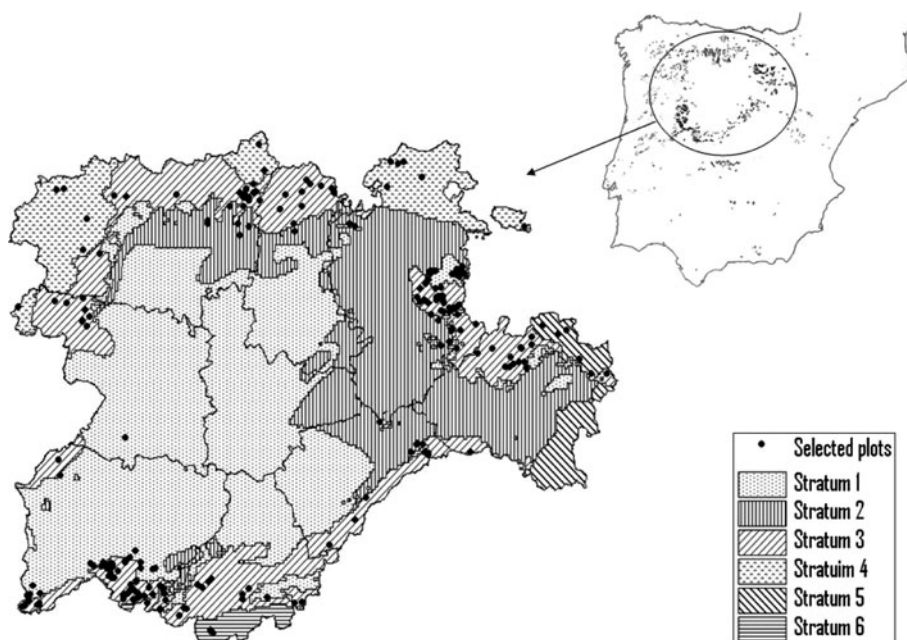
The primary objective of this study was to develop an ingrowth model for pyrenean oak stands in north-western Spain in order to be included in the individual-tree growth model for coppices stands of the species. The model was developed using the two-stage approach, predicting first, the ingrowth probability and secondly, the number of recruits. The independent variables included in the model were restricted to variables directly or indirectly available from forest inventories for practical forest management planning.

Materials and methods

Data collection

The study area is the Castilla y León region, which covers 94.225 km^2 in north-western Spain (Fig. 1). Data belonging to the second (DGCN 1996) and third (DGCN 2006) Spanish National Forest Inventory (SNFI) were used to develop the ingrowth model. The SNFI is a systematic sample of permanent plots distributed on a 1-km square grid, with a remeasurement interval of 10 years. The plots are circular in shape with variable radius and are composed of four sub-plots with radii of 5, 10, 15 and 25 m. The minimum diameters at breast height (dbh) thresholds on the different radii are 75, 125, 225 and 425 mm, respectively. At each inventory, the data recorded for each sample tree are species, diameter at breast height, total height, distance and azimuth from the plot centre.

Ingrowth modelling is a very stochastic process, and available information from SNFI is plenty of limited (any silvicultural information), so an initial selection of suitable plots was made from all the SNFI plots located in the Castilla y León area, selecting pure *Q. pyrenaica* stands (basal area proportion over 90%). Because of the high variability in silvicultural and ecological conditions in pyrenean oak stands as well as the lack of past management information in the SNFI, a second selection was made according to a stand typology classification for *Q. pyrenaica* stands (Roig et al. 2007). The coppices selected were characterized by medium to high densities (Reineke's stand density index over 200) and regular diameter distributions, excluding impoverished or very sparsely populated stands. Since no information was available with respect to why and how harvest was carried out, plots that received any harvesting treatment between the two measurements were also excluded. This selection was made in the same way as in other components of the growth model. Figure 1 shows the distribution of the plots used in this work.

Recruitment was calculated for each plot as the number of new trees that reached or exceeded a dbh of 75 mm between two measurements (10-year growth period). The data in this study are taken from the first sub-plot (5 m radius and minimum dbh threshold of 75 mm). The plots were established between 1992 and 1993 and remeasured after 10 years, i.e. between 2002 and 2003.

Six strata, according to biogeoclimatic characteristics, can be identified in the region (Elena Roselló 1997). Precipitation, temperature, altitude and soil characteristics were taken into account to classify these strata. Ecological characteristics calculated according to each location (altitude, geographic position and catchment area) for each plot are evaluated using the climatic estimation models proposed by Sánchez et al. (1999).

A total of 222 permanent plots were used to fit the ingrowth model. Stand variables from the sample plots, including the mean and standard deviation for the number of ingrowth trees are shown in Table 1. Climatic characteristics are indicated in Table 2 according to the biogeoclimatic stratum defined by Elena Roselló (1997).

Independent data sets from permanent sample plots located in ecorregion 2 according to Elena Roselló (1997), in central Spain and on the southern slopes of the Central mountain range (Madrid region), were used to test the performance of the proposed model. Measurements were taken, and data collected at these sample plots using an equal procedure to that which was used in the modelling data set. Therefore, 19 plots were analysed (Table 1).

Ingrowth model

Much of the variability seen in recruitment data reflects the fact that during any period regeneration may or may not occur. This characteristic (no/some recruitment) becomes especially apparent when individual species is modelled. Such data may be modelled using a two stage approach (Vanclay 1994a). In the first step, an equation to predict the probability of ingrowth occurrence in a specific plot was fitted. The second step predicts the number of trees recruited, which depends on different independent variables.

Probability of ingrowth occurrence

Logistic regression is a useful method for analysing data that includes categorical response variables and does not require large sample sizes for the data analysis to be feasible (Bergerud 1996). Assume that there are k explanatory variables $x = (x_1, x_2, \ldots, x_k)$ at the beginning of a growth period, the response variable $\pi(x)$ is a binary variable to indicate the presence status of ingrowth trees in a plot at the end of the period, where $0 =$ absence and $1 =$ presence of recruitment. The probability of presence of recruitment over a 10-year period $\pi(x)$ was modelled with the following logistic function:

Table 1 Main characteristics of second national forest inventory plots used to fit and to test the model

	Number of plots	N	D_g	D_m	D_O	H_m	H_O	BA	N_{ingr}
Fitting data set									
Plots with ingrowth	106	1315.2	13.5	12.6	20.8	7.5	9.4	15.9	357.9
		(695.8)	(4.9)	(4.0)	(8.8)	(1.7)	(2.2)	(6.9)	(274.5)
Plots without ingrowth	116	948.6	18.1	16.9	26.3	9.2	11.4	18.8	–
		(661.1)	(7.4)	(7.0)	(9.0)	(2.5)	(3.0)	(8.6)	–
All plots	222	1126.8	15.9	14.8	23.6	8.4	10.4	17.4	–
		(701.1)	(6.7)	(6.1)	(9.2)	(2.3)	(2.8)	(7.9)	–
Test data set									
Plots with ingrowth	15	1509.6	12.4	11.8	19.2	7.7	9.5	15.1	375.6
		(699.5)	(4.4)	(4.0)	(7.9)	(1.6)	(2.5)	(4.8)	(266.9)
Plots without ingrowth	4	1188.8	19.0	17.0	31.3	10.2	13.0	21.7	–
		(1035.7)	(8.6)	(7.1)	(11.9)	(1.2)	(1.9)	(8.7)	–
All plots	19	1442.1	13.8	12.9	21.8	8.2	10.2	16.5	–
		(759.9)	(5.9)	(5.0)	(9.9)	(1.8)	(2.8)	(6.2)	

N = density (trees/ha); D_g = quadratic mean diameter (cm); D_m = average diameter (cm); D_O = dominant diameter (cm); H_m = average height (m); H_O = dominant height (m); BA = Basal area (m^2/ha); N_{ingr} = number of ingrowth trees per hectare. Dominant variables were defined as mean variable of the thickest 100 stems per hectare. Standard deviation in *brackets*

Table 2 Climatic characteristics of national forest inventory plots used to fit the model classified according to the biogeoclimatic strata (Elena Roselló 1997)

	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	Stratum 6	All strata
Number of plots	36	20	145	13	5	3	222
Plots with ingrowth	16	16	71	2	2	0	107
Plots without ingrowth	20	4	74	11	3	3	115
Altitude (m)	965.7	1006.3	1114.6	957.1	1171	676.4	1058.1
P annual (mm)	840.5	723.8	997.1	1224.4	583.2	1276.7	942.9
P winter (mm)	276.3	227.0	337.9	433.1	142.5	441.0	315.8
P spring (mm)	236.3	199.6	275.1	318.5	170.8	359.9	260.4
P summer (mm)	89.8	102.8	97.9	142.1	123.5	62.4	98.5
P autumn (mm)	238.1	194.5	286.2	330.7	146.7	413.6	268.1
T annual (°C)	11.0	9.9	10.1	9.5	10.0	14.0	10.3
T winter (°C)	3.7	2.6	2.8	3.4	2.8	6.1	3.0
T spring (°C)	19.0	17.8	18.0	15.9	18.0	22.9	18.2
T max (°C)	28.3	27.4	27.4	23.7	27.1	32.4	27.5
T min (°C)	−1.1	−2.4	−2.0	−1.0	−1.6	1.2	−1.8
Fluctuation (°C)	29.4	29.7	29.4	24.7	28.7	31.2	29.3
ETP (mm)	664.6	633.5	636.1	615.4	633.4	768.6	642.9
Surplus (mm)	463.2	348.2	613.2	776.6	190.5	873.5	558.0
Deficit (mm)	287.3	257.8	252.2	167.7	240.5	365.3	258.1
Water index	43.7	30.6	72.6	111.5	7.4	85.7	62.9
Drought period (month)	1.9	1.8	1.7	0.5	1.5	2.5	1.7

P = annual and seasonal (spring, summer, autumn and winter) rainfall (mm), T = average annual and seasonal (summer and winter) temperature (°C), T_{max} = hottest month average maximum temperature (°C), T_{min} = coldest month average minimum temperature (°C), fluctuation = difference between hottest month average temperature and coldest month average temperature in °C, ETP = potential evapotranspiration (mm), *Surplus* = water surplus in mm, *Deficit* = water deficit in mm

$$\pi(x) = \frac{\exp^{\alpha+\beta x}}{1 + \exp^{\alpha+\beta x}} = \left[1 + \exp^{-(\alpha+\beta x)}\right]^{-1} \qquad (1)$$

where α and β are parameters defining the S-shaped function of x. The concept of the generalized linear model is based on an assumption that a function of $\pi(x)$ may be expressed as a linear combination of x, $g[\pi(x)]$. The most frequently used link function is the logistic link:

$$g[\pi(x)] = \log\left[\frac{\pi(x)}{1 - \pi(x)}\right] = \alpha + \beta x. \qquad (2)$$

All the stand-level variables defined in Table 1 and the ecological variables presented in Table 2, as well as different combinations of them were tested as independent variables in the general logistic model. The biogeoclimatic strata inclusion in the model was tested using dummy variables Eq. 2 was estimated with PROC LOGISTIC using maximum likelihood estimation (SAS/STAT 2000). Stepwise selection to identify the variables was used, and a significance level of 0.05 was required for a variable to be retained in the model. Both biologically meaningful selection and statistically significant selection were considered. An empirical ratio estimator suggested by Snowdon (1991) was applied to correct the bias created when the model predictions are transformed back to the original units (Flewelling and Pienaar 1981).

Pearson statistic and standard deviance have been used as standard procedures to asses goodness-of-fit in logistic regression. These tests do not work well with sparse data, since they do not follow an χ^2 distribution under the null hypothesis and are as such not valid measures of model fit (McCullagh and Nelder 1986). Kuss (2001) found some deficiencies in the traditional Hosmer–Lemeshow test, so other tests were suggested to test the goodness-of-fit for the model: the Osius and Rojek test (Osius and Rojek 1992), the McCullagh-test (McCullagh 1985), the information matrix test (IM-test) (White 1982) and the RSS-(Residual Sum of Squares)-test (Copas 1989). These statistics were assessed with the %GOFLOGIT macro (Kuss 2001) using the IML module from the SAS/IML User's Guide (SAS/STAT 2000).

The model predicts the continuous probabilities, so it is necessary to define a threshold or cut-off point to convert these back to dichotomous results and to maximize the rate of correct classification. 'Sensitivity' is defined as the proportion of events correctly classified, and 'specificity' is defined as the proportion of non-events correctly classified. Sensitivity and specificity facilitate the model's *ROC* (receiver operating characteristic) curve, which is plotted as sensitivity against 1-specificity with all possible cut-off points or thresholds (Pepe 2000). If the model discriminates perfectly, the ROC curve passes through the point (0, 1), and the area below the curve is one. Each point on the ROC curve provides the sensitivity and specificity measures associated with a cut-off point in the probability scale which allows classification of each observation as either a predicted event or a predicted non-event.

Amount of ingrowth

The curvilinear response provided by the logarithmic transformation of relative abundance gave a slightly better fit than a linear model and ensured more conservative extrapolations (Vanclay 1992). Furthermore, log-transformation stabilizes variance, which is necessary for regression analysis. The relationship between the number of recruits (N_{ing}) and the independent variables was expressed as follows:

$$\ln(N_{ing}) = \mathbf{X}\prime \cdot \beta + \varepsilon \qquad (3)$$

where X' is the transpose of the vector of independent variables, β is a vector of regression coefficients and ε is a vector of random errors. It was assumed that the error term ε had $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$ and that they were uncorrelated.

The linear least squares fits for Eq. 3 were accomplished using the PROC REG procedure on SAS software (SAS/STAT 2001). Some of the independent variables were transformed in order to express a linear relationship with the dependent variable. Stepwise selection of the explanatory variables was used, and a significance level of 0.05 was required for a variable to be retained in the model. Explanatory variables tested were the same that for the probability model (Tables 1, 2).

A transformation can bias the results because logarithmic regression theoretically estimates "medians" instead of the desired "means". Therefore, a bias correction was needed to transform the model predictions back to the original units (Flewelling and Pienaar 1981). An empirical ratio estimator suggested by Snowdon (1991) was applied: $\frac{N_{ing10}}{\exp\left[\ln \hat{N}_{ing10}\right]}$, where N_{ing10} is the mean number of recruits for 10 years and $\hat{N}_{ing10}$ is the mean estimated number of recruits.

Model evaluation

The model was evaluated quantitatively by examining the magnitude and distribution of residuals to detect any obvious patterns and systematic discrepancies, and by testing for bias and precision to determine the accuracy of model predictions (absolute and relative bias, root mean square error, coefficient of variation and coefficient of determination of the linear regression of observed versus predicted data) (Vanclay 1994b; Soares et al. 1995; Gadow and Hui 1998).

Independent data set was used to test the model. The ingrowth probability for each plot was calculated, and the predicted and observed ingrowths were compared applying the same previous evaluation criteria.

Fig. 2 Relationship between ingrowth occurrence probability and significant variables in logistical model (Eq. 4): quadratic mean diameter D_g (**a**) and average stand height H_m (**b**)
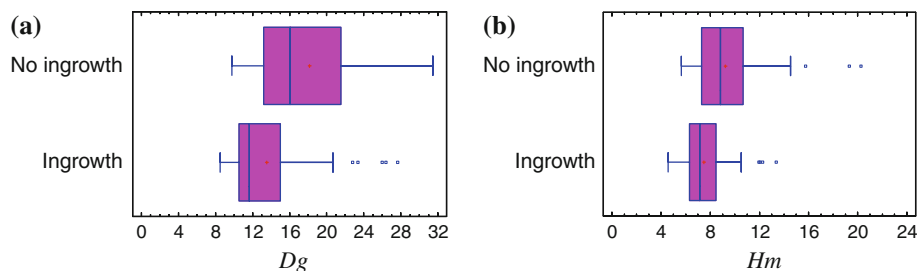
Results

Probability of ingrowth occurrence

The significant variables were selected based on the likelihood ratio test, and a criterion that rational interpretation is more important than model fit. An appropriate biological interpretation of the ingrowth process is essential in considering variable selection. Variables describing stand conditions (quadratic mean diameter D_g and average height H_m) at the beginning of the period were found to be highly significant in predicting the probability of recruitment (Fig. 2). The relationship between both H_m and D_g and recruitment was negative. Strata variable was tested as categorical variable, and it was defeated.

The logistical model for *Quercus pyrenaica* was:

$$\pi = \left[1 + \exp\left(-\left(a_0 + a_1 D_g + a_2 H_m\right)\right)\right]^{-1} \qquad (4)$$

where D_g is quadratic mean diameter (cm), and H_m is average height (m). The parameters a_0, a_1 and a_2 and standard deviations are shown in Table 3.

A threshold value of 0.58 gives the best hit rate and allows 71.7% of plots to be classified correctly, with sensitivity equal to 61.2% and specificity equal to 81.0%. Figure 3 shows the classification of plots into those with ingrowth observed and those with no ingrowth observed, and Fig. 4 is the *ROC* curve based on the sensitivity and the specificity. The area under the ROC curve was 0.8272.

The conducted goodness-of-fit tests are shown in Table 4, and they did not reveal any indication of a poor
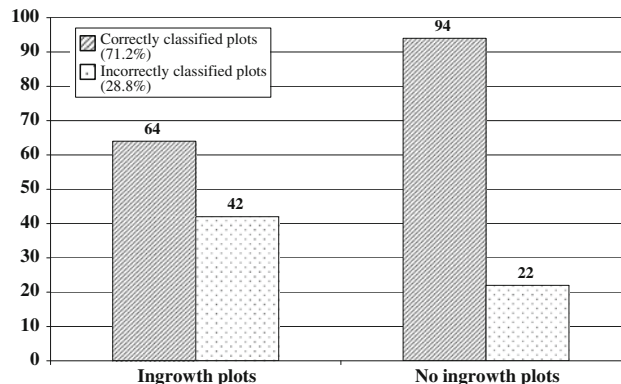
Fig. 3 Classification of the studied plots according to the logistical model (Eq. 4) in relation to the observed situation in each of them
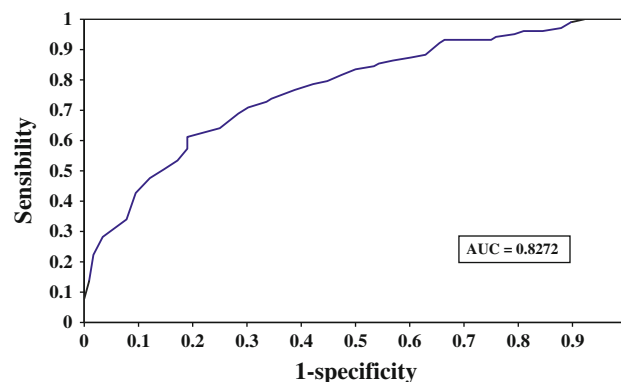
Fig. 4 Receiver operating characteristic (*ROC*) curve and area under the curve (*AUC*) for the ingrowth occurrence probability model

Table 3 Parameters and standard deviations of both logistical and amount of ingrowth models

Parameter	Value	Standard deviation
Logistical model		
a_0	4.4765	0.7994
a_1	−0.1167	0.0393
a_2	−0.3492	0.1044
Amount of ingrowth		
b_0	8.60716	0.85419
b_1	−0.64247	0.14346
b_2	197.99212	30.88893

Table 4 Goodness-of-fit tests applied in logistic model (Eq. 4) (Kuss 2001)

Goodness-of-fit tests	Value	P-value
Standard Pearson test	218.605	0.438
Standard deviance	249.559	0.058
Osius-test	−0.009	0.504
McCullagh-test	0.071	0.472
IM-test	13.293	0.004
RSS-test	42.491	0.466

model fit. Tests which are based on a summation of residuals (Standard Pearson test, Osius-test, McCullagh-test and RSS-test) behave better than those which rely on the principle of comparing two different estimators of the information matrix that should give comparable results under a satisfactory model fit (IM-test).

Amount of ingrowth

Once the probability of recruitment is determined, the next stage is to predict the conditional number of trees. Only the 106 sample plots where ingrowth occurred were used to fit this equation. An exploratory analysis of the data suggested that recruitment was related to N and D_m^2 and that none of the climatic variables had a significant influence. The empirical ratio estimator suggested by Snowdon (1991) was 1.1897. The amount of ingrowth is worked out according to the following formulation:

$$\ln(N\text{ing}) = b_0 + b_1 \ln N + b_2 {1}/{D_m^2} \tag{5}$$

where N_{ing} is the number of recruitments (trees/ha), N is the number of trees per hectare, and D_m is average diameter (cm). All parameters (b_0, b_1 and b_2) and standard deviations are shown in Table 3 and were significant ($P < 0.001$).

Parameter of the number of trees per hectare was negative, and inverse of square average diameter of the stand were positively affected, so ingrowth was lower where competition was greater. In general, the model results revealed significant differences between the predicted and the observed number of recruits (Table 5), shown on the coefficient of determination (R_{adj}^2) for linear regression of observed versus predicted data of this conditional model (0.364) and the coefficient of variation (0.28). The corrected residual was −3.19 E-13, although the absolute mean residual was 161.31. The residuals form lines (Fig. 5) owing to small plot size, meaning that the measured values

Table 5 Fit statistics for the conditional model for number of recruits applied to both fitting and test data set

Evaluation criteria	Fitting data	Test data
Mres	−3.19 E-13	10.18
VR	0.28	2.24
RMS	433302.66	88945.14
Amres	161.35	188.38
R²/Mef	0.358	0.108

Mres = mean residual; *VR* = variance ratio; *RMS* = residual mean of squares; *Amres* = absolute mean residual; *R²/Mef* = coefficient of determination/model efficiency of linear regression of observed versus predicted data
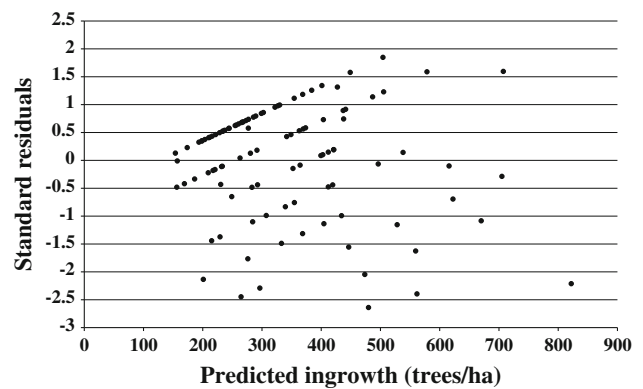
Fig. 5 Plot of standard residual (predicted ingrowth—observed ingrowth) against number of predicted ingrowth trees per hectare calculated according to Eq. 5

can only have certain values (maximum of 10 ingrowth trees for radius 5 m subplot).

Discussion

The first stage of the two step model for pyrenean oak predicts the probability of ingrowth occurrence from square mean diameter and average stand height. The second stage predicts the expected amount of ingrowth (where this is predicted to occur) using density and average stand diameter as input variables.

The data set was drawn from the Spanish National Forest Inventory (SNFI) permanent plots with variable radius, which ensured a representative distribution of ecological characteristics. The circular plot with variable radius is often the only feasible method to sample irregular stand structures as coppices efficiently (Trasobares et al. 2004). The problem with this data source was that the only suitable data to fit the ingrowth model was that obtained in the smallest radius subplots (5 m), since these were the only ones in which the minimum size for NFI measurement of stems was 7.5 cm, so larger random errors can be expected because ingrowth often is clustered. Although the probability of recruitment is lower in a smaller sized plot, the conditional number of recruits per unit area will be relatively large, so an appropriate final level can be determined (Lexerød and Eid 2005). Another weakness associated with this data source is that all the independent variables used were registered in the larger plot (radius 25 m), including those of the ingrowth sample plot (radius 5 m). It is unlikely that the potential differences in these variables biased the parameter estimates since they can be treated as random measurement errors, and the relatively large number of plots reduces the effect of extreme combinations of variables (Lexerød and Eid 2005).

The two-stage model presents several advantages over a single model (Hamilton and Brickell 1983): (1) The residuals from a single model are rarely either normally distributed or homogeneous; (2) the two stage model is a much higher resolution model than the single model, since the former is capable of describing the process more completely than would be possible with a single model, regardless of its functional form.

Current recruitment depends on regeneration conditions of previous years (canopy, climate, harvesting, etc.…) (Leak and Graber 1976) and the minimum diameter of the dbh-distribution used in the inventory (here 75 mm). Gap dynamics is characterized by small or micro-scale disturbance of the mature forest canopy. Trees die standing, snap off, or are blown to the ground, creating a "hole" in the canopy. The death of a single stem or a few stems releases available growing space. This growing space is then occupied by tree regeneration, usually as a result of released advance regeneration or recruitment from buried or dispersed seed (McCarthy 2001). In general, the longer the time after harvesting, the greater the amount of natural regeneration (Hao et al. 2005). Furthermore, there is an anthropogenic influence in terms of recurrent fires, grazing livestock, collection of firewood and poles as well as harvesting of non-timber products on pyrenean oak stands. Silvicultural practices are not common in these stands and in any case, information regarding such interventions does not appear in the SNFI data set. Therefore, these variables were not considered in this study.

Ingrowth models for *Quercus* species are scarce. Pyrenean oak is characterized by its stump and root sprouting capacity. This type of structure makes modelling the dynamics of the stand more difficult in comparison with other species which only regenerate through stump sprouting or natural seeding. Published results for broadleaf trees have displayed relatively large differences between the expected and the observed number of recruits as obtained results in this paper. Bendzsak (2004) concluded in his work, that the only variables that were directly related to recruitment, for various species including red and white oak, were the pH of the A soil horizon and the elevation. Lexerød and Eid (2005) developed recruitment models for some broadleaf species (including oaks) in young growth forests in Norway with variables describing location, site and stand conditions. However, for *Q. pyrenaica*, neither biogeoclimatic strata nor ecological tested variables have been found to be significant, although soil data were not available.

Negative parameters for both average height (H_m) and square mean diameter effects imply that the ingrowth rate of trees generally decreases as they increase. Greater average height and square mean diameter indicate older trees or mature stands, so, they show smaller ingrowth

occurrence probability than younger stands. Whereas, others studies result density variables as significant (i.e. basal area (Buongiorno et al. 1995; Hao et al. 2005) or density (Qin 1998; Lexerød and Eid 2005).

In stands with ingrowth occurrence, the number of ingrowth trees is related to stand density, where the number of recruits decreases as N values increase. For a given stand development stage, this variable is closely related to growing space and availability of resources to trees, so high density values lead to increased competition and mortality (Lexerød and Eid 2005). Finally, the mean stand diameter is also significant in the recruitment quantification model, since the ingrowth process is negatively affected by the presence of greater competition and older stands.

The correlation between the amount of ingrowth, certain independent variables, and the statistical results are low, indicating either that the relationships are strongly nonlinear or that there is a high degree of natural variability in the ingrowth process (Qin 1998). Variability in ingrowth can be due to factors that are not contemplated in the models (site index, soil characteristics, fires, grazing livestock, etc.…), to the stochastic nature of the ingrowth process, and to the small plot size. Significant increases in the precision of ingrowth predictions seem unlikely unless the suite of independent variables considered is expanded or the model recalibrated for a more limited set of forest conditions. In other words, improved predictions may be possible by incorporating additional site-specific information or by performing the analysis on a more geographically restricted data set (Shifley et al. 1993).

Independent data were available for model evaluation. The application of logistical model to find out which plots will present recruitment was satisfactory (71.7% of plots classified correctly), although the amount of ingrowth produces worse result (coefficient of determination equal to 0.108). These results are similar to fitting results. Huang et al. (2003) asserted that the quality of the model fit can also be assumed to reflect the quality of predictions when the data are collected through a well-designed and representative sampling-process, when the models will be used within the range of the data and when the models are biologically realistic.

This study is the first to focus on ingrowth models for mediterranean oak species in Spain and provides a foundation for further investigation. Although a selection of plots according to a stand typology was made to fit the ingrowth model, data set still presents high variability in silvicultural and ecological conditions. The regeneration of pyrenean oak generally takes place through stump and/or root sprouting; therefore, it may be essential to consider factors such as silvicultural treatments, forest fires or livestock grazing in the ingrowth model. This information is not available in the Spanish National Forest Inventory,

the source from which the data for this study was drawn, although additional measurements could significantly improve our understanding of pyrenean oak stand dynamics. Because of the biological importance of the species, this study also represents a starting point for further studies concerning yield optimisation and the evaluation of alternative management regimes for the species, and it is an important component of stand growth and yield models.

References

Adame P, Cañellas I, Roig S, del Rio M (2006) Modelling dominant height growth and site index curves for rebollo oak (*Quercus pyrenaica* Willd.). Ann For Sci 63:929–940

Adame P, Del Río M, Cañellas I (2008a) A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). For Ecol Manage 256:88–98

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

Adams DM, Ek AR (1974) Optimizing the management of uneven-aged forest stands. Can J For Res 4:274–287

Bendzsak M (2004) A survey of hardwood recruitment under red pine (*Pinus resinosa* Ait.) plantations at Joker's Hill. Faculty of Forestry, University of Toronto, Ontario

Bergerud WA (1996) Introduction to regression models: with worked forestry examples. Biometrics information handbook no. 7. British Columbia. Ministry of Forests Research Program, Victoria, p 147

Buongiorno J, Michie BR (1980) A matrix model for uneven-aged forest management. For Sci 26:609–625

Buongiorno J, Peyron J, Houllier F, Bruciamacchie M (1995) Growth and management of mixed-species, uneven-aged forests in the French Jura: implications for economic returns and tree diversity. For Sci 41:397–429

Burk TE (1978) Characterizing the growth dynamics of Lake States aspen. University of Minnesota, Minnesota, p 90

Cañellas I, Del Río M, Roig S, Montero G (2004) Growth response to thinning in *Quercus pyrenaica* Willd. coppice stands in Spanish central mountain. Ann For Sci 61:243–250

Copas JB (1989) Unweight sum of squares test for proportions. Appl Stat 38:71–80

DGCN (1996) II Inventario Forestal Nacional 1986–1996, Dirección General de Conservación de la Naturaleza, Ministerio de Medio Ambiente, Madrid

DGCN (2006) III Inventario forestal nacional 1996–2006, Dirección General de Conservación de la Naturaleza, Ministerio de Medio Ambiente, Madrid

Elena Roselló R (1997) Clasificación Biogeoclimática de España Peninsular y Balear. MAPA, Madrid, p 446

Flewelling JW, Pienaar LV (1981) Multiplicative regression with lognormal errors. For Sci 27:281–289

Gadow KV, Hui G (1998) Modelling forest development. Kluwer Academic Publishers, The Netherlands

García O (1988) Growth modelling—a (re)view. New Zealand Forestry 33:14–17

Grubb PJ (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. Biol Rev 52:107–145

Hamilton DA, Brickell JE (1983) Modeling methods for a two-state system with continuous responses. Can J For Res 13:1117–1121

Hao Q, Meng F, Zhou Y, Wang J (2005) A transition matrix growth model for uneven-aged mixed-species forests in the Changbai Mountains, northeastern China. New For 29:221–231

Huang S, Yang Y, Wang Y (2003) A critical look at procedures for validating growth and yield models. In: Amaro A, Reed D, Soares P (eds) Modelling forest systems. CAB Internacional, pp 271–293

Hyink DM, Moser JW (1983) A generalized framework for projecting forest yield and stand structure using diameter distributions. For Sci 29:85–95

Kariuki M (2004) Modelling dynamics including recruitment, growth and mortality for sustainable management in uneven-aged mixed-species rainforests. School of Environmental Science and Natural Resources Management. Southern Cross University, p 212

Kuss O (2001) A SAS/IML macro for goodness-of-fit testing in logistic regression models with sparse data. In: 26th Annual SAS users group international conference

Leak WB, Graber RE (1976) Seedling input, death, and growth in uneven-aged northern hardwoods. Can J For Res 6:368–374

Lexerød N, Eid T (2005) Recruitment models for Norway spruce, scots pine, birch and other broadleaves in young growth forests in Norway. Silva Fenn 39:391–406

Liu J, Ashton PS (1998) FORMOSAIC: an individual-based spatially explicit model for simulating forest dynamics in landscape mosaics. Ecol Model 106:177–200

McCarthy J (2001) Gap dynamics of forest trees: a review with particular attention to boreal forests. Environ Rev 9:1–59

McCullagh P (1985) On the asymptotic distribution of Pearson's statistic in linear exponential-family models. Int Stat Rev 53:61–67

McCullagh P, Nelder JA (1986) Generalized linear models. Chapman & Hall, London

Monserud RA, Sterba H (1996) A basal area increment model for individual trees growing in even- and uneven-aged forests stands in Austria. For Ecol Manage 80:57–80

Moser JW (1972) Dynamics of uneven-aged forest stands. For Sci 18:184–191

Osius G, Rojek D (1992) Normal goodness-of-fit tests for multinomial models with large degrees of freedom. J Am Stat Assoc 87:1145–1152

Pepe MS (2000) Receiver operating characteristic methodology. J Am Stat Assoc 95:308–311

Qin Y (1998) Ingrowth models and juvenile mixedwoods stand dynamics. Department of renewable resources, University of Alberta, Edmonton, p 132

Roig S, del Río M, Ruíz-Peinado R, Cañellas I (2007) Tipología dasométrica de los rebollares (*Quercus pyrenaica* Willd.) de la zona centro de la Península Ibérica. Los sistemas forrajeros: entre la producción y el paisaje. Actas de la XLVI Reunión Científica de la Sociedad Española para el Estudio de los Pastos, Vitoria, pp 535–542

Sánchez O, Sánchez F, Carretero MP (1999) Modelos y cartografía de estimaciones climáticas termopluviométricas para la españa peninsular. INIA, Madrid, p 192

SAS/STAT (2000) SAS/STAT user's guide, version 8. SAS Institute Inc, Cary

SAS/STAT (2001) The SAS system for windows release 6.12. SAS Institute Inc, Cary

Shifley SR, Ek AR, Burk TE (1993) A generalized methodology for estimating forest ingrowth at multiple threshold diameters. For Sci 39:776–798

Snowdon P (1991) A ratio estimator for bias correction in logarithmic regressions. Can J For Res 21:720–724

Soares P, Tomé M, Skovsgaard JP, Vanclay JK (1995) Evaluating a growth model for forest management using continuous forest inventory data. For Ecol Manage 71:251–265

Trasobares A, Pukkala T, Miina J (2004) Growth and yield model for uneven-aged mixtures of *Pinus sylvestris* L. and Pinus nigra Arn. in Catalonia, north-east Spain. Ann For Sci 61:9–24

Vanclay JK (1989) A growth model for North Queensland rainforests. For Ecol Manage 27:245–271

Vanclay JK (1992) Modelling regeneration and recruitment in a tropical forest. Can J For Res 22:1235–1248

Vanclay JK (1994a) Modelling forest growth and yield—application to mixed tropical forests. CAB Internat, UK

Vanclay JK (1994b) Sustainable timber harvesting: simulations studies in the tropical rainforests of north Queensland. For Ecol Manage 69:299–320

White H (1982) Maximum likehood estimation of misspecified models. Econometrica 50:1–25