

Modeling stem increment in individual *Pinus occidentalis* Sw. trees in La Sierra, Dominican Republic

S. Bueno^{1*} and E. Bevilacqua²

¹ State University of New York. College of Environmental Science and Forestry. 1 Forestry Drive.
320 Bray Hall. Syracuse. NY 13210. USA

² Associate professor. State University of New York. College of Environmental Science and Forestry. USA

Abstract

One of the most common and important tree characteristics used in forest management decision-making is tree diameter-at-breast height (DBH). This paper presents results on an evaluation of two growth functions developed to model stem diameter increases in individual *Pinus occidentalis* Sw. trees in La Sierra, Dominican Republic. The first model was developed in order to predict future DBH (FDM) at different intervals of time and the other for predicting growth, that is, periodic annual diameter increment (PADIM). Each model employed two statistical techniques for fitting model parameters: stepwise ordinary least squares (OLS) regression, and mixed models. The two statistical approaches varied in how they accounted for the repeated measurements on individual trees over time, affecting standard error estimates and statistical inference of model parameters. Each approach was evaluated based on six goodness-of-fit statistics, using both calibration and validation data sets. The objectives were 1) to determine the best model for predicting future tree DBH; 2) to determine the best model for predicting periodic annual diameter increment, both models using tree size, age, site index and different indices of competitive status; and 3) compare which of these two modeling approaches predicts better the future DBH.

OLS provided a better fit for both of the growth functions, especially in regards to bias. Both models showed advantages and disadvantages when they were used to predict growth and future diameter. For the prediction of future diameter with FDM, accuracy of predictions were within one centimeter for a five-year projection interval. The PADIM presented negligible bias in estimating future diameter, although there was a small increase in bias as time of prediction increased. As expected, each model was the best in estimating the response variable it was developed for. However, a closer examination of the distribution of errors showed a slight advantage of the FDM against the PADIM. Based on this, it is proposed that the FDM model is used to estimate future diameter and periodic diameter increment (growth) of *P. occidentalis*.

Additional key words: diameter growth, repeated measures, mixed models, OLS, site index, indices of competitive status.

Resumen

Modelando el incremento en diámetro del fuste en árboles individuales de *Pinus occidentalis* Sw. en La Sierra, República Dominicana

Una de las características más empleadas e importantes utilizada en la toma de decisiones en el manejo forestal es el diámetro normal —DBH— de los árboles. En este artículo se presentan los resultados de la evaluación de dos funciones de crecimiento utilizadas para predecir el incremento futuro en diámetro normal (FDM) en intervalos de tiempo diferentes, y el crecimiento o incremento periódico anual en diámetro normal (PADIM) en árboles individuales de *Pinus occidentalis*, en La Sierra, República Dominicana. Se emplearon dos técnicas estadísticas para el ajuste de los parámetros: cuadrados mínimos ordinarios (OLS) por etapas de regresión y modelos mixtos. Las técnicas variaron en la forma la en que se tuvieron en cuenta las medidas repetidas sobre los árboles, afectando a la estimación de los parámetros y al error estándar. La bondad de ajuste de cada modelo fue evaluada con seis estadísticos, usando conjuntos de datos separados para la calibración y la validación de los mismos. Los objetivos fueron 1) determinar el mejor modelo para predecir el incremento futuro en DBH y 2) el incremento periódico anual en diámetro de los árboles usando en ambos modelos el tamaño y la edad del árbol, el índice de sitio y diferentes índices de competencia, y 3) comparar cuál de las técnicas de ajuste empleadas predice mejor el DBH futuro.

* Corresponding author: swbueno@syr.edu ; swbueno@gmail.com

Received: 26-05-09; Accepted: 12-11-09.

El método de los cuadrados mínimos ordinarios ofreció un ajuste mejor en la estimación del incremento en diámetro normal en ambos modelos, especialmente en lo que respecta al sesgo. La precisión de la estimación del diámetro futuro con el modelo FDM fue menor de un centímetro para la predicción del intervalo entre uno y cinco años, y de poco más de un centímetro para la predicción del diámetro a los seis años. El error en la predicción del diámetro futuro fue insignificante con el modelo PADIM, aunque se detectó un pequeño aumento del sesgo conforme aumentó el intervalo de tiempo de predicción. Según lo esperado, cada modelo fue el mejor para la predicción de la variable para la que fue diseñado. Sin embargo, un examen más exhaustivo de la distribución de los errores mostró una ligera ventaja del modelo FDM frente al PADIM, por lo que proponemos el uso del primero para la predicción del diámetro normal futuro y el incremento periódico anual en diámetro (crecimiento) de *P. occidentalis*.

Palabras clave: crecimiento en diámetro, medidas repetidas, modelos mixtos, OLS, índice de sitio, índices de competencia.

Introduction

One of the most common and important tree characteristics used in forest management decision-making is tree diameter-at-breast height (DBH). This variable has numerous beneficial attributes, including being easy to measure (Zhang *et al.*, 2004) and having strong correlations with other tree characteristics. The distribution of trees by DBH class allows foresters and ecologists to understand stand structure, stand dynamics, and future forest yield. Individual-tree diameter growth models are among the most basic and essential components of forest growth models (Sánchez-González *et al.*, 2006). They allow one to project and describe the state of a tree at some future time.

Many forest modeling systems that focus on individual-tree growth functions are designed to produce detailed tree and stand information. Compared with stand-class models, individual-tree models ensure reliable predictions for a wide range of tree sizes, sites and stand condition (Zhao *et al.*, 2004). Individual-tree models are capable of predicting the growth of an average tree of a given size for a specified stocking level and site productivity. Detailed information about stand dynamics and structure can be derived from models based on individual tree-growth. A major advantage of single-tree distance independent models lies in eliminating the need for stand tables and having a much faster computer execution time. The only serious disadvantage to individual-tree distance-independent models

is that they are incapable of predicting the growth of a specific individual tree with any reliability (low R^2 values).

Modeling systems with a focus on individual trees commonly consist of several basic components, such as: 1) a diameter-growth component; 2) a height-growth component; 3) a tree crown component, generally related to tree vigor; and 4) a mortality component, which may be stochastically generated or predicted as a function of growth rate (Avery and Burkhart, 2002). Each tree is modeled by simulating its growth in diameter, height and crown, and then deciding whether it survives or dies. Subsequently, stem volume growth is calculated. Aggregating tree lists provides per-hectare volumes and levels of production (Davis and Johnson, 1987).

Explanatory variables used in tree growth models attempt to capture three fundamental biological components known to affect tree growth: a tree's current age, size or stage of development; site productivity; and competitive status (Alders 1979; Vanclay, 1994). Diameter increment increases to a maximum in the early stages of a tree's life, then decreases, and finally approaches zero when a tree approaches senescence. Greater increments occur in trees with long and healthy crowns. Many growth and yield models include age as predictor of tree growth. For even-aged stands, the average age of dominant and codominant trees is a reasonable predictor variable, but is not meaningful for uneven-aged stands. In the absence of age, current tree size is often a substitute, although this prevents us from

Abbreviations used: B: absolute bias, B%: relative bias, BAL_0 : cumulative basal area of trees larger than subject at first measurement ($t=0$), $BAPH_0$: calculated basal area per hectare at first measurement ($t=0$), CCI_0 : ratio of the subject basal area to average basal area at first measurement ($t=0$), DBH_0 : diameter-at-breast height at first measurement ($t=0$), DR: Dominican Republic, D_t : future diameter at end of measurement period (time t), FDM: future diameter model, MAD: mean absolute deviation, MSE: mean square error, OLS: ordinary least squares, PADI: periodic annual diameter increment (cm yr^{-1}), PADIM: periodic annual diameter increment model, R^2 : coefficient of determination, RMSE: root mean square error, SDI_0 : calculated Reineke's site density index at first measurement ($t=0$), SI: site index (base age 40), t : elapsed time since first measurement (years), TPH_0 : calculated trees per hectare at first measurement ($t=0$).

separating past cumulative competitive interactions from current stage of development. For even-aged stands, it is generally accepted that smaller trees have less diameter growth than larger trees, and trees in stands with low basal area exhibit greater growth than trees in stands with higher basal area (Knowe *et al.*, 1997). The most significant predictor of a tree growth is its current size, which is a reflection of its past competitive interactions (Bevilacqua, 1999). Site productivity is often quantified using site index (Wycoff, 1990). In a similar way, site index may not be appropriate in even-aged stands (Schroder *et al.*, 2000).

Trees growing in the open or dominant positions in closed stands are also expected to grow faster than trees of the same size but growing in dense stands or in subordinate crown positions (Wycoff, 1990). According to Munro (1974), there are two recognized philosophies for quantifying competitive position within individual tree models: distance-dependent and distance-independent. Distance-dependent models require spatial information of the trees which are assigned a coordinate location. The competitive status of each individual is modeled as a function of its size and that of the competitors, weighted by the distance separating the trees. Distance-independent models determine competitive position by comparing the relative size and condition of the subject tree to various stand characteristics, such as basal area and/or average diameter. Distance-independent growth models assume that spatially, all sizes are uniformly distributed throughout the stand (Davis and Johnson, 1987) and estimate tree growth individually or by size classes.

Even though individual tree increment depends to a great extent on the vigor of the subject, it also depends on the competition with other trees. Only individual-tree models have the ability to simulate the competitive environment around each tree (Davis and Johnson, 1987). Competition tends to create negative dependence in size growth of individuals in spatial proximity, while micro-site variation tends to create positive dependence as neighbors are subject to similar environmental conditions (Fox *et al.*, 2007). The development of an index or proxy for competitive status has received the most attention in the recent literature. Through the use of competition indices, modelers attempt to quantify the influence of neighbors on the growth of individual trees in a forest stand.

All competition indices assume that the level of competitive stress around a tree can be quantified by taking account of the number of competitors and their

size within a defined neighborhood (Bevilacqua, 1999). Competition indices may be expressed in absolute or relative terms. They can also be spatially related or not. Individual-tree growth models predict basal area increment over a given period of time from variables such as tree size, vigor, age and competitive stress measured at the beginning of the period (Wimberly and Bare, 1996).

Competition indices based on stem size include: 1) the ratio of subject tree size to average tree size (Davis and Johnson, 1987); and 2) the cumulative basal area of trees larger than the subject tree (BAL). As the ratio of the subject tree basal area to the average tree basal area gets numerically larger, the vigor of the tree is supposedly greater and the grow rate approaches the genetic potential. BAL has been useful in predicting diameter increment, and should be considered complementary to stand basal area. As BAL decreases, the predicted increment increases.

There are two commonly used modeling strategies: 1) the formulation of a potential or average growth component combined with a modifier function and; 2) a composite model, with a single equation that models diameter growth as a function of tree and stand level predictor variables (Murphy and Shelton, 1996; Tang *et al.*, 1997; Westfall, 2006). Studies have been carried out to compare the suitability of these two variants for modelling growth of individual trees, and their performance has been fundamentally the same (Zhang *et al.*, 2004).

Many individual-tree growth models use the change in diameter over a specific time period as the response variable. However, modelling diameter increment is not the only alternative to predicting tree growth. Other variables have been modeled including basal area increment, future basal area and future diameter. All these are alternatives for estimating increase in stem size. They are mathematically related and little differences in the outcome of the modeling process are expected (Vanclay, 1994) if the assumptions regarding the error term are met (Zhang *et al.*, 2004).

The decision to model diameter or basal area increment in growth studies has been subjective. Many authors have used diameter increment (Cao, 1999; Lessard *et al.*, 2001; Bragg, 2003), while others have chosen basal area increment (Opie, 1968; Sadiq, 1982; Quicke *et al.*, 1994; Mailly *et al.*, 2003; Zhang *et al.*, 2004). Diameter increment and its integral are closely related to many stand characteristics and can be measured more easily and reliably than other increments.

Mathematically, two trees having the same diameter increment might differ in basal area increment, if one has a greater initial diameter, it will have a larger basal area increment (Zeide, 2001).

Potential predictor variables for individual tree basal area increment are DBH and/or some transformation of DBH, alone or in combination (Zhao *et al.*, 2004). When used in combination, the coefficients of these variables have both positive and negative values, confirming the assumed relationship that, for a given set of stand conditions, the basal area increment increases to a certain level and then decreases thereafter (Zhao *et al.*, 2004).

Predictive equations are often calibrated with re-measurement data from permanent plots having individually identified trees. Linear regression analysis, with specific assumptions regarding the error term — *i.e.*, normality and independence— has often been the preferred statistical method for modeling growth. However, the nature of the data often nullifies those assumptions and biological processes that influence tree growth are generally nonlinear in nature. Therefore, different statistical techniques have been developed for estimating the parameters of linear and nonlinear growth and yield models including mixed models, simultaneous equations, log-linear regression and seemingly unrelated regression (Zhang *et al.*, 2004).

Often, multiple observations are obtained from the same sampling unit (Littell *et al.*, 1996) over a period of time and/or space. The nature of repeated measures experiments is often ignored and independence between observations is assumed (Zhang *et al.*, 2004; Uzoh and Oliver, 2008). However, two measurements taken at adjacent units of time and/or space are more highly correlated than two measurements taken several time points or space apart. High autocorrelations can make two mutually exclusive variables appear to be related. While ignoring the least-squares regression assumptions of independence still results in unbiased estimates of the model parameters, estimates of model errors are biased (Uzoh and Oliver, 2008). If error estimates are biased, inferences of model coefficients are in question.

Covariates at the tree and stand level, such as DBH, density measures, site index and competition indices, are included in the model as fixed effects. Under such circumstances, the appeal of using mixed models is to obtain improved estimates of model variance when observations are correlated. To account for an appropriate error term, methods based on mixed models with

special parametric structures on the covariance matrices are applied. The specification of covariance structures addresses the bias in the standard error of parameter estimates (Zhang and Gove, 2005). The most commonly used covariance structures for modelling repeated measurements data are compound symmetric, first-order autoregressive and unstructured (Littell *et al.*, 1996). The most appropriate goodness-of-fit statistic used to evaluate and compare models includes Akaike's and Baye Information Criteria.

The objectives of this paper are to: 1) determine the best fitting non-linear regression model for predicting DBH over time in *P. occidentalis* trees using current size, elapsed time, site index and different indices of competitive status as predictor variables; 2) determine the best fitting non-linear regression model for predicting periodic annual diameter increment using current size, age, site index and different indices of competitive status as predictor variables; and 3) compare which of these two modelling approaches predicts better the future DBH.

Materials and methods

The study area is a region of approximately 1,800 km² in the north central portion of Cordillera Central, Dominican Republic. The data available for model development were from 25 natural *Pinus occidentalis* stands in three different life zones in La Sierra region, Dominican Republic. Nine stands are in the humid zone, 6 in the intermediate zone, and 10 stands in the dry region. The humid-life zone is denominated formally as Subtropical Very Humid Forest; the intermediate zone located between the two previously mentioned zones, is called Subtropical Humid Forest and; the dry life zone corresponds to the formal denomination Subtropical Dry Forest (Holdridge, 1967).

In each stand, one permanent rectangular plot was located. Individual trees in each plot were marked for identification. The diameter at breast height (DBH, cm) outside bark was measured to the nearest 0.1 cm for all trees larger than 5 cm. The individual tree diameters were measured with a diameter tape every measurement year. Total tree height (m) was measured with a hypsometer on each tree each year. Permanent plot data were used because they provide a direct measure of individual tree DBH growth over time and are considered the best kind of DBH growth data. Every effort was made to cover a wide range of

stand conditions. Plots selected for sampling were unburned and appear free of damages from insects and or fungi.

The plots were established at random from 1984 through 1991. See Table 1 for plot details. The youngest stand measured was 21 years in its first measurement year (1988), and the oldest stand was 46 years, also in its first measurement year. The growth and yield data spanned an age range of 25 years. The stands ranged in density from 192 to 950 stems per ha; from 9.26 to 33.39 m² in basal area per hectare; site density index (Reineke) range from 91.18 to 273.22; and site index (40 year index age) from 13 to 30.

The data for each life zone were separated into two parts, with one used for model estimation and the second for model validation. The estimation data set had a total of 830 trees and the validation data set 200

trees. The combined data set includes 4,362 DBH measurements for an average of 4.2 measurements per tree.

Model development

To model change in DBH, two approaches were employed: (1) prediction of future diameter (FDM); and (2) prediction of periodic annual diameter increment (PADIM). The predictor variables used for model building were related to stand competition, productivity, and individual tree size. Stand competition was measured using variables such as calculated number of trees per hectare (TPH), basal area per hectare (BAPH, m² ha⁻¹), Reineke's site density index (SDI). Two individual-tree competition indices were also computed, cumulative basal area of trees larger than the subject

Table 1. Summary statistics on sample plots used for growth measurements

Plot ID	Size (ha)	SDI	TPH	SI	BA	Measurement year		Interval (years)	Number of measurements
						Initial	Final		
<i>Dry zone</i>									
101	0.1000	29.12	250	22	14.5	1984	1989	5	6
102	0.1000	24.99	240	13	12.3	1984	1991	7	8
103	0.1000	29.80	440	20	13.8	1987	1990	3	4
108	0.1000	34.91	660	21	16.2	1989	1994	5	4
109	0.1000	29.60	580	18	13.7	1989	1994	5	4
110	0.1000	48.59	950	23	22.4	1989	1994	5	4
111	0.1000	20.70	580	15	09.3	1989	1994	5	4
112	0.1000	28.63	740	15	12.9	1989	1994	5	4
115	0.1000	32.45	350	21	15.8	1991	1995	4	2
116	0.1000	37.12	420	25	18.0	1991	1995	4	2
<i>Intermediate zone</i>									
1	0.1000	42.13	470	29	20.5	1988	1994	6	5
3	0.1250	40.14	336	28	15.7	1988	1994	6	5
5	0.1250	43.39	352	22	17.0	1988	1994	6	5
6	0.0625	17.01	192	23	13.8	1988	1991	3	4
11	0.0625	23.40	656	20	17.5	1988	1990	2	3
12	0.0625	25.06	672	20	18.7	1988	1990	2	3
<i>Humide zone</i>									
8	0.0625	39.48	832	27	30.2	1988	1994	6	5
9	0.0625	30.31	720	25	22.9	1988	1994	6	5
10	0.0625	31.73	528	30	31.3	1988	1991	3	4
13	0.0625	27.14	576	24	20.8	1988	1994	6	5
14	0.0625	30.07	656	23	22.9	1988	1994	6	5
15	0.0625	24.64	304	28	19.8	1988	1994	6	5
16	0.0625	35.56	400	31	28.8	1988	1994	6	5
17	0.0625	24.83	544	27	18.9	1988	1993	5	3
18	0.0625	30.30	592	25	23.8	1988	1991	3	4

tree (BAL), and the ratio of the subject tree basal area to average basal area (CCI). As a measure of site productivity, site index (base age 40) was used. The individual tree size was represented by DBH.

These predictor variables, alone and in combination, were tested for predicting future diameter (D_t) and periodic annual diameter increment (PADI) using both OLS and mixed models analysis procedures. The mixed effects model approach addresses the problem of correlations among observations from repeated measurements on the same experimental unit at different points in time. Such correlations were accounted for in the model-fitting process by incorporating random effect parameters in the model. Using a mixed-effects approach allows the specification of the variance-covariance matrix to estimate model parameters. The coefficients vary from tree to tree, providing a model suitable to each tree in the sample (Westfall, 2006) and increasing the accuracy of predictions of future DBH and diameter increment based on current stand conditions.

A mixed model includes both fixed and random effects components. Between plots, between and within-tree differences were accounted for by including random effect parameters specific at those levels. Two different sizes of experimental units could be differentiated: 1) a spatial unit which is an individual tree; and 2) a set of temporal units that are the repeated measurements on individual trees. We also assumed that the increment of an individual tree is also dependent on its competitive status relative to neighboring trees, and the impact of management.

The statistical procedures employed to get the parameters of both models (FDM and PADIM) are based on the fact that the stochastic component of growth variability is a consequence of different factors acting simultaneously. Considering that the effect of some of these unobservable factors remains constant for a given period, and then it is possible to calibrate future increment by introducing into the model the stochastic effects predicted for a prior period.

In general, a linear mixed model, including both fixed and random components had the general expression for the multilevel linear mixed model proposed was:

$$y = X\beta + Zb + \varepsilon \quad [1]$$

where y is a n -dimensional vector including the n observations for the response variable; X is a $n \times p$ design matrix, including covariates and terms associated with fixed parameters of the model; β is a p -dimensional vector of fixed parameters of the model; Z is a $n \times q$ design matrix for the random components of the model;

b is a q -dimensional vector of random components acting at tree, plot and period levels; and ε is a n -dimensional vector of conditional (Littell *et al.*, 1996).

With the PADIM, the dependent variable was periodic annual diameter increment (PADI). The independent tree and stand variables were diameter at the start of the measurement period (D_0), two measures of competition (CCI and BAL) calculated also at the start of the measurement period, three density variables (TPH, BAPH and SDI) and a site quality measure (SI). The model was as follows.

$$PADI = \beta_0 + \beta_1 DBH_0 + \beta_2 CCI_0 + \beta_3 BAL_0 + \beta_4 TPH_0 + \beta_5 BAPH_0 + \beta_6 SDI_0 + \beta_7 SI + \varepsilon \quad [2]$$

where PADI = periodic annual diameter increment (cm yr⁻¹), DBH₀ = diameter-at-breast height at start of measurement year, CCI₀ = ratio of the subject basal area to average basal area at start of measurement year, TPH₀ = calculated trees per hectare at start of measurement year, BAPH₀ = calculated basal area per hectare at start of measurement year, SDI₀ = calculated Reineke's site density index at start of measurement year, BAL₀ = cumulative basal area of trees larger than subject at start of measurement year, SI = site index (base age 40)

As with periodic annual diameter increment, to obtain the best equation to predict future diameter, the same two modeling procedures as above were explored, namely OLS and mixed linear models. In each procedure, the dependent variable was diameter at the end of a measurement period. The independent tree and stand variables were diameter at the start of the measurement period (D_0), time (linear and quadratic terms), two measures of competition (BAL and CCI) also calculated at the start of the measurement period, three stand density variables (TPH, BAPH and SDI), the interactions of time with the competition and density variables, and a site quality measure (SI). The model was as follows.

$$D_t = \beta_0 + \beta_1 DBH_0 + \beta_2 t + \beta_3 t^2 + \beta_4 CCI_0 + \beta_5 BAL_0 + \beta_6 TPH_0 + \beta_7 BAPH_0 + \beta_8 SDI_0 + \beta_9 SI + \beta_{10} t \times CCI_0 + \beta_{11} t \times BAL_0 + \beta_{12} t \times TPH_0 + \beta_{13} t \times BAPH_0 + \beta_{14} t \times SDI_0 + \beta_{15} t^2 \times CCI_0 + \beta_{16} t^2 \times BAL_0 + \beta_{17} t^2 \times TPH_0 + \beta_{18} t^2 \times BAPH_0 + \beta_{19} SDI_0 + \varepsilon \quad [3]$$

where D_t = estimated diameter at end of measurement period (time t), DBH₀ = diameter-at-breast height at first measurement ($t=0$). CCI₀ = ratio of the subject basal area to average basal area at first measurement ($t=0$), TPH₀ = calculated trees per hectare at first measurement ($t=0$), BAPH₀ = calculated basal area per hectare at first measurement ($t=0$), SDI₀ = calculated Reineke's site density index at first measurement ($t=0$), BAL₀ =

cumulative basal area of trees larger than subject at first measurement ($t=0$), t = elapsed time since first measurement (years).

Four variance-covariance matrices were used for the mixed model methods to estimate the regression coefficients: compound symmetry, autoregressive, unstructured specified in the repeated statement and unstructured specified with a random statement.

Non-significant variables were dropped from the resulting models developed above.

Finally, the best of the two approaches (mixed models and OLS) for each dependent variable, the periodic annual diameter increment and future diameter, were quantitatively evaluated using independent verification data by examining the distribution, bias and precision of residuals to determine the accuracy of model estimations (Vanclay, 1994). Mean square error (MSE), root mean square error (RMSE), pseudo coefficient of determination (R^2), absolute and relative bias (B, B%), and mean absolute deviation (MAD) were calculated as follows:

$$MSE = \frac{\sum_{i=1}^n (D_i - \hat{D}_i)^2}{n - m} \quad [4]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (D_i - \hat{D}_i)^2}{n}} \quad [5]$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (D_i - \hat{D}_i)^2}{\sum_{i=1}^n (D_i - \bar{D})^2} \quad [6]$$

$$B = \frac{\sum_{i=1}^n (D_i - \hat{D}_i)}{n} \quad [7]$$

$$B\% = 100 \times \frac{\left| \sum_{i=1}^n \frac{(D_i - \hat{D}_i)}{n} \right|}{\bar{d}} \quad [8]$$

$$MAD = \sum_{i=1}^n \frac{|(D_i - \hat{D}_i)|}{n} \quad [9]$$

Results

Modeling periodic annual diameter increment

Estimating PADIM with the mixed model procedure resulted in a model that includes D_0 , CCI, BAL and SI

as significant predictor variables. The best variance-covariance structure among those evaluated was the unstructured error structure (UN) specified as repeated statement, producing the smallest AIC, AICC and BIC values. After performing stepwise regression and dropping all non-significant variables, the estimation of PADI by OLS resulted in three measures of competition and site index as significant predictor variables.

Based on six goodness-of-fit statistics, the best approach to estimate PADI was the OLS procedure. In the estimation and validation phases, respectively, mean square errors were 0.069 and 0.067 (cm) and relative biases were 29.1% and 30.2% for the mixed model. The corresponding quantities of MSE, pseudo- R^2 and relative bias for the OLS procedure were 0.055 and 0.55 (cm); 0.165 and 0.170; and 0.1% and 4.7%, respectively.

Modeling future diameter

The dependent variable of the future diameter model (FDM) model was the DBH achieved at each remeasurement. For a particular tree, time between two measurements was irregular. Most trees were measured every year for the first three years and then the intervals were two years in one plot, three years in 14 plots, and 4 years in two plots. The remaining eight plots were measured every year. In the initial model tested [3], interaction terms were included to measure the joint effect of initial diameter and the remaining independent variables. Interaction effects occur when independent variables not only have separate effects but have also combined effects on the dependent variable.

Estimating future diameter with the mixed model procedure resulted in 18 variables being statistically significant (p value < 0.05), being the best variance-covariance among those evaluated the unstructured error structure (UN) specified as repeated statement. It produced the smallest AIC, AICC and BIC values. Stepwise OLS regression performed on the independent variables to estimate future diameter resulted in a model with 18 independent predictor variables. Parameter estimates, standard errors of estimates and p -values associated with each of the estimates are displayed in Table 2.

As with periodic annual diameter increment, the best modeling approach to estimate future diameter was chosen based on six goodness-of-fit statistics. The comparisons of both models in regards to each of the

Table 2. Parameter estimates, standard errors and p-values to estimate future diameter using ordinary least squares procedure

Variable	Estimate	Standard error	Probability
T	2.3045	0.13704	<0.0001
t ²	-0.3219	0.02412	<0.0001
D ₀	0.8861	0.00661	<0.0001
CCI	1.9533	0.09128	<0.0001
t*CCI	-1.3651	0.07257	<0.0001
t ² *CCI	0.2507	0.01265	<0.0001
BAL ₀	0.0805	0.00780	<0.0001
t*BAL ₀	-0.1431	0.00674	<0.0001
t ² *BAL ₀	0.0246	0.00114	0.0001
BAPH	-0.0354	0.00787	<0.0001
t*BAPH	0.1388	0.00517	<0.0001
t ² *BAPH	-0.0245	0.00085	<0.0001
TPH	-0.0015	0.00019	<0.0001
t*TPH	-0.0011	0.00013	<0.0001
t ² *TPH	0.0001	0.00002	<0.0001
SDI	0.0059	0.000625	<0.0001
t*SDI	-0.0053	0.000590	<0.0001
t ² *SDI	0.0008	0.000098	<0.0001

statistics in the estimation and validation data sets are shown on Table 3.

Use of FDM to estimate growth

The best FDM obtained by means of performing stepwise regression was employed in estimating periodic annual diameter increment. First, future diameter was predicted at the end of every measurement interval (one-year, three-year and four-year intervals) and then subtracted from the diameter at the start of the interval. The resulting diameter increment was then compared with the observed corresponding diameter increment. Goodness-of-fit statistics were computed and compare

with their corresponding counterparts calculated by using the best PADIM (OLS).

To estimate periodic annual diameter increment in one-year measurement intervals, the PADIM was best in terms of MSE, RMSE, and MAD. Bias and relative bias were lower for the FDM (Table 4). For three and four-year measurement intervals the PADIM was better than the FDM in all statistics.

The distribution of errors for predicted values of periodic annual diameter increment in the three different measurement intervals did not show any conspicuous differences when plotted against the predicted increment by both the FDM and PADIM.

Likewise, the scatter plot of MSE and bias averaged values against initial averaged diameter by four cm diameter classes, showed minimum differences between these two models for predicting growth of *P. occidentalis* (Figs. 1 and 2).

Use of PADIM to estimate yield

The best PADIM obtained by OLS was used for estimating future diameter by first predicting periodic annual diameter increment, and then adding the prediction to the diameter at the start of the measurement period. Six sets of future diameter were computed. The resulting diameters were then compared with their corresponding observed diameters and goodness-of-fit statistics were computed and compared with the corresponding goodness-of-fit statistics calculated by using the best FDM.

For estimating diameter one year hence, the PADIM shows a slight advantage over the FDM in all statistics. From two to six years hence, the FDM is best in all but one goodness-of-fit statistics (MAD in two-year hence estimation) (Table 4). From two to six years, most of the errors corresponding to the PADIM are allocated

Table 3. Comparisons of the mixed and ordinary least square models based on goodness-of-fit statistics from the estimation and validation phases, to estimate future diameter

Statistic	Estimation data		Validation data	
	Mixed model	OLS model	Mixed model	OLS model
MSE	12.2156	0.4108	15.257	0.075
RMSE	3.4951	0.6409	3.906	0.037
Bias	0.5863	-0.0610	0.199	0.407
B%	2.8456	0.2959	0.905	1.853
MAD	2.7170	0.4489	15.257	1.075
R ²	0.7575	0.9918	0.716	0.980

Table 4. Indicators of fit for estimating periodic annual diameter increment for one, three and four-year diameter increment, using both the FDM and PADIM

Statistic	Time period		
	1-yr	3-yr	4-yr
<i>Period annual diameter increment prediction by FDM</i>			
MSE	0.1166	0.0529	0.0403
RMSE	0.3415	0.2299	0.2007
Bias	-0.0075	0.0905	-0.1375
B%	0.0061	30.5908	60.5447
MAD	0.2694	0.1833	0.1774
R ²	N/A	N/A	N/A
<i>Periodic annual diameter increment prediction by PADIM</i>			
MSE	0.0609	0.0390	0.0254
RMSE	0.2468	0.1974	0.1593
Bias	0.0234	0.0233	-0.0809
B%	7.5518	7.8829	35.3411
MAD	0.1854	0.1466	0.1286
R ²	0.1591	0.2009	N/A

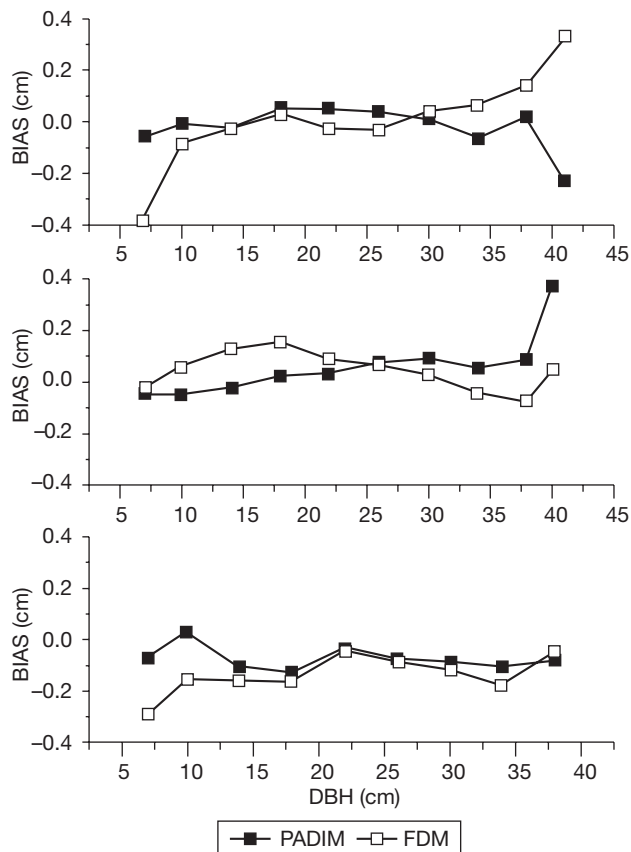


Figure 1. Bias behavior versus initial observed diameter values for periodic annual diameter increment prediction by both OLS-FDM and OLS-PADIM (one-year, three-year and, four-year periodic annual increments).

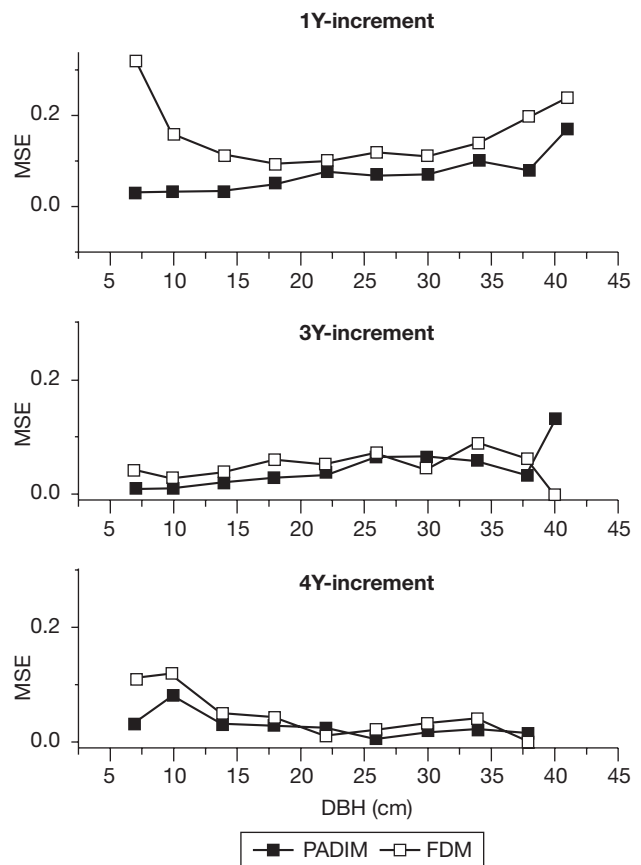


Figure 2. MSE behavior versus initial observed diameter values for periodic annual diameter increment prediction by both OLS-FDM and OLS-PADIM (one-year, three-year and, four-year periodic annual increments).

above the zero reference line, indicating that this model underestimates future diameter in a monotonic increasing manner as the year of prediction increases. Bias and MSE are also larger for the PADIM and also increase from year two to year six when plotted against initial diameter (Fig. 3 and 4).

Discussion

Diameter growth and yield prediction is one of the primary components of individual-tree growth models. These models allow detailed analyses on stand structure, but need additional equations to describe other components (e.g., tree mortality and recruitment) of tree or stand growth to make a complete stand or tree projection system. Two statistical techniques, stepwise OLS regression and mixed models were evaluated for predicting both future diameter and PADI.

Table 5. Indicators of fit for estimating future diameter one to six years hence, using both the FDM and PADIM

Statistic	Time period					
	1-yr hence	2-yr hence	3-yr hence	4-yr hence	5-yr hence	6-yr hence
<i>Future diameter prediction by FDM</i>						
MSE	0.1582	0.3408	0.5761	0.7867	0.7159	1.0784
RMSE	0.3977	0.5838	0.7590	0.8870	0.8461	1.0385
Bias	-0.0620	0.0105	-0.1006	-0.3841	-0.2115	0.1105
B%	0.3062	0.0506	0.4200	1.5739	1.1179	0.4330
MAD	0.2744	0.4542	0.5938	0.7390	0.6328	0.8435
R ²	0.9967	0.9929	0.9869	0.9877	0.9855	0.9740
<i>Future diameter prediction by PADIM</i>						
MSE	0.1304	0.3617	1.0135	1.4639	2.1034	4.5195
RMSE	0.3612	0.6014	1.0067	1.2099	1.4503	2.1259
Bias	0.0063	0.3472	0.7191	1.0029	1.2068	1.7232
B%	0.0312	1.6791	3.0022	4.1097	6.3785	6.7536
MAD	0.2082	0.4159	0.7561	1.0278	1.2246	1.7464
R ²	0.9973	0.9925	0.9827	0.9826	0.9588	0.9364

To validate which approach was the best, we randomly partitioned the data into a model estimation and validation component (80:20 split), and compared the observed and predicted DBH at last measurement for these observations, as well as the periodic annual diameter growth over the total measurement period, using the best equations from the resulting models.

The best modelling approach to estimate PADI was chosen based on the six goodness-of-fit statistics. Between the two alternatives, mixed and ordinary least squares models, the latter resulted in a better fit, especially in regards to bias and relative bias. These were 300 and 7 times smaller in the estimation and validation data sets, respectively.

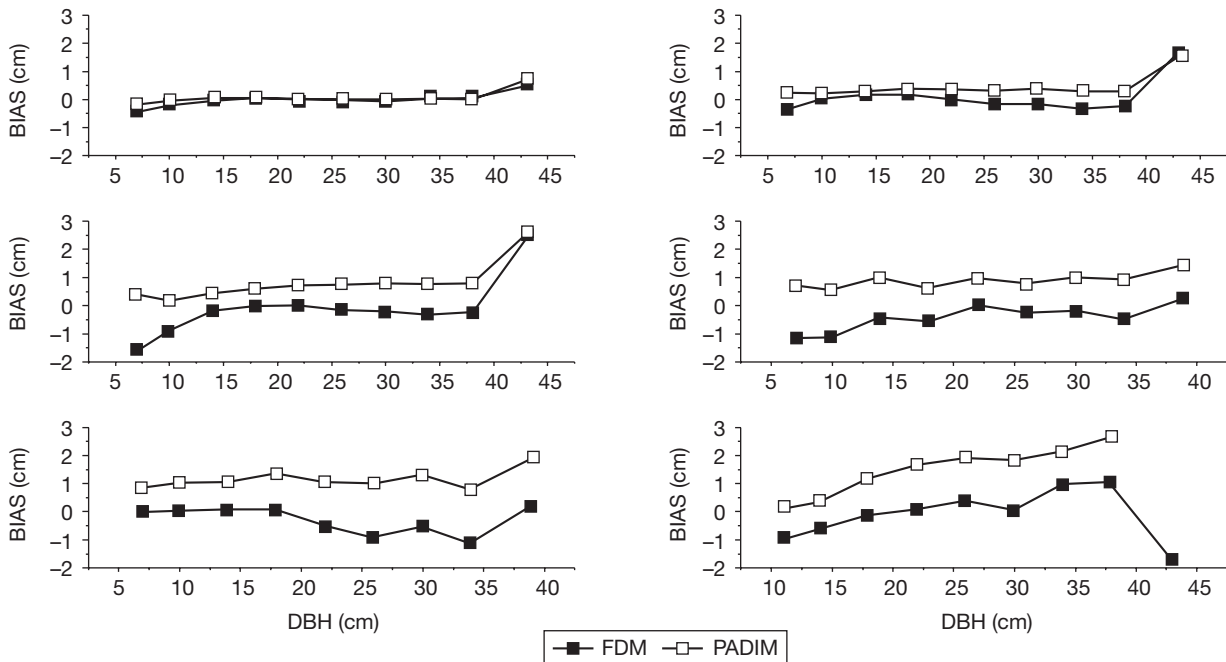


Figure 3. Bias behavior versus initial observed diameter values for future diameter prediction by both OLS-FDM and OLS-PADIM for one, two, three, four, five and, six years hence.

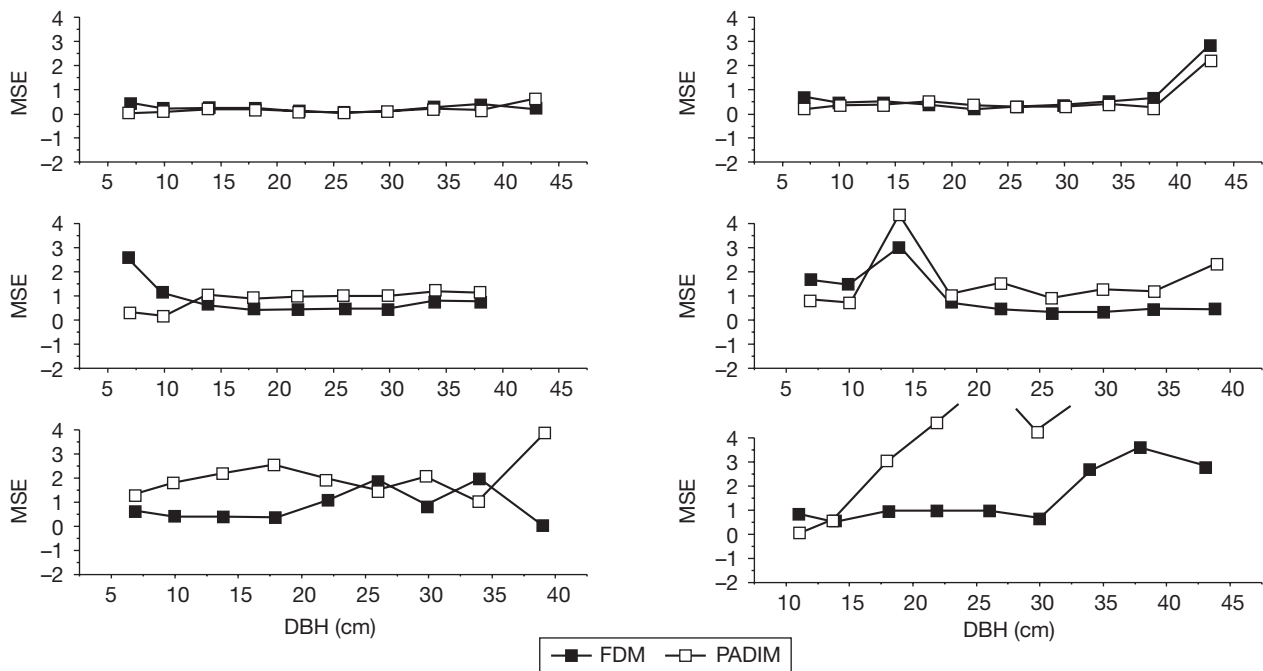


Figure 4. MSE behavior *versus* initial observed diameter values for future diameter prediction by both OLS-FDM and OLS-PADIM for one, two, three, four, five and six years hence.

The remaining significant parameters in the stepwise selection procedure to predict periodic annual diameter increment for *P. occidentalis* in the study area are: initial diameter at the start of the growing period (D_0), basal area larger than the subject tree (BAL), the ratio of the subject basal area to average plot basal area (CCI), and Reineke's site density index (SDI).

To estimate future diameter, once again, OLS was the best modelling approach. Here, MSE was particularly different between the approaches, being almost 30 and 14 times smaller for the OLS procedure in both estimation and validation phases. Bias was 10 and 2 times bigger for the mixed model in both phases. The pseudo-coefficient of determination (R^2) was very high for the OLS approach (0.99 and 0.98) in both estimation and validation, as compared to the corresponding quantities in the mixed procedure (0.75 and 0.71, respectively). The model resulted with eighteen significant predictor variables selected by the stepwise procedure (Table 3).

Both models showed advantages and disadvantages when they were used to predict growth and yield of *P. occidentalis* in La Sierra, D.R. As expected, the FDM was the best in predicting yield and the PADIM was the best in predicting growth, although by closely examining the distribution of errors, one can notice a slight advantage of the FDM. Three of the desirable proper-

ties of good estimators were present in both models; precision was less than one centimeter while bias was less than 0.5 cm. The difference in precision as measured by the coefficient of variation between the two models was negligible.

For the prediction of future diameter with FDM, precision measured (RMSE and MAD) less than one centimeter in all years except for predictions six years hence (RMSE = 1.03 cm). The PADIM also presented negligible bias (mean error) in estimating future diameter, although it increased as time of prediction increased. All goodness-of-fit statistics except pseudo R^2 displayed this monotonic increase as the prediction moved from one to six years hence.

Although the prediction of periodic annual increment by the FDM showed larger residuals quantities, relative bias for three- and four-year intervals and goodness-of-fit statistics were very similar in value to those with the PADIM. In predicting growth one-year hence, FDM was better than the PADIM in terms of bias and relative bias. Figure 1 shows that bias was larger for the PADIM when predicting periodic annual diameter increment in small trees for one-year measurement intervals, and was larger for the FDM in larger trees. In predicting periodic annual diameter increment for three-year measurement intervals with the PADIM, bias was larger for trees over 40 cm.

For trees <20 cm in diameter, bias was larger for the PADIM when the measurement interval was four years. MSE was conspicuously large for predicting periodic annual diameter increment of trees 15 cm with the FDM in one- and four-year measurement intervals, and larger for 40 cm trees in three-year measurement intervals.

For predicting future diameter (yield) one-year hence, the PADIM was better according to all goodness-of-fit statistics. Future diameter predictions from two to six years hence were better predicted by the FDM as indicated by goodness-of-fit statistics, the distribution of errors around the zero reference line, and the behavior of bias and MSE.

Each model was the best in estimating the response variable it was developed for. But if a decision to keep just one model would have to be made, the goodness-of-fit statistics showed that in estimating future diameter, there is an increasing trend in bias and MSE presented by the PADIM. The error distribution for this model was also concentrated mostly above the zero reference line, indicating that there is a systematic underestimation of future diameter. In predicting periodic annual diameter increment for the three different measurement intervals, goodness-of-fit statistics for both models were very similar with few exceptions. The distribution of errors was quite similar around the zero reference line, and MSE and bias were very close except for small and large trees. In addition, when predicted values for future diameter are plotted against observed values, based on the above, it is recommended that the future diameter model is used to estimate both future diameter (yield) and periodic diameter increment (growth) of *P. occidentalis*, if necessary.

Predictor variables in FDM

The predictor variables in this model were related to time of measurement (*i.e.* t , t^2), tree size (*i.e.* DBH), tree status (*i.e.* CCI), competition index (*i.e.* BAL), and stand density (*i.e.* BAPH, TPH, SDI). These variables are commonly included as predictors in single-tree diameter growth and yield models. Site productivity (*i.e.* SI) was initially included in the model and dropped by the stepwise regression procedure. Eighteen predictor variables were statistically significant, including the linear and quadratic time terms and DBH. The remaining variables were CCI, BAL, BAPH, TPH and SDI by themselves, and each with their interaction with both time terms.

The future diameter (yield) model supports the hypothesis that time is an important factor as a complement to other predictor variables to describe the competitive status and future yield potential. The linear term (with a negative coefficient) and quadratic terms imply that the model is capturing a non-linear trend in the DBH growth. Since most of the sites are relatively poor, the model may be capturing a gradual deceleration in growth, even for the short time period represented by the data. Slow growth was noticed during the first few remeasurements, with little or no growth in the last. Some considerations can be hypothesized as responsible for the absence of growth.

Subsequent to the data collection for this study, many of the trees in plots located in the dry and intermediate zones started to deteriorate. This manifested in the drop of the needles and the reduction of their length, and finally, the death of many trees (Torres J.G., personal communication). It was thought that the problem was related with a plague (probably a fungus) affecting only the needles. Another hypothesis is that the roots were being affected and were unable to assimilate water and nutrients by the attack of two fungi, *Leptographium serpens* and *Phytophthora cinamomi*. Two conclusions can be reached from the above statement: 1) a considerable number of the plots were affected at least in an initial stage by fungi on the roots affecting nutrients and water assimilation, and limiting growth; and 2) in the two years previous to the last measurement, annual precipitation was considerably lower than average, which, along with the effects caused by the fungi in the roots of many of the trees, may have reduced the radial growth.

Tree size influence on diameter increment was expressed by initial DBH, which in itself is an expression of competition within a plot or stand, as current tree size is a good indicator of future growth. It reflects past competitive interactions. Current DBH provides information about the developmental history of a tree. Trees in an even-aged stand having a large DBH have experienced less competitive stress in the process of stand development, relative to trees with smaller DBH that have been oppressed. If a significant alteration of stand structure does not occur, the competitive status for a tree of a given size will remain relatively similar in the near future. Accordingly, DBH can be as effective as more complex competition indices at predicting future growth rates. In our study, a positive change in diameter was greater for larger diameter classes.

As the basal area of larger trees (BAL) decreases, the tree is increasingly competitive. BAL is included

in three terms in the FDM (Table 2). Its influence has to be evaluated by taking into consideration the combined effect of all three terms on the response of future diameter. The combined effects of these terms for times different than zero diminish future diameter of subject trees as BAL increases. This fact is an indication that these trees are being subjected to predominantly asymmetric (one-sided) competition.

For CCI, which represents the competitive status as a tree gets larger, the tree is considered to be progressively more vigorous and will grow at rates closer to its genetic potential. CCI is also represented in the FDM by three terms (Table 2). The combined effects of the terms in CCI provoked an increased in tree growth as the variable increased. The coefficients for the BAPH variables, which included the terms BAPH, t^*BAPH and t^2*BAPH , were negative, positive and negative, respectively (Table 2). As a group they would decrease future diameter as their main component (*i.e.*, BAPH) increases.

The group composed by the predictor variables TPH, t^*TPH and t^2*TPH had coefficients which were negative, negative and positive, respectively. Their combined effects would decrease future diameter. Signs were positive, negative and positive for the coefficients related to SDI, t^*SDI and t^2*SDI , respectively. Their magnitudes were also relatively small and their influence on future diameter is negligible.

Validation of FDM

The final FDM model was validated using the estimation data set, by computing the goodness-of fit statistics described above. Mean squared and root mean square errors were 1.075 and 1.037; bias and relative bias (%) were 0.407 and 1.853; MAD was 1.705 and the pseudo coefficient of determination was 0.98.

The data set in this study had limitations, which caused problems in the modelling work. That may have affected model predictions. Since sample trees were randomly selected and are reflective of the diameter distribution, most were medium sized; small and large trees were under-represented (the quadratic mean diameter range was from 14.25 to 32.46 cm). Their development was observed for at most 7 years. In addition, Vanclay (1994) reported problems in the growth equation when measurement intervals vary greatly. This may have resulted in low precision of the diameter increment observations, estimated as the differences of

two successive DBH measurements. For the periodic annual diameter increment PADIM, this may have been reflected in the value of the coefficient of determination ($R^2 = 0.16$). The procedure applied in the calculation of annual increments, based on successive measurements, assumes that tree growth is constant during the interval between measurements.

Several statistical techniques and combinations of predictor variables were explored to develop change in the diameter models for *P. occidentalis* Sw. in La Sierra, Dominican Republic. Even though the measurements were repeated over time and mixed models are supposed to better account for the correlated nature of observations, OLS was the best model for both PADIM and FDM.

Comparing these two approaches, it was concluded that the FDM fits did not differ conspicuously from the PADIM in estimating diameter growth, and that it logically was better for estimating yield. While individual tree predictions may not be as accurate as one would like, it is impossible to measure all the factors affecting growth. The application of this model is simple, requiring a tree list from a given stand with information on DBH, CCI, basal area in trees larger than or equal to the subject trees, basal area per hectare, trees per hectare, site density index and the extent of time for the desired projection. Although goodness-of-fit statistics values were similar for both models, the scatter showed greater precision for FDM when predicting periodic annual diameter increment. Therefore, it is recommended that the FDM is used to model the increment in stem size of the species in the study region.

This study is the first attempt to model individual tree growth of *P. occidentalis* in the Dominican Republic. These models enable individual tree simulation of diameter growth based on past growth information from common stand inventories. The species has great economic, ecologic and social importance, and the two models developed can provide valuable information for decision-makers, forest managers and researchers. The models can be used to facilitate the sustainable management of *P. occidentalis*.

Acknowledgements

We want to express the collaboration of Drs. Ralph D. Nyland, Lianjun Zhang and Edwin H. White for reviewing and suggesting their valuable inputs. Funding for this project was provided by the J. William

Fulbright Foreign Scholarship Board (FSB), and by the Graduate Assistantship program at the State University of New York, College of Environmental Science and Forestry and the Collegiate Science and Technology and contributions of minor grants from the Dominican government through Secretaria de Estado de Educacion Superior, the Program on Latin America and the Caribbean at Syracuse University, Cooperativa San Jose, Universidad ISA, Plan Sierra Inc., and PROCARYN.

References

- ALDERS D., 1979. A distance independent tree model for exotic plantations in east Africa. *Forest Science* 25, 59-71.
- AVERY T.E., BURKHART H.E., 2002. *Forest measurements*. McGraw Hill, New York. 456 pp.
- BEVILACQUA E., 1999. Growth responses in individual eastern white pine (*Pinus strobus* L.) trees following partial cutting treatments. Ph. D. University of Toronto. 137 pp.
- CAO Q.V., 1999. Prediction of annual diameter growth and survival for individual trees from periodic measurements. *Forest Science* 46(1), 127-131.
- DAVIES L.S., JOHNSON K.N., 1987. *Forest management*. McGraw-Hill Company. 790 pp.
- FOX J.C., BI H., ADES P.K., 2007. Spatial dependence and individual-tree growth models. *Forest Ecology and Management* 245, 10-19.
- KNOWE S.A., AHRENS G.R., DEBELL D.S., 1997. Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modelling approaches for young red alder plantations. *Forest Ecology and Management* 98, 49-60.
- LESSARD V.C., MCROBERTS R.E., HOLDAWAY M.R., 2001. Diameter growth models using Minnesota forest inventory and analysis data. *Forest Science* 47(3), 301-310.
- LITTELL R.C., MILLIKEN G.A., STROUP W.W., WOLFINGER R.D., 1996. *SAS System for mixed models*. SAS Institute Inc Cary, NC, USA. 633 pp.
- MAILLY D., HIRBIS S., POTHIER D., 2003. Predicting basal area increment in a spatially explicit, individual tree model. *Canadian Journal of Forest Research* 33(3), Research Library, 435.
- MUNRO D.D., 1974. Forest growth models – A prognosis. In: *Growth models for tree and stand simulation* (Fries J., ed). Department of Forest Yield Research, Research Notes Nr 30, Royal College of Forestry, Stockholm. pp. 7-21.
- MURPHY P.A., SHELTON M.G., 1996. An individual-tree basal area growth model for loblolly pine stands. *Canadian Journal of Forest Research* 26, 327-331.
- OPIE J.E., 1968. Predictability of individual tree growth using various definitions of competing basal area. *Forest Science* 14(3), 314-323.
- QUICKE H.E., MELDAH R.S., KUSH J.S., 1994. Basal area growth of individual trees: a model derived from a regional longleaf pine growth study. *Forest Science* 40(3), 528-542.
- SADIQ R.A., 1982. Estimation of stand basal area growth and yield with a reverse logistic function. *Canadian Journal of Forest Research* 13, 289-297.
- SÁNCHEZ-GONZÁLEZ M., DEL RÍO M., CANELLAS I., MONTERO G., 2006. Distance independent tree diameter growth model for cork oak stands. *Forest Ecology and Management* 225, 262-270.
- SCHRODER J., SOALLEIRO R.R., ALONSO G.V., 2000. An age-independent basal area increment model for maritime pine trees in northwestern Spain. *Forest Ecology and Management* 157, 55-64.
- TANG S., WANG Y., ZHANG L., MENG C.H., 1997. A distribution-free method to predict stand diameter distribution. *Forest Science* 43, 491-500.
- UZOH F.C.C., OLIVER W.W., 2008. Individual tree diameter increment model for managed even-aged stands of ponderosa pine throughout the western United States using multilevel linear mixed effects models. *Forest Ecology and Management* 256, 438-445.
- VANCLAY J.K., 1994. *Modelling forest growth and yield: applications to Mixed Tropical Forests*. CAB International, Wallingford. 312 pp.
- WESTFALL J.A., 2006. Predicting past and future diameter growth for trees in the northeastern United States. *Canadian Journal of Forest Research* 36, 1551-1562.
- WIMBERLY M.C., BARE B.B., 1996. Distance-dependent and distance-independent models of Douglas-fir and western hemlock basal area growth following silvicultural treatment. *Forest Ecology and Management* 89(19), 1- 11.
- WYKOFF W.R., 1990. A basal area increment model for individual conifers in the northern Rocky Mountains. *Forest Science* 36(4), 1077-1104.
- ZEIDE B., 2001. Density and the growth of even-aged stands. *Forest Science* 48(4), 743-754.
- ZHANG L., GOVE J.H., 2005. Spatial assessment of model errors from four regression techniques. *Forest Science* 51(4): 334-346.
- ZHANG L., PENG C., DANG Q., 2004. Individual-tree basal area growth models for jack pine and black spruce in northern Ontario. *Forestry Chronic* 80(3), 366-374.
- ZHAO D., BORDERS B., WILSON M., 2004. Individual-tree diameter growth and mortality models for bottomland mixed-species hardwood stands in the lower Mississippi alluvial valley. *Forest Ecology and Management* 199, 307-322.