



A comparison of artificial neural networks and regression modeling techniques for predicting dominant heights of Oriental spruce in a mixed stand

Ilker ERCANLI^{1*}, Ferhat BOLAT¹ and Hakkı YAVUZ²

¹ Çankırı Karatekin University, Faculty of Forestry, 18200, Çankırı, Türkiye. ² Karadeniz Technical University, Faculty of Forestry, 61080, Trabzon, Türkiye

*Correspondence should be addressed to İlker Ercanli: ilkerercanli@karatekin.edu.tr

Abstract

Aim of study: This paper introduces comparative evaluations of artificial neural network models and regression modeling techniques based on some fitting statistics and desirable characteristics for predicting dominant height.

Area of study: The data of this study were obtained from Oriental spruce (*Picea orientalis* L.) felled trees in even-aged and mixed Oriental spruce and Scotch pine (*Pinus sylvestris* L.) stands in the northeast of Türkiye.

Materials and methods: A total of 873 height-age pairs were obtained from Oriental spruce trees in a mixed forest stand. Nonlinear mixed-effects models (NLMEs), autoregressive models (ARM), dummy variable method (DVM), and artificial neural networks (ANNs) were compared to predict dominant height growth.

Main results: The best predictive model was NLME with a single random parameter (root mean square error, RMSE: 0.68 m). The results showed that NLMEs outperformed ARM (RMSE: 1.09 m), DVM in conjunction with ARM (RMSE: 1.09 m), and ANNs (RMSE: from 1.11 to 2.40 m) in the majority of the cases. Whereas considering variations among observations by random parameter(s) significantly improved predictions of dominant height, considering correlated error terms by autoregressive correlation parameter(s) enhanced slightly the predictions. ANNs generally underperformed compared to NLMEs, ARM, and DVM with ARM.

Research highlights: All regression techniques fulfilled the desirable characteristics such as sigmoidal pattern, polymorphism, multiple asymptotes, base-age invariance, and inflection point. However, ANNs could not replicate most of these features, excluding the sigmoidal pattern. Accordingly, ANNs seem insufficient to assure biological growth assumptions regarding dominant height growth.

Additional key words: dominant height; mixed-effects; dummy variable; machine learning; growth curve; biological interpretation.

Abbreviations used: AAE (average absolute error); ADA (algebraic difference approach); AIC (Akaike's information criterion); ANNs (artificial neural networks); AR (1) (first order autoregressive structure); ARM (autoregressive models); ARMA (1, 1) (first-order autoregressive and moving average structures); Bias (average bias); Bias (%) (percentage of average bias); BIC (Bayesian information criterion); CAR (1) (first-order continuous autoregressive error structure); DVM (dummy variable method); FDD (forest district directorate); FI (fit index); GAM (generalized additive model); GADA (generalized algebraic difference approach); ISSA (iterative screening and structure analysis); NLME (nonlinear mixed-effects model); NLS (nonlinear least square); max. AE (maximum absolute error); RMSE (root mean squared error); RMSE (%) (percentage of root mean squared error); SFE (state forest enterprise); SI (site index); SRR (sum of relative ranks).

Citation: Ercanli, I; Bolat, F; Yavuz, H (2023). A comparison of artificial neural networks and regression modeling techniques for predicting dominant heights of Oriental spruce in a mixed stand. *Forest Systems*, Volume 32, Issue 1, e004. <https://doi.org/10.5424/fs/2023321-19134>

Received: 28 Dec 2021. **Accepted:** 20 Feb 2023.

Copyright © 2023 CSIC. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Funding: The authors received no specific funding for this work.

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

Introduction

The average height of dominant or co-dominant trees at a specified index age, defined as the site index (SI), is used as an indicator of forest site quality in both pure and mixed stands (Carmean, 1972). The prediction of SI is required for efficient forest management planning since it is an imperative starting point for forest-level plans and silvicultural strategies. In particular, forest managers require site quality predictions to assess the amount of wood volume to be produced for a specific species within a time period (Skovsgaard & Vanclay, 2008). In the forestry literature, many statistical techniques have been proposed to evaluate site qualities with SI predictions. Importantly, Bailey & Clutter (1974) pioneered the algebraic difference approach (ADA) in the forestry literature for formalizing the base-age invariance feature in SI models. ADA is independent of the initial selected base age and produces polymorphic SI curves with single asymptotes for a given SI class. Thereafter, Cieszewski & Bailey (2000) introduced the generalized algebraic difference approach (GADA) as a generalization of ADA. It can produce polymorphic site index curves with multiple asymptotes.

Many nonlinear growth models such as Chapman-Richards, Bertalanffy, Gompertz, Soloboda, & Korf models have been intensively used for developing SI models and constructing SI curves. These growth models have provided desirable characteristics including polymorphism, multiple horizontal asymptotes, sigmoid trend, base-age invariance, and inflection point presence in SI predictions. In spite of the attractive performance of these growth models in predicting height-age relations, some data obtained from permanent sample plots over time or stem analyses of dominant trees cause them to violate independence among observations due to highly correlated data features in spatial and time terms. Ignoring basic statistical assumptions often results in biased standard errors in parameter predictions with these growth models (Grégoire et al., 1995). Historically, Clutter (1961) and Lappi & Bailey (1988) are the earliest pioneering studies that have reported that both spatial and temporal correlation structures are possible with data from permanent sample plots or stem analysis of individual trees. For dealing with statistical issues arising from temporal and spatial correlations in height-age relations of dominant trees, autoregressive models (ARMs) such as the first autoregressive structure AR (1), the first-order continuous autoregressive error structure (CAR (1)), and the first-order autoregressive and moving average structures (ARMA (1, 1)) have often been implemented (Weiskittel et al., 2009; Wang et al., 2011). Similarly, linear and nonlinear mixed-effects models (NLM-Es) (Lappi & Bailey, 1988; Wang et al., 2007; 2008) and dummy variable models (DVM) (Martín-Benito et al., 2008; Wang et al., 2008) have been applied to minimize impacts of spatial autocorrelations resulted from variation between sample trees or plots. Many studies have shown

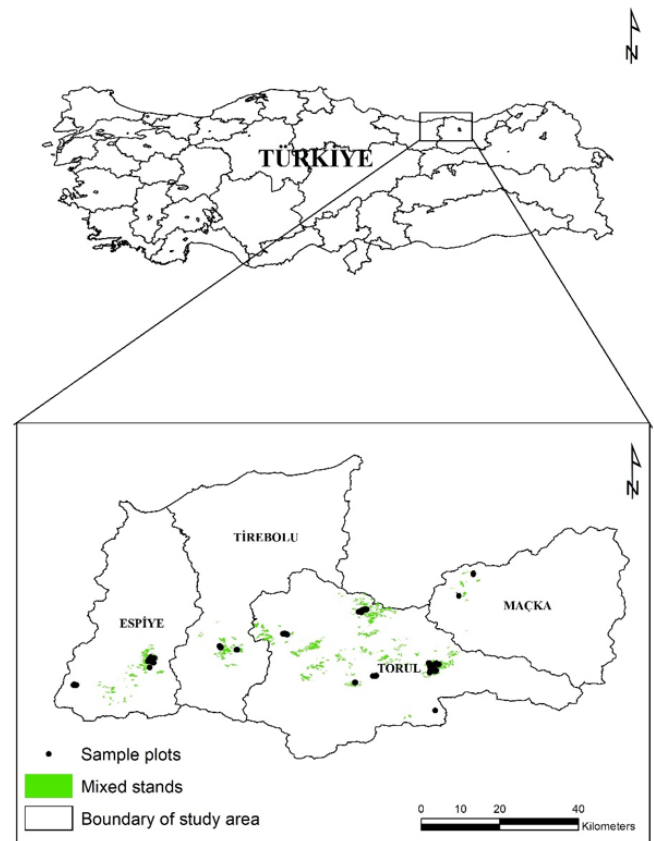


Figure 1. Location of the study areas

an improvement ranging from moderate to high upgrading fitting statistics accompanied by NLM-Es or ARMs. However, the above-mentioned statistical issues limit the use of regression models for SI predictions and require new methods for satisfactory predictions of site quality.

Artificial neural networks (ANNs), as an application and a type of artificial intelligence system, differ from many innovative prediction techniques since they need no statistical assumptions. Furthermore, ANNs can provide efficient predictions in complex and highly nonlinear data from dynamic forest systems. Therefore, forest modelers have considered ANNs as an effective tool to obtain sound and competent predictions of some individual trees as well as stand and forest attributes. Many studies have compared ANNs and regression models (e.g. NLM-Es) using fit statistics (e.g. bias and root mean square error (RMSE)). ANNs need to be tested in the context of desirable characteristics of dominant height growth, including sigmoidal growth, polymorphism, and multiple asymptotes for a thorough assessment of its prediction performance.

The objectives of this study were to: (1) evaluate the performance of ANNs to predict dominant height growth obtained from stem analysis; (2) compare ANNs and regression models (i.e., NLME, DVM, and ARM) in terms of fit statistics; and (3) evaluate the ability of ANNs in the context of desirable characteristics of dominant height growth.

Material and methods

Study site

This study was conducted in mixed Oriental spruce (*Picea orientalis* L.) and Scotch pine (*Pinus sylvestris* L.) stands in the northeast of Türkiye. These mixed stands are located at Torul State Forest Enterprise (SFE) and Maçka SFE, in the Trabzon Forest District Directorate (FDD); Tirebolu SFE and Espiye SFE in the Giresun FDD (Fig. 1). The altitudes and slopes of the studied locations vary from 650 to 1450 m and 15 to 50%, respectively. These areas are characterized geomorphologically as high mountainous land with moderate to steep slopes. The annual mean temperature is 10.9 °C and mean daily maximum temperature ranges from 11.8 °C to 28.6 °C; mean daily minimum temperature ranges from -3.8 °C to 17.4 °C, respectively. This climate is a typical Black Sea climate, characterized by a mild winter and a cool summer. The mean annual rainfall varies from 900 to 1350 mm.

Data

The sample plots were carefully investigated for traces of intensive silvicultural treatments or clear-cutting within naturally regenerated and regularly stocked stands (60–90% of tree layer cover). The stands inventoried, that were thinned moderately and had no evidence of historical damage such as fire or storms, were sampled in this study. The Oriental spruce and Scotch pine mixed stands were selected to have uniform stratification of these species when both of them had been in the upper stratum, such that there were site trees of both species in the plot. A total of 161 temporary circular plots were sampled subjectively by Ercanli (2010) to represent the existing range of stand ages, density indexes and site qualities. The sampling plot sizes ranged from 600 to 1200 m² according to stand density to ensure a minimum of 30 trees per sample plot. In each sampling stand, one dominant Oriental spruce tree

in accordance with 100 trees of the greatest height per hectare was sampled for stem analysis. The area where the study was carried out is under the management and administration of the Turkish National Forestry Directorate, but an authorization for only one felled tree was enabled in these plots. Therefore, the study was limited to only one felled tree at each sample point among the 100 trees of the greatest height per hectare.

These trees were felled, leaving stumps of a height of 0.30 m, and cross-sectional discs were cut at 3, 4, or 5 m intervals above ground surface to avoid distorting economic appraisal for tree stems in harvest applications of Turkish forestry. The total height of these felled trees was measured to the nearest 0.01 m. The number of rings was counted at each cross-sectioned point, and then these values were converted to stump age, which can be considered equal to tree age. As cross section lengths do not coincide with periodic height growth, it was necessary to adjust height and age data from stem analysis. Fabbio et al. (1994) presented the ISSA method that uses second differences in ring counts to have smoother height/age curves. Height at each age of the felled trees was calculated by the ISSA method based on height and age data at cross-sectional cuts. The data set in this study consisted of a total of 161 stem analyses with 873 height/age pairs. The mean, standard deviations, minimum, and maximum values for the data used are presented in Table 1. Also, profile plots of stem analyses for the felled trees are presented in Fig. 2.

Methods

Base nonlinear model

In the literature on forests, a variety of statistical growth functions have been employed to describe the dominant height-age relationship and to predict site index of forest stands. Besides the fitting abilities of these statistical models, the model used to obtain site index pre-

Table 1. Descriptive statistics of tree and stand data for these dominant trees.

Variables ^[1]	Minimum value	Maximum value	Mean value	Std. dev.
h (m)	0.3	28.8	8.9	6.3
t (years)	2.0	175.0	50.9	35.6
dbh (cm)	19.5	63.5	35.4	9.45
Stand basal area (m ² ha ⁻¹)	7.09	101.2	45.5	17.1
Number of trees (ha ⁻¹)	270	2275	1056.8	367.2
Stand relative density (Curtis et al., 1981)	1.82	18.12	9.32	3.20
Stand age (years)	23.3	115.4	75.4	16.2

^[1] h: total tree height, t: total tree age, dbh: diameter at breast height.

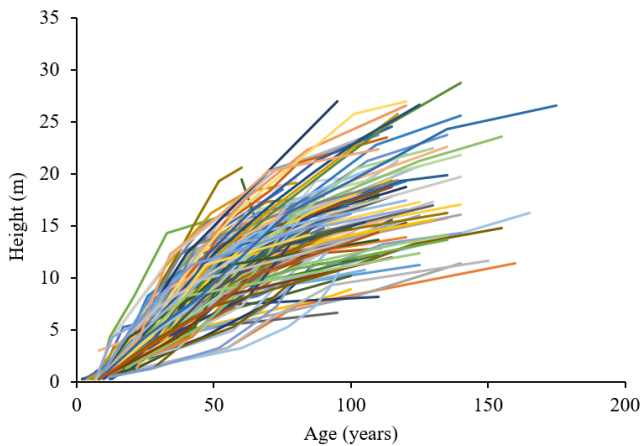


Figure 2. Line plots of 873 dominant height data obtained from 161 stem analysis trees.

dictions has some desirable growth properties, including: polymorphism, base-age invariance, sigmoid growth pattern with an inflection point, passing a zero point, multiple asymptotes with function of site index (increasing with increasing site index), parallel asymptote at older ages and a non-decreasing curve over time (Corral-Rivas et al., 2004; Diéguez-Aranda et al., 2006; Martín-Benito et al., 2008). In the present study, the GADA expansion of Chapman-Richards growth model (Chapman, 1961; Richards, 1959) was selected as the base function due to its better predictive results compare to other functions in our preliminary analyses, in which some preliminary analyses were performed based on both fitting criteria and the desirable growth properties to determine the predictive activation function of choice from these statistical functions. The Chapman-Richards growth model $h = a[1 - \exp(-bt)]^c$ was solved by a and assuming $b = c \left(\frac{h}{t}\right)^d t^f$ (Corral-Rivas et al., 2004), and so the parameter was included as an intermediate variable in this function structure. Thus, the GADA dynamic age-height models based on model Chapman-Richards base growth model have the form given in Eqs. (1) and (2):

$$h = h_0 \left[\frac{(1 - \exp(-X_0 t))}{(1 - \exp(-X_0 t_0))} \right]^c \quad (1)$$

$$X_0 = d \left(\frac{h_0}{t_0}\right)^f t_0^g \quad (2)$$

where h is the height at an age t ; and h_0 is the height (site index, e.g., 10, 12, 14, 16, 18, 20 or 30 m site index values) at age t_0 (base age, e.g., 100 years) for the sample trees.

Regression analysis techniques

The nonlinear least square (NLS) technique, based on the Marquardt algorithm, was performed using the PROC MODEL procedure in SAS/STAT® 9 software (SAS Inst.,

2004). When fitting the Chapman-Richards growth model by NLS, some local and global parameters for each tree variable (h_0) were replaced with $[\sum_{i=1}^n h_i k_i]$, which is a sum of terms containing a tree-specific parameter (h_0). The Chapman-Richards growth model was fitted by expanding each parameter, adding a related parameter and a dummy variable to differentiate the site represented by these felled trees, and it corresponds to various sets of parameters for each sample tree. Thus, a dummy variable for each tree (k_i), that was equal to 1 for the i^{th} tree and 0 otherwise, was used as an application of the DVM, described by Cieszewski et al. (2000). It is necessary to replace the h_0 variable with a sum of terms that includes a tree specific parameter (associated with the site by including its simultaneous site value) and a dummy variable for each sample tree such as $h_{01}k_1 + h_{02}k_2 + \dots + h_{0n}k_n$ in Eq. (1). In this Eq. (1), h_{0i} is the site-specific parameter for each sample trees i , and k_i is a dummy variable equal to 1 for individual i and 0 otherwise.

ARMA (1, 1) was performed to account for possible temporal autocorrelations in residuals that are important issues in dominant height predictions. Also, ARMA (1, 1) incorporating dummy variables was used to simultaneously consider temporal autocorrelations and spatial variations. This analysis was programmed using the SAS/ETS PROC MODEL procedure.

Some studies [e.g., Wang et al. (2007; 2008)] have evaluated NLMEs as an alternative technique for predicting tree-specific parameters. In the forestry literature, NLMEs have been used as a statistical method to deal with autocorrelation problems that arise in hierarchical data. In NLMEs, population-specific fixed parameters and sampling unit-specific random parameters are predicted concurrently by incorporating a covariance matrix into the model structure (Calama & Montero, 2004). NLMEs provide a significant quantification for spatial inter-individual variations that can be generated by essentially environmental factors in the context of SI predictions. In DVM, a more common method of modeling SI, global and local parameters are predicted for quantifying local growth variations with general height trends. However, in NLME, the inclusion of random parameters into the model structure enables estimation of local height variations among clustered or nested sample units located in different stands. Wang et al. (2008) compared the performance of NLME with DVM to predict dominant heights and proposed that these methods were proper for constructing models with specific or local parameters. A general expression of the model structure of NLMEs is as follows:

$$H_{ij} = f(\theta_{ij} + v_{ij}) + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2) \quad (3)$$

$$\theta_{ij} = A_{ij} \beta + B_{ij} b_i, \quad b_i \sim N(0, \sigma^2) \quad (4)$$

where H_{ij} represents a dependent variable for j^{th} tree at i^{th} time; θ is a vector of tree-specific parameter; v is a co-

variate vector; ε is an error term that randomly distributes around mean zero; N and σ^2 denote normal distribution and variance, respectively; β shows effect of fixed parameter(s) that is/are computed for whole population; and b_i represents effect of random parameter(s) for i th tree.

We estimated the variance components and fixed parameters of the base nonlinear model using the PROC NLMIXED procedures of the SAS/STAT 9 package (SAS Inst., 2004). The maximum-likelihood method was used to fit NLMEs, and the adaptive Gaussian quadrature was used in the computation of integral over random effects as described by Pinheiro & Bates (2006). In NLMEs, some parameters of the model must be set to random or fixed effect parameters in a given model structure. Eq. (1) has four parameters: c , d , f , and g . It is important to determine which of these parameters are fixed or random ones, which alternatives for fixed and random parameters were compared based on distributions of models' residuals and some goodness-of-fit statistics. Spatial inter-individual variations among individual trees and autocorrelation problems caused by hierarchical data structures can be solved by NLME. However, since the data set used in this study originated from stem analysis, the autocorrelation problem among observations obtained from the same tree were occurred, and this feature results in significant correlation between residuals and invalidates standard hypothesis testing (Grégoire et al., 1995). Therefore, Wang et al. (2007; 2013) and Sharma et al. (2014) used NLMEs and ARMs such as AR (1), ARMA (1, 1), CAR (1), TOEPLITZ and spatial power structure (POW (1)) to model dominant height-age relations. We used different combinations of NLMEs and ARMs for further modeling of residual autocorrelation that appeared in stem analysis data. We excluded POW (1) and CAR (1) from evaluations due to non-convergence fitting results. We performed these analyses using the SAS NLINMIX macro, which uses the expansion around zero method. Also, the SAS NLINMIX macro was used in this study to combine temporal and spatial correlations by using these ARM and NLME procedures.

Artificial neural network techniques

We used multiple layer feed forward (FF) and cascade correlation (CC) learning algorithms with the Levenberg-Marquardt algorithm as an alternative prediction technique to the above-mentioned regression techniques to predict dominant height values (h) at any tree ages (t) from input variables of tree ages (t), dominant height (h_0) at age of t and base age (t_0). These networks consist of an input layer, a hidden layer and an output layer and activation functions with hyperbolic tangent sigmoid (tansig), logistic sigmoid (logsig) and linear (purelin). These activation functions link together network layers that critically affect the fitting performance of neural networks. We compared the following alternative combinations of the activation

functions: (A1) tansig function between input layer and hidden layer and tansig function between hidden layer and output layer; (A2) tansig function between input layer and hidden layer and logsig function between hidden layer and output layer; (A3) tansig function between input layer and hidden layer and Purelin function between hidden layer and output layer; (A4) logsig function between input layer and hidden layer and logsig function between hidden layer and output layer; (A5) logsig function between input layer and hidden layer and tansig function between hidden layer and output layer; (A6) logsig function between input layer and hidden layer and purelin function between hidden layer and output layer; (A7) purelin function between input layer and hidden layer and purelin function between hidden layer and output layer; (A8) purelin function between input layer and hidden layer and logsig function between hidden layer and output layer; and (A9) purelin function between input layer and hidden layer and tansig function between hidden layer and output layer to decide the best predictive combination. The combination of A2, A4 and A8 resulted in the non-convergence of ANN models, thus these three alternatives were excluded. Another important factor affecting the performance of the network structure is the number of neurons in hidden layers. In our study, the number of neurons ranged from 1 to 100 (1, 2, 3,20, 50, 70, 90 and 100). Thus, a total of 600 trained networks, including 100 number of neurons and 6 transfer function alternatives based on multiple layer feed forward (FF) and cascade correlation (CC) learning algorithms with 600 alternatives, were trained and used to obtain the most accurate dominant height predictions. All these alternatives to artificial neural network models were trained using the "Neural Network Toolbox" (NNtool module) of MATLAB (The MathWorks Inc., 2015a).

Evaluations for various prediction techniques

The performance of the models used to predict the dominant height growth of study forests was compared using statistical criteria and graphical evaluation. The statistical criteria were coefficient of correlation between observed and predicted heights (r), average absolute error (AAE), maximum absolute error (max. AE), root mean squared error (RMSE), percentage of root mean squared error (RMSE%), average bias (Bias), percentage of average bias (Bias%), fit index (FI), Akaike's information criterion (AIC), and Bayesian information criterion (BIC). These criteria are defined as follows:

$$r = \left(\frac{\sum_{i=1}^n [(h_i - \bar{h}_i) \times (\hat{h}_i - \hat{\bar{h}}_i)]}{\sqrt{\sum_{i=1}^n (h_i - \bar{h}_i)^2 \times \sum_{i=1}^n (\hat{h}_i - \hat{\bar{h}}_i)^2}} \right) \quad (5)$$

$$AAE = \sum_{i=1}^n |h_i - \hat{h}_i| / n \quad (6)$$

$$\max. AE = \max(|h_1 - \hat{h}_1| \dots \dots \dots |h_n - \hat{h}_n|) \quad (7)$$

$$RMSE = \sqrt{\sum_{i=1}^n (h_i - \hat{h}_i)^2 / (n - k)} \quad (8)$$

$$RMSE\% = \left(\left[\sqrt{\sum_{i=1}^n (h_i - \hat{h}_i)^2 / (n - k)} \right] / \bar{h}_i \right) \times 100 \quad (9)$$

$$Bias = \sum_{i=1}^n (h_i - \hat{h}_i) / n \quad (10)$$

$$Bias\% = \left(\left[\sum_{i=1}^n (h_i - \hat{h}_i) / n \right] / \bar{h}_i \right) \times 100 \quad (11)$$

$$FI = 1 - \frac{\sum_{i=1}^n (h_i - \hat{h}_i)}{\sum_{i=1}^n (h_i - \bar{h}_i)} \quad (12)$$

$$AIC = n \times \ln(RMSE) + 2 \times k \quad (13)$$

$$BIC = n \times \ln(RMSE) + n \times \ln(k) \quad (14)$$

where \hat{h}_i is the predicted height of i th tree; h_i is the observed height of i th tree; and k is the number of independent variables. Smaller values for AAE, max. AAE, RMSE, RMSE%, Bias, Bias%, AIC, and BIC, as well as higher values for r and FI, indicate better fit results. We calculated the sum of relative ranks (SRR) (Poudel & Cao, 2013) for these prediction techniques. We calculated relative ranks separately for the criteria and added up those relative ranks to calculate SRR. Therefore, the SRR is a combination of different model fit criteria. The prediction method with the lowest sum of rank values was considered the best predictive method.

In addition, assessments of desirable characteristics of dominant height growth have been an important requirement in modeling dominant height. As stated in many studies such as Bailey & Clutter (1974) and Cieszewski (2002), desirable characteristics of height growth curves are: (1) a sigmoidal structure (known as an ‘‘S curve’’); (2) a logical behavior with a zero value of height at age zero; (3) polymorphism; (4) an asymptotic behavior; and (5) base-age invariance. In this study, the model compatibility with these desirable features was assessed graphically using site index curves (5, 10, 15, 20, 25, and 30 m at base age 100).

Results

The predictive ability with AAE, max. AAE, RMSE, RMSE%, AIC, BIC, Bias, and Bias%, r and FI of regression techniques based on NLME and ARM for temporal and spatial autocorrelations had significant variability depending on whether these models including the single and double random parameters. The NLME model, which randomly includes the single and double parameters with f and g , resulted in better predictive ability with lower AAE, max. AAE, RMSE, RMSE%, AIC, BIC, Bias, and Bias%, and the higher r and FI compared to the prediction fit statistics of the base nonlinear model with NLS (Table 2). However, the NLME model with two random alternatives (d & g and d & f) and also three random alternatives (f , d

& g), gave the poorest fitting results in dominant height predictions. RMSE was significantly reduced (approx. 17% and 48%, respectively) when the effects of temporal and spatial autocorrelations with NLME and ARM were included in the base nonlinear model. However, combining simultaneous temporal and spatial autocorrelations decreased the ability of the base nonlinear model to fit dominant height growth data in this study. In this context, simultaneously combining temporal and spatial autocorrelations with NLINMIX macro and model structure with two or three random parameters in NLMEs resulted in an increase in RMSE, ranging from 0.03 to 4.20 m. Based on this information, NLMEs with d - f , d - g , and d - f - g random parameters were inadequate to fit dominant height growth data in this study. Based on the relative rank of Poudel & Cao (2013), NLME- f resulted in the best predictive ability with at least %99 of the total variation in tree heights (Table 3). Table 4 shows parameter estimates including parameters of fixed and random effects for NLME- f .

From the results for these ANN models, a network architecture with a tan- or log-sigmoid transfer function for the input layer improved model accuracy for both FF and CC network architectures (Table 2). In contrast, selecting the purelin transfer function for the input layer considerably reduces model accuracy and efficiency. For instance, the RMSE value for CC-tansig-purelin (cascade correlation learning algorithms with tan-sig function between input layer and hidden layer and Pure-Lin function between hidden layer and output layer, abbreviated as tan-pure) with 46 neurons was 1.14, while it was 2.40 for CC-purelin-tansig with 21 neurons. Also, FI for CC-tansig-purelin with 46 neurons was 0.97, while it was 0.86 for CC-purelin-tansig with 21 neurons. In summary, the selection of this transfer function, purelin, for the input layer for dominant height growth was inappropriate, and the selection of the tansig or logsig transfer function for the input layer was more suitable for our data set. Note that our network architecture with a logsig transfer function for the output layer for both ANN architectures did not converge.

In general, ANNs resulted in moderately worse predictive ability with higher AAE, max. AAE, RMSE, RMSE%, AIC, BIC, Bias, and Bias%, and the lowest r and FI compared to given regression techniques. While the RMSE differences between ARMA (1,1) and ANNs with tan-or-log-sigmoid transfer functions for the input layer were negligible (0.2 m), the differences between the ANNs and NLME- f were significant (> 0.5 m), implying a 40% reduction in RMSE (Table 2). These results suggested that spatial autocorrelations were more influential than temporal autocorrelations in the dominant height growth trends of Oriental spruce located in these studied stands.

Figs. 3, 4, and 5 depict graphical analyses of RMSE% and lagged-residuals, as well as observed vs. predicted values. Fig. 3 shows that NLME- f had a smaller RMSE% across all ages than other models. CC-tansig-tansig with

Table 2. Fit statistics of the regression techniques, feed-forward and cascade-forward network architectures to predict dominant height growth of Oriental spruce stands growing in Oriental spruce and Scot pine forest associations.

Prediction techniques ^[1]	RMSE	RMSE%	AIC	BIC	BIAS	BIAS%	FI	AVEBE	MAXAVEBE	r
Base nonlinear model	1.32	14.95	251.01	265.32	-0.01	-0.14	0.96	0.90	6.29	0.98
NLME-c	0.88	9.91	-108.08	-93.77	-0.22	-2.43	0.98	0.57	5.47	0.99
NLME-c-d	1.66	18.74	447.59	461.90	-0.34	-3.82	0.93	0.62	23.30	0.97
NLME-c-f	1.94	21.89	583.18	597.49	-0.59	-6.68	0.91	0.78	26.54	0.96
NLME-c-g	1.79	20.19	512.72	527.03	-0.58	-6.50	0.92	0.76	22.93	0.96
NLME-d	0.67	7.61	-337.67	-323.36	0.06	0.67	0.99	0.45	4.00	0.99
NLME-d-f	3.37	38.03	1064.81	1079.12	2.30	25.97	0.72	2.34	11.72	0.93
NLME-d-g	5.53	62.43	1497.08	1511.39	3.40	38.36	0.24	3.46	25.80	0.84
NLME-f	0.68	7.66	-332.25	-317.94	-0.06	-0.64	0.99	0.46	3.67	0.99
NLME-f-g	0.76	8.59	-232.84	-218.53	-0.13	-1.49	0.99	0.52	3.85	0.99
NLME-g	0.68	7.69	-328.49	-314.18	-0.06	-0.67	0.99	0.46	3.64	0.99
NLME-f-d-g	4.20	47.48	1258.33	1272.64	1.99	22.49	0.56	2.20	26.60	0.90
ARMA (1, 1)	1.09	12.26	77.74	92.06	-0.02	-0.23	0.97	0.78	5.47	0.99
ARMA and dummy variable	1.09	12.29	79.66	93.97	-0.01	-0.07	0.97	0.79	5.45	0.99
NLME-f-AR1	1.35	15.29	270.18	284.49	0.00	0.03	0.95	0.89	6.18	0.98
NLME-f-ARMA (1, 1)	1.35	15.29	270.18	284.49	0.00	0.03	0.95	0.89	6.18	0.98
NLME-f-TOEPH	1.35	15.29	270.25	284.57	0.00	0.03	0.95	0.89	6.18	0.98
FF-log-tan (52)	1.11	12.56	99.04	113.35	0.00	0.05	0.97	0.76	5.80	0.98
FF-log-pure (75)	1.14	12.85	118.89	133.20	0.08	0.94	0.97	0.79	7.06	0.98
FF-pure-pure (27)	2.24	25.26	708.13	722.44	-0.03	-0.36	0.88	1.75	8.53	0.93
FF-pure-tan (28)	2.40	27.10	769.32	783.63	-0.17	-1.92	0.86	1.96	8.09	0.93
FF-tan-tan (100)	1.13	12.77	113.16	127.47	0.02	0.28	0.97	0.80	6.25	0.98
FF-tan-pure (87)	1.14	12.85	118.67	132.98	0.03	0.33	0.97	0.78	7.50	0.98
CF-log-tan (24)	1.13	12.77	113.55	127.87	-0.06	-0.63	0.97	0.79	5.91	0.98
CF-log-pure (28)	1.27	14.36	215.82	230.14	0.15	1.68	0.96	0.94	5.56	0.98
CF-pure-pure (90)	2.24	25.25	707.73	722.05	-0.03	-0.39	0.88	1.75	8.64	0.93
CF-pure-tan (21)	2.40	27.10	769.28	783.59	-0.19	-2.16	0.86	1.96	8.07	0.93
CF-tan-tan (53)	1.12	12.70	108.49	122.81	-0.08	-0.85	0.97	0.78	5.35	0.98
CF-tan-pure (46)	1.14	12.84	118.01	132.32	-0.02	-0.28	0.97	0.80	5.05	0.98

^[1]Lowercase letters denoted as c, d, f, and g show fixed and random parameters for nonlinear mixed-effect models (NLMEs) for accounting to spatial autocorrelation. NLMEs consisted of one, two or three random parameters (e.g. it is called as NLME-c for one random parameter). AR1 and ARMA (1, 1) are the first order autoregressive and the first order autoregressive moving average structure, respectively for taking temporal autocorrelation into consideration. NLME-f-AR1, -ARMA (1, 1) or -TOEPH shows NLINMIX macro for combining simultaneously temporal and spatial autocorrelations. FF and CF sign feed-forward and cascade-forward network architectures. If data is transferred with log-sigmoid function into input layer, it is called "FF-log" for feed-forward architecture. Then, if data weighted in hidden layer is transferred with tan-sigmoid function into output layer, it is called "FF-log-tan". Values into parenthesis for FF and CF present the number of neurons in given ANN architectures.

53 neurons and the ARMA (1,1) displayed a very similar pattern in terms of RMSE% and had poorer estimates across ages over 100 years. For lower ages than 100 years, the base model and NLME-f-AR1 yielded worse predictions. While NLME-f, CC-tansig-tansig with 53 neurons and ARMA (1, 1) behaved consistently above 50 years, the base model and NLME-f-AR1 showed a steady decrease in terms of RMSE%. All models produced the

largest RMSE% for young trees below 50 years. While RMSE% for NLME-f, CC-tansig-tansig with 53 neurons, and ARMA (1, 1) was less than 10% for ages older than 50 years, it was less than 10% and RMSE% for the base model and NLME-f-AR1 over 90 years, indicating that NLME-f, and ARMA (1, 1) models for over 50 years, and NLME-f-AR1 over 90 years can be used reliably in prediction of dominant height growth in the studied stands.

Table 3. Relative rank orders of the regression techniques, feed-forward and cascade-forward network architectures in terms of fit statistics

Prediction techniques ^[1]	RMSE	RMSE%	AIC	BIC	BIAS	BIAS%	FI	AVEBE	MAXAVEBE	r	Rank ^[2]
Base nonlinear model	5.15	5.15	10.95	10.95	5.50	5.50	2.34	5.63	4.58	4.39	60.14
NLME-c	2.30	2.30	4.88	4.88	3.93	3.93	1.32	2.29	3.47	1.69	30.98
NLME-c-d	7.29	7.29	14.27	14.27	2.97	2.97	3.37	2.78	27.54	6.64	89.39
NLME-c-f	9.07	9.07	16.56	16.56	1.00	1.00	4.40	4.45	31.92	8.58	102.61
NLME-c-g	8.11	8.11	15.37	15.37	1.12	1.12	3.82	4.21	27.04	7.22	91.50
NLME-d	1.00	1.00	1.00	1.00	6.06	6.06	1.00	1.00	1.48	1.00	20.61
NLME-d-f	18.20	18.20	24.70	24.70	23.47	23.47	12.21	20.48	11.91	15.36	192.68
NLME-d-g	32.00	32.00	32.00	32.00	32.00	32.00	32.00	32.00	30.92	32.00	318.92
NLME-f	1.03	1.03	1.09	1.09	5.16	5.16	1.01	1.13	1.05	1.02	18.75
NLME-f-g	1.55	1.55	2.77	2.77	4.58	4.58	1.13	1.72	1.28	1.24	23.16
NLME-g	1.05	1.05	1.16	1.16	5.14	5.14	1.01	1.14	1.00	1.02	18.85
NLME-f-d-g	23.54	23.54	27.97	27.97	21.07	21.07	18.73	19.04	32.00	19.90	234.84
ARMA (1, 1)	3.63	3.63	8.02	8.02	5.44	5.44	1.75	4.46	3.48	2.88	46.74
ARMA and dummy variable	3.64	3.64	8.05	8.05	5.55	5.55	1.75	4.52	3.44	2.89	47.09
NLME-f-AR1	5.34	5.34	11.27	11.27	5.62	5.62	2.42	5.53	4.42	4.58	61.41
NLME-f-ARMA (1, 1)	5.34	5.34	11.27	11.27	5.62	5.62	2.42	5.53	4.42	4.58	61.41
NLME-f-TOEPH	5.34	5.34	11.27	11.27	5.62	5.62	2.42	5.53	4.43	4.58	61.42
FF-log-tan (52)	3.80	3.80	8.38	8.38	5.63	5.63	1.81	4.27	3.92	3.04	48.66
FF-log-pure (75)	3.96	3.96	8.71	8.71	6.24	6.24	1.87	4.54	5.62	3.26	53.12
FF-pure-pure (27)	10.98	10.98	18.67	18.67	5.35	5.35	5.68	14.44	7.60	14.03	111.76
FF-pure-tan (28)	12.02	12.02	19.70	19.70	4.27	4.27	6.46	16.54	7.01	15.05	117.05
FF-tan-tan (100)	3.92	3.92	8.62	8.62	5.79	5.79	1.85	4.69	4.53	3.14	50.86
FF-tan-pure (87)	3.96	3.96	8.71	8.71	5.83	5.83	1.87	4.45	6.21	3.38	52.90
CF-log-tan (24)	3.92	3.92	8.62	8.62	5.16	5.16	1.85	4.55	4.07	3.15	49.03
CF-log-pure (28)	4.82	4.82	10.35	10.35	6.76	6.76	2.20	6.06	3.59	4.61	60.31
CF-pure-pure (90)	10.97	10.97	18.66	18.66	5.33	5.33	5.68	14.40	7.75	14.22	111.99
CF-pure-tan (21)	12.02	12.02	19.70	19.70	4.11	4.11	6.46	16.54	6.99	15.04	116.70
CF-tan-tan (53)	3.88	3.88	8.54	8.54	5.02	5.02	1.83	4.44	3.31	3.10	47.54
CF-tan-pure (46)	3.96	3.96	8.70	8.70	5.41	5.41	1.86	4.61	2.90	3.19	48.68

^[1] Lowercase letters denoted as c, d, f, and g show fixed and random parameters for nonlinear mixed-effect models (NLMEs) for accounting to spatial autocorrelation. NLMEs consisted of one, two or three random parameters (e.g. it is called as NLME-c for one random parameter). AR1 and ARMA (1, 1) are the first order autoregressive and the first order autoregressive moving average structure, respectively for taking temporal autocorrelation into consideration. NLME-f-AR1, -ARMA (1, 1) or -TOEPH shows NLINMIX macro for combining simultaneously temporal and spatial autocorrelations. FF and CF sign feed-forward and cascade-forward network architectures. If data is transferred with log-sigmoid function into input layer, it is called "FF-log" for feed-forward architecture. Then, if data weighted in hidden layer is transferred with tan-sigmoid function into output layer, it is called "FF-log-tan". Values into parenthesis for FF and CF present the number of neurons in given ANN architectures. R: coefficient of correlation between observed and predicted heights, AAE: average absolute error, max. AE: maximum absolute error, RMSE: root mean squared error, RMSE%: percentage of root mean squared error, Bias: average bias, Bias%: percentage of average bias, FI: fit index, AIC: Akaike's information criterion, and BIC: Bayesian information criterion. ^[2] Sum of relative order values calculated for the regression techniques and ANNs architectures was denoted as Rank.

The one-lagged residuals of all models, except for CC-tansig-tansig with 53 neurons, suggested no temporal autocorrelations, and were identically distributed around

zero (Fig. 4). CC-tansig-tansig with 53 neurons showed a slight temporal autocorrelation, but it was not clear evidence of temporal autocorrelation. Correcting for tem-

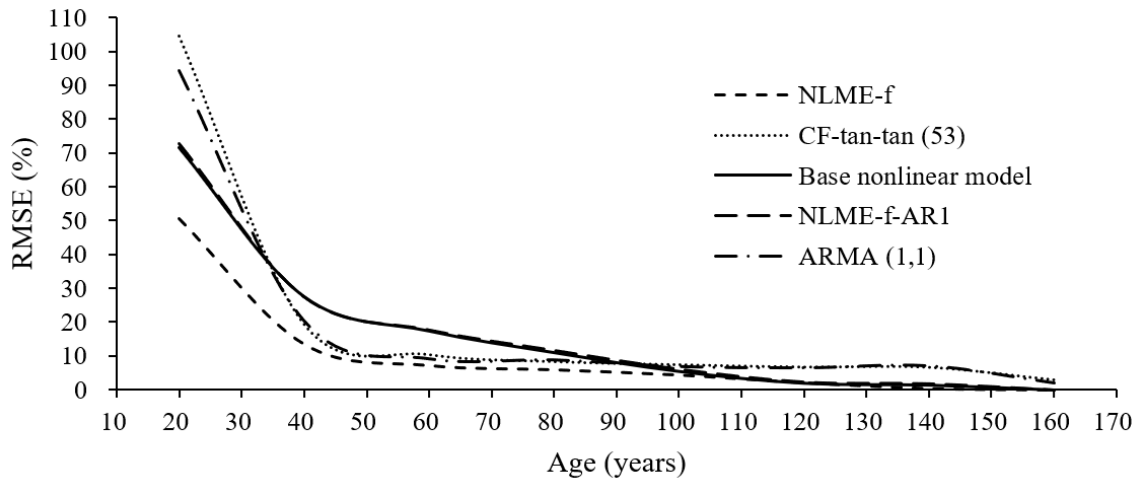


Figure 3. Line plots of regression techniques, the base nonlinear model, ARMA (1, 1), NLME-f, NLME-f-AR1 and cascade-forward network architecture with tan-sigmoid transfer function for both input and output layers (CF-tan-tan-53) in terms of RMSE

poral autocorrelations in residuals of the base model did not improve prediction accuracy. While the RMSE for the base model was 1.32 m, it was 1.09 m for ARMA (1, 1). Likewise, Bias was -0.01 and -0.02 for the base model and ARMA (1, 1), respectively. Graphics of observed vs. predicted dominant heights showed NLME-f to be better than other models (Fig. 5). Patterns of the base model, NLME-f-AR1, and CF-tan-tan with 53 neurons were highly similar, but were slightly different from ARMA (1, 1).

Some graphs of dominant height growth for five site index values (i.e. 5, 10, 15, 20, 25, and 30 m at base age 100) for given regression techniques, feed-forward and cascade-forward network architectures are shown in Fig. 6, Fig. 7 and Fig. 8, respectively. All regression techniques provided expected behaviors of dominant height growth (e.g. polymorphism, multiple asymptotes, and sigmoid pattern). It is interesting that while FF-purelin-tansig with 28 neurons and CC-purelin-tansig with 21 neurons performed worse than other ANN architectures in terms of given fit statistics, these network architectures provided most of the desirable properties of dominant height growth. In other

words, ANN architectures that were fitted well, failed to assure the desirable properties of dominant height growth in this study. In detail, dominant height growth for the regression techniques, FF-purelin-tansig with 28 neurons and CC-purelin-tansig with 21 neurons showed a sigmoidal pattern (known as the “S curve”). Site index curves for these models are polymorphic for varying site indexes, and follow multiple asymptotes at older ages (i.e. beyond 150 years). While the regression techniques had strong asymptotic behaviors for five site index curves, that is a desirable property for dominant height growth and provide the biological interpretation, FF-purelin-tansig with 28 neurons and CC-purelin-tansig with 21 neurons, had fairly weak asymptotic behaviors. For the predictions from these ANN models, it is obvious that h is not 0 at age 0 and extrapolating beyond the measurements range indicate the SI curves go to a common asymptote. In the same way, whereas dominant heights for the regression techniques were equal to zero at age zero, those for all ANNs were not at age zero. Also, whereas dominant heights for the regression techniques were equal to site index values at base age

Table 4. Parameter estimates, standard errors (S.E.), t-values, and p-values for NLME-f

Parameter ^[1]	Estimate	S.E.	t-value	p-value
c	2.3696	0.0574	41.31	<0.0001
d	0.5614	0.1871	3.00	0.0031
f	-0.2934	0.0915	3.21	0.0016
g	-0.7551	0.0913	8.28	<0.0001
σ_f^2	0.0232	0.0031	7.44	<0.0001
σ_e^2	0.5189	0.0276	18.81	<0.0001

^[1] σ_f^2 and σ_e^2 , variance components. ϵ , error terms.

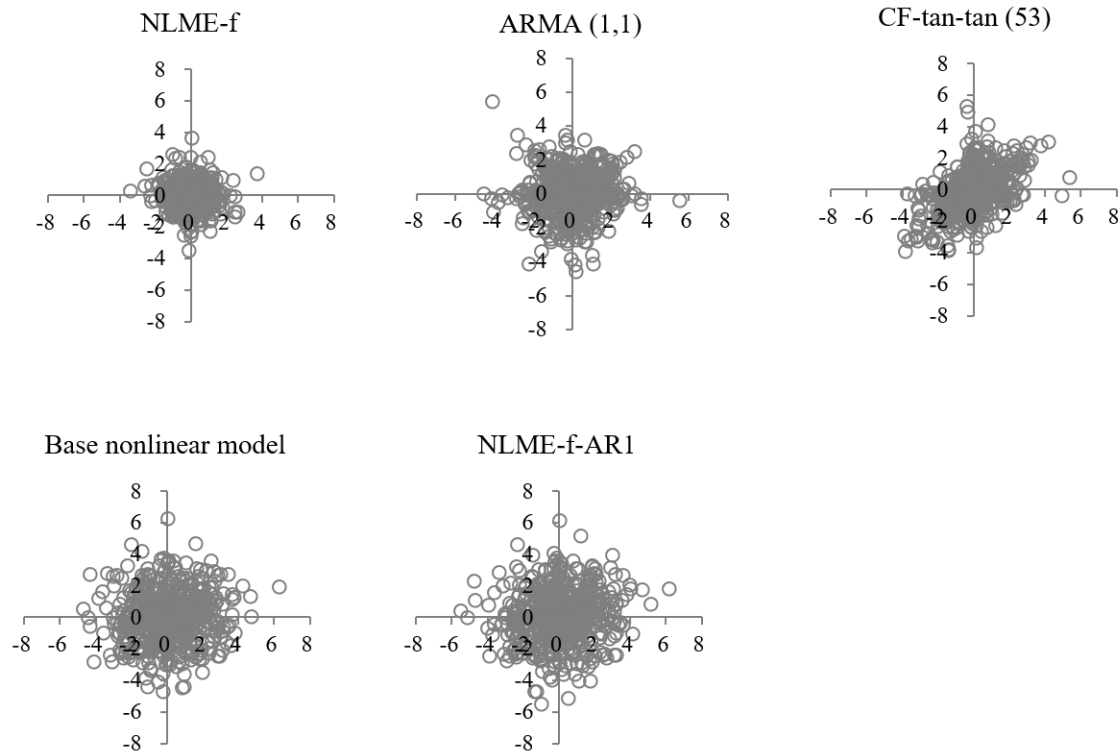


Figure 4. The one lagged residuals of regression techniques, the base nonlinear model, ARMA (1, 1), NLME-f, NLME-f-AR1 and cascade-forward network architecture with tan-sigmoid transfer function for both input and output layers (CF-tan-tan-53).

100, those for all ANNs architecture were different from site index values at a given base age. All these findings on site index curves show that Chapman-Richards site index model, which is an n-degree polynomial regression model to a certain extent, is more effective than the different ANN model structures in obtaining site index predictions.

Discussion

In this study, dominant height data were obtained from stem analysis and these data, which were obtained from the same tree, were serially correlated, and this feature may cause a larger error variance as reported by many studies (Lappi & Bailey, 1988; Wang *et al.*, 2007; Martín-Benito *et al.*, 2008; Wang *et al.*, 2008). We implemented ARMA (1, 1) to remove the effects of the covariance structure on predictions and achieved a significant improvement in predictions (Table 2). Also, mixed-effects models have been recommended to account for variation among observations driven by unobserved influences. This unobservable variability is taken into account by adding subject-specific parameters at the tree or plot level (even region level) to a given linear or nonlinear model (Wang *et al.*, 2008). In context of site index, Calegario *et al.* (2005) stated that subject-specific parameters are needed to achieve desirable site curves that vary with a site quality. In this study, the nonlinear models with a single random parameter gener-

ally outperformed those with an autoregressive structure (Table 2). NLME with two and three random parameters performed generally worse than those with one random parameter. Conversely, Wang *et al.* (2011) reported that NLME with two or more random parameters performed worse relative to those with a single random parameter, which probably resulted from predicting two or more random parameters using a single height measurement. Wang *et al.* (2008) also reported that in the case of using multiple measurements, NLME with two random parameters outperformed the dummy variable method since NLME considers both the subject-specific parameter and the error parameter.

Fortin *et al.* (2007) demonstrated that combining correlated error terms that appeared in repeated measurements with random effects provided the smallest fit statistics in predicting merchantable basal area. However, they also stated that considering correlated error terms only may provide more accurate estimates than combining correlations in residuals and random effects in the case of less variability among subjects (e.g., trees). Likewise, Martín-Benito *et al.* (2008) used dummy variables as subject-specific responses and could not achieve higher accuracy because of less variability resulting from compensating effects among regions. Conversely, in some experiments, as evidenced by our study, the correlated errors seemed to be negligible, whereas the variability among subject trees was significant (Wang *et al.*, 2011). One possible reason is that correla-

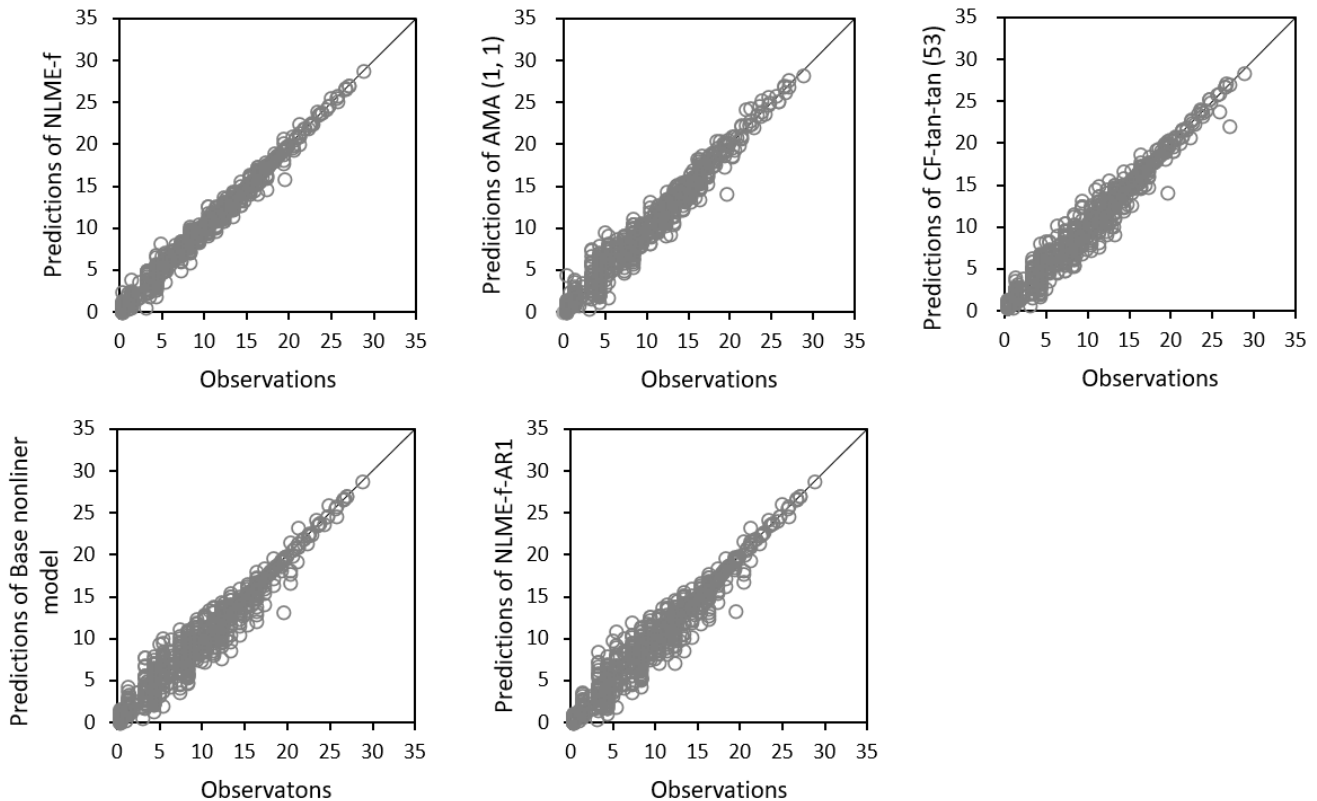


Figure 5. Scatter plots relating to observations vs. predictions using regression techniques, the base nonlinear model, ARMA (1, 1), NLME-f, NLME-f-AR1 and cascade-forward network architecture with tan-sigmoid transfer function for both input and output layers (CF-tan-tan-53).

tions between closer measurements in time were insignificant (<1%) (Nothdurft et al., 2006). Additionally, since data sets were collected from different ecological regions in this study, adding tree-specific parameters as random effects to the base nonlinear model enhanced the prediction accuracy significantly (Gregorie, 1987; Wang et al., 2008).

Most of the dominant height models developed resulted in erratic estimates at especially younger ages below 50 years. This could be due to changes in the status of dominant trees over time, which could be influenced by factors such as a lack of knowledge about the selection of dominant trees at these ages, as well as temperature and rainfall (García, 2005; Martn-Benito et al., 2008; Weiskittel et al., 2009). Therefore, forest managers should treat young stands with care before applying this model.

The discussions associated with ANN have focused on prediction accuracy, but few studies have focused on biological reality. Since several factors affect the growth of a tree (or a forest stand), relations between variables and the significance of the variables are needed to biologically interpret parameter estimates of the models. The black-box structure into the internal workings of an ANN used for training, on the other hand, complicates biological interpretations of model parameters, inferences on the population of interest, and comparisons of differences among subjects (e.g., trees in this study). Yee et al. (1993), Aert-

sen et al. (2010) and Jevšenak & Levanič (2016) evaluated some modeling techniques such as ANN and generalized additive model (GAM) in terms of statistical, ecological interpretability, and user-friendliness, and suggested that other modeling techniques (e.g. GAM) may be preferred over ANN to predict site index. Also, Yee et al. (1993) reported the lack of biological interpretation as one of the disadvantages of ANN, and stated that assessment of an ecological model should be based on both the fit statistics and the ability to describe growth curves. Brosofske et al. (2014) commented that if a researcher's aim is to improve the accuracy of predictions, ANN may be a good choice in the case of a large dataset, but not a small dataset. In a case study (Castaño-Santamaría et al., 2013), ANN resulted in poor predictions in an uneven-aged forest due to very high height-diameter diversity, which resulted from different site qualities that obscured the training of ANN. They also explained that including diameter at breast height variance to each diameter class resulted in ANN being able to provide similar results with a local-mixed model.

The shape of the fitted curve is highly important for a reliable dominant height model (Guan & Gertner, 1991). However, the overfitting issue, as reported in previous studies by Aertsen et al. (2010), Özçelik et al. (2013) and Jevšenak & Levanič (2016) turns out to be decisive in assuring desirable properties of dominant height growth. To

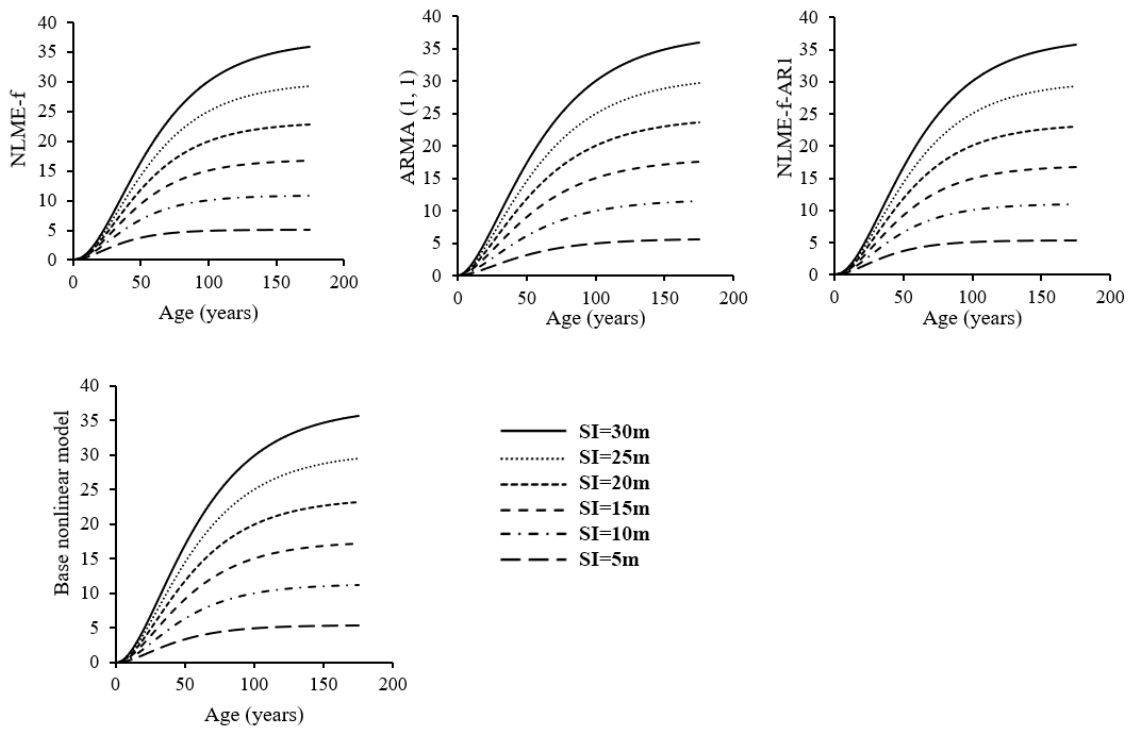


Figure 6. Site index curves for five site index values for Oriental spruce stands using regression techniques, the base nonlinear model, ARMA (1, 1), NLME-f, and NLME-f-ARI.

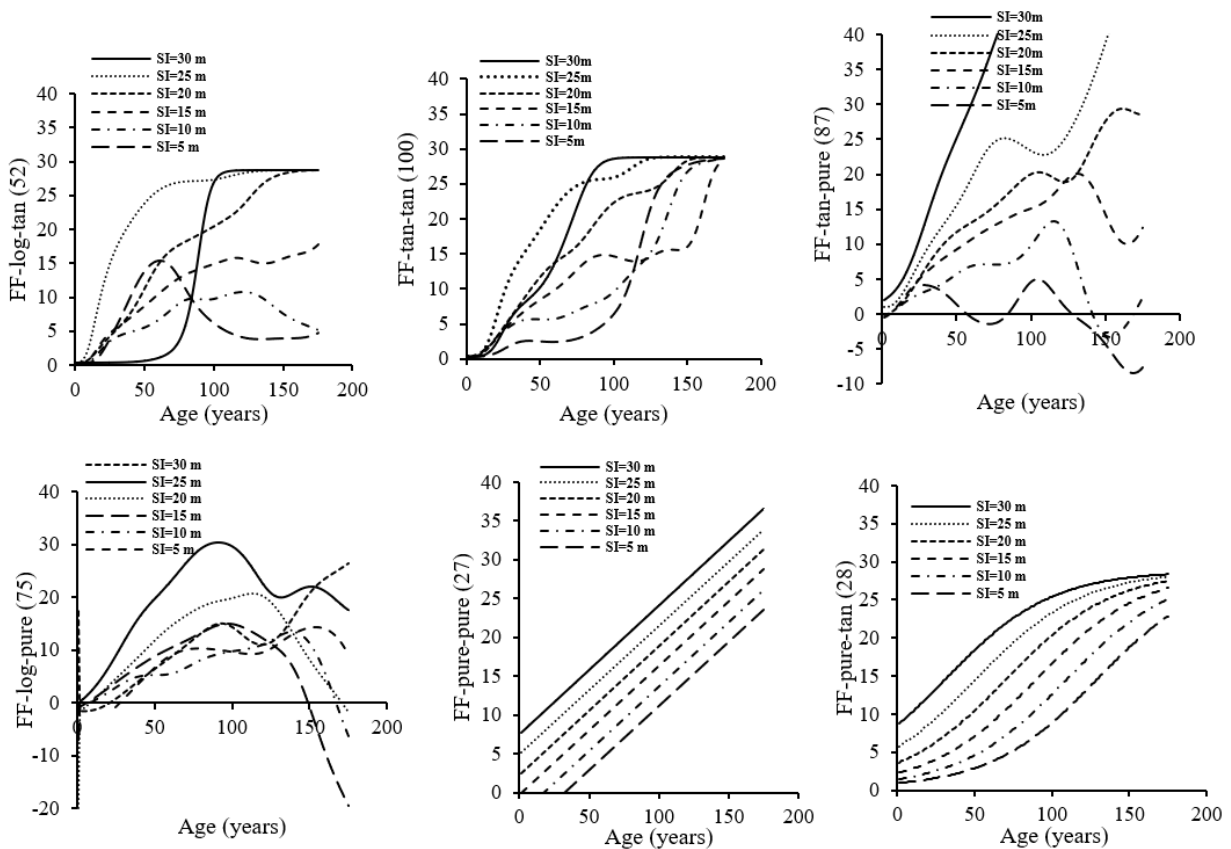


Figure 7. Site index curves for five site index values for Oriental spruce stands using feed-forward network architectures, FF-log-tan (52), FF-tan-tan (100), FF-tan-pure (87), FF-log-pure (75), FF-pure-pure (77), and FF-pure-tan (28).

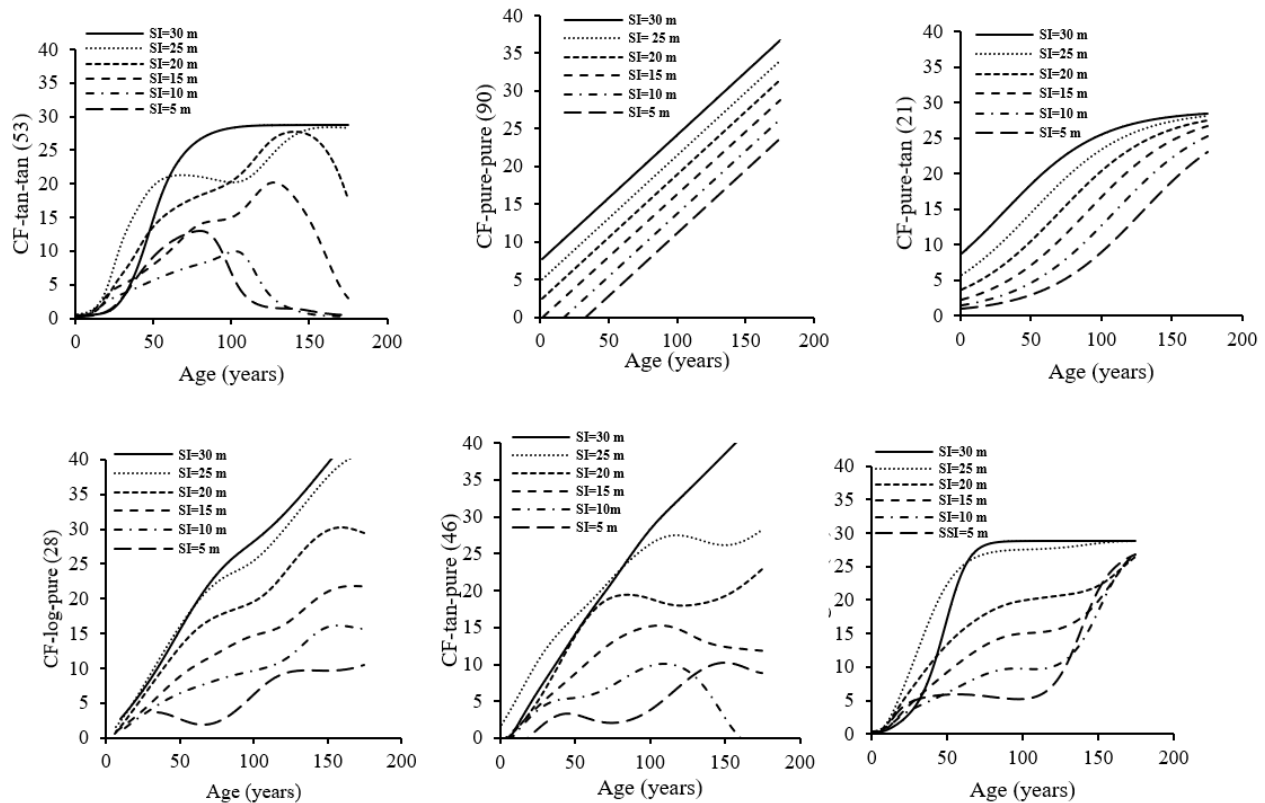


Figure 8. Site index curves for five site index values for Oriental spruce stands using cascade-forward network architectures, CF-log-tan (24), CF-tan-tan (53), CF-tan-pure (46), CF-log-pure (28), CF-pure-pure (90), and CF-pure-tan (21).

overcome this problem, an alternative way is to use the Levenberg-Marquardt training algorithm, as recommended by Jevšenak & Levanič (2016). However, in our study, we failed to avoid overfitting in the validation data set. The required properties of dominant height growth were reported by Goelz & Burk (1992). Accordingly, our results showed that most ANN architectures could not provide well-suited fitted curves in terms of dominant height growth (Figs. 4 and 5). Only FF-purelin-tansig with 28 neurons and CC-purelin-tansig with 21 neurons enabled a few required properties, including polymorphism and sigmoidal pattern, suggesting that the ANN needs improving to meet biological presumptions in terms of individual dominant height growth.

Conclusion

NLMs with a single random parameter were better at fitting the stem analysis data when compared with NLMs with two or three parameters. The worst estimates among the NLMs were obtained by those with three parameters having approximately four times greater RMSE when compared to the base nonlinear model. In the data obtained by the stem analysis, the spatial autocorrelation was more evident than the temporal autocor-

relation. These regression techniques had a homogeneous residual variance and had greater correlations with the measurements. The desirable properties for dominant height growth were provided by these regression techniques, such as multiple asymptotes, base age invariance, and sigmoidal pattern.

ANNs had worse estimates than the regression techniques and did not fulfill the desirable properties for dominant height growth. ANNs suggested very different curves depending on the transfer functions. This study showed that while selecting tan or log sigmoid transfer functions for the input layer was more appropriate in terms of the fit statistics, the purelin transfer function was more suitable in terms of ensuring the desirable properties. Selecting a purelin transfer function for the input layer and a tan-sigmoid transfer function for the output layer provided some desirable properties, including polymorphism and sigmoidal pattern. In the output layer, choosing a tan-sigmoid other than the purelin function, which presents a linear trend, will be more likely to provide a nonlinear trend compatible with polymorphism and sigmoidal pattern. Also, this study showed that determining an accurate transfer function was essential for the biological assumptions but was not sufficient. Considering the learning rate factor may increase the model efficiency and provide opportunities for ensuring the assumptions of interest.

Authors' contributions

Conceptualization: İ. Ercanli, H. Yavuz
Data curation: İ. Ercanli, F. Bolat
Formal analysis: İ. Ercanli, F. Bolat
Funding acquisition: Not applicable.
Investigation: İ. Ercanli, F. Bolat
Methodology: İ. Ercanli, F. Bolat
Project administration: Not applicable.
Resources: İ. Ercanli, H. Yavuz
Software: İ. Ercanli, F. Bolat
Supervision: İ. Ercanli, H. Yavuz
Validation: İ. Ercanli, F. Bolat
Visualization: İ. Ercanli, F. Bolat
Writing – original draft: İ. Ercanli, F. Bolat, H. Yavuz
Writing – review & editing: İ. Ercanli, F. Bolat, H. Yavuz

References

- Aertsen W, Kint V, Van Orshoven J, Özkan K, Muys B, 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol Modell* 221: 1119-1130. <https://doi.org/10.1016/j.ecolmodel.2010.01.007>
- Bailey RL, Clutter JL, 1974. Base-age invariant polymorphic site curves. *For Sci* 20: 155-159.
- Brosofske KD, Froese RE, Falkowski MJ, Banskota A, 2014. A review of methods for mapping and prediction of inventory attributes for operational forest management. *For Sci* 60: 733-756. <https://doi.org/10.5849/forsci.12-134>
- Calama R, Montero G, 2004. Interregional nonlinear height diameter model with random coefficients for stone pine in Spain. *Can J For Res* 34: 150-163. <https://doi.org/10.1139/x03-199>
- Calegario N, Daniels RF, Maestri R, Neiva R, 2005. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal eucalyptus plantation case study. *For Ecol Manag* 204: 11-21. <https://doi.org/10.1016/j.foreco.2004.07.051>
- Carmean WH, 1972. Site index curves for upland oaks in the Central States. *For Sci* 18: 109-120.
- Castaño-Santamaría J, Crecente-Campo F, Fernández-Martínez JL, Barrio-Anta M, Obeso JR, 2013. Tree height prediction approaches for uneven-aged beech forests in north-western Spain. *For Ecol Manag* 307: 63-73. <https://doi.org/10.1016/j.foreco.2013.07.014>
- Chapman DG, 1961. Statistical problems in population dynamics. *Proc 4th Symp on Mathematical Statistics and Probability*, Univ of California Press, Berkeley, pp. 153-168.
- Cieszewski CJ, 2002. Comparing fixed-and variable-base-age site equations having single versus multiple asymptotes. *For Sci* 48: 7-23.
- Cieszewski CJ, Bailey R, 2000. Generalized algebraic difference approach: theory based derivation of dynamic site equations with polymorphism and variable asymptotes. *For Sci* 46: 116-126.
- Cieszewski CJ, Harrison M, Martin SW, 2000. Practical methods for estimating non-biased parameters in self-referencing growth and yield models. University of Georgia PMRC-TR 7.
- Clutter JL, 1961. Development of compatible analytic models for growth and yield of loblolly pine. Ph.D. dissertation, Duke University, Durham, NC, USA.
- Corral-Rivas JJ, Álvarez González JG, Ruíz González AD, von Gadow K, 2004. Compatible height and site index models for five pine species in El Salto, Durango (Mexico). *For Ecol Manage* 201(2-3): 145-160. <https://doi.org/10.1016/j.foreco.2004.05.060>
- Curtis RO, Clendenan GW, Demars DJ, 1981. A new stand simulator for coast douglas-fir: DFSIM users guide. U.S. Forest Service General Technical Report PNW-128.
- Diéguez-Aranda U, Burkhart HE, Amateis RL, 2006. Dynamic site model for loblolly pine (*Pinus taeda* L.) plantations in the United States. *For Sci* 52 (3): 262-272.
- Ercanli İ, 2010. Trabzon ve Giresun Orman Bölge Müdürlükleri sınırları içerisinde yer alan Doğu Ladini (*Picea orientalis* (L.) Link)-Sarıçam (*Pinus sylvestris* L.) karışık meşcerelerine ilişkin büyüme modelleri. PhD, Graduate School of Natural and Applied Sciences, Karadeniz Technical University, Trabzon, Türkiye.
- Fabbio G, Frattegiani M, Manetti MC, 1994. Height estimation in stem analysis using second differences. *For Sci* 40: 329-340.
- Fortin M, Daigle, G, Ung C-H, Bégin J, Archambault L, 2007. A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling. *Eur J For Res* 126: 573-585. <https://doi.org/10.1007/s10342-007-0179-1>
- García O, 2005. Comparing and combining stem analysis and permanent sample plot data in site index models. *For Sci* 51: 277-283.
- Goelz J, Burk T, 1992. Development of a well-behaved site index equation: jack pine in north central Ontario. *Can J For Res* 22: 776-784. <https://doi.org/10.1139/x92-106>
- Grégoire TG, Schabenberger O, Barrett JP, 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can J For Res* 25: 137-156. <https://doi.org/10.1139/x95-017>
- Gregorie TG, 1987. Generalized error structure for forestry yield models. *For Sci* 33: 423-444.
- Guan BT, Gertner G, 1991. Modeling red pine tree survival with an artificial neural network. *For Sci* 37: 1429-1440.
- Jevšenak J, Levanič T, 2016. Should artificial neural networks replace linear models in tree ring based climate reconstructions? *Dendrochronologia* 40: 102-109. <https://doi.org/10.1016/j.dendro.2016.08.002>
- Lappi J, Bailey RL, 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *For Sci* 34: 907-927.

- Martín-Benito D, Gea-Izquierdo G, Del Río M, Canellas I, 2008. Long-term trends in dominant-height growth of black pine using dynamic models. For Ecol Manag 256: 1230-1238. <https://doi.org/10.1016/j.foreco.2008.06.024>
- MathWorks Inc., 2015. MATLAB Users guide, vers 2015a. The MathWorks, Inc., Natick, MA, USA. Matlab user manual.
- Nothdurft A, Kublin E, Lappi J, 2006. A non-linear hierarchical mixed model to describe tree height growth. Eur J For Res 125: 281-289. <https://doi.org/10.1007/s10342-006-0118-6>
- Özçelik R, Diamantopoulou MJ, Crecente-Campo F, Eler U, 2013. Estimating Crimean juniper tree height using non-linear regression and artificial neural network models. For Ecol Manag 306: 52-60. <https://doi.org/10.1016/j.foreco.2013.06.009>
- Pinheiro J, Bates D, 2006. Mixed-effects models in S and S-PLUS. Springer-Verlag, New York.
- Poudel KP, Cao QV, 2013. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. For Sci 59: 243-252. <https://doi.org/10.5849/forsci.12-001>
- Richards FJ, 1959. A flexible growth function for empirical use. J Exp Bot 10 (29): 290-300. <https://doi.org/10.1093/jxb/10.2.290>
- SAS Inst., 2004. SAS/ETS 9.1 Users Guide. SAS Institute, Cary, NC, USA.
- Sharma M, Subedi N, Ter-Mikaelian M, Parton J, 2014. Modeling climatic effects on stand height/site index of plantation-grown jack pine and black spruce trees. For Sci 61: 25-34. <https://doi.org/10.5849/forsci.13-190>
- Skovsgaard JP, Vanclay JK, 2008. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. Forestry 81: 13-31. <https://doi.org/10.1093/forestry/cpm041>
- Wang Y, LeMay VM, Baker TG, 2007. Modelling and prediction of dominant height and site index of *Eucalyptus globulus* plantations using a nonlinear mixed-effects model approach. Can J For Res 37: 1390-1403. <https://doi.org/10.1139/X06-282>
- Wang M, Borders BE, Zhao D, 2008. An empirical comparison of two subject-specific approaches to dominant heights modeling: The dummy variable method and the mixed model method. For Ecol Manag 255: 2659-2669. <https://doi.org/10.1016/j.foreco.2008.01.030>
- Wang M, Bhatti J, Wang Y, Varem-Sanders T, 2011. Examining the gain in model prediction accuracy using serial autocorrelation for dominant height prediction. For Sci 57: 241-251.
- Wang M, Kane MB, Borders BE, Zhao D, 2013. Direct variance-covariance modeling as an alternative to the traditional guide curve approach for prediction of dominant heights. For Sci 60: 652-662. <https://doi.org/10.5849/forsci.13-019>
- Weiskittel AR, Hann DW, Hibbs DE, Lam TY, Bluhm AA, 2009. Modeling top height growth of red alder plantations. For Ecol Manag 258: 323-331. <https://doi.org/10.1016/j.foreco.2009.04.029>
- Yee D, Prior M, Florence L, 1993. Development of predictive models of laboratory animal growth using artificial neural networks. Bioinformatics 9: 517-522. <https://doi.org/10.1093/bioinformatics/9.5.517>