

Total tree height predictions *via* parametric and artificial neural network modeling approaches

Yasin Karatepe ⁽¹⁾,
Maria J Diamantopoulou ⁽²⁾,
Ramazan Özçelik ⁽¹⁾,
Zerrin Sürücü ⁽³⁾

Height-diameter relationships are of critical importance in tree and stand volume estimation. Stand description, site quality determination and appropriate forest management decisions originate from reliable stem height predictions. In this work, the predictive performances of height-diameter models developed for Taurus cedar (*Cedrus libani* A. Rich.) plantations in the Western Mediterranean Region of Turkey were investigated. Parametric modeling methods such as fixed-effects, calibrated fixed-effects, and calibrated mixed-effects were evaluated. Furthermore, in an effort to come up with more reliable stem-height prediction models, artificial neural networks were employed using two different modeling algorithms: the Levenberg-Marquardt and the resilient back-propagation. Considering the prediction behavior of each respective modeling strategy, while using a new validation data set, the mixed-effects model with calibration using 3 trees for each plot appeared to be a reliable alternative to other standard modeling approaches based on evaluation statistics regarding the predictions of tree heights. Regarding the results for the remaining models, the resilient propagation algorithm provided more accurate predictions of tree stem height and thus it is proposed as a reliable alternative to pre-existing modeling methodologies.

Keywords: Tree Height Model Prediction, Generalized Models, Mixed-Effects Models, Levenberg-Marquardt Algorithm, Resilient Propagation

Introduction

Natural forest areas have continued to decline in the last 30 years while plantations have increased in the same timeframe, now comprising some 7% of the global forest area. Plantations are of great importance in economic and ecological terms. Thus, robust and reliable information on the growth and development characteristics of these species are essential in order to sustainably manage plantations and determine the success of afforestation efforts. Diameter at breast height (d) of trees and total tree-bole height (h) are two basic commonly used parameters in forest inventory, planning and management of plantations. These parameters are fundamental variables in several forest practices such as prediction of standing yield,

growth projection, carbon accounting, identification of stand structural diversity, determination of the extent of damage in forest or stands, and to estimate missing tree heights (Curtis 1967, Adame et al. 2008).

In forest management applications or field inventories, d in trees can be measured rapidly and precisely, while h is costly, more time consuming, and challenging to measure (Meng & Huang 2009, Sharma & Parton 2007, Özçelik et al. 2018). It is also known that there is a strong correlation between d and h in the stem of trees. Because of this, models can be created to discover the relationship between d and h based on the tree diameter at breast height and the total stem height of trees that are measured. Ultimately, these mod-

els are used to estimate the height of any tree in the stand with unknown height (Robinson & Wykoff 2004, Lei et al. 2009).

There are many studies describing tree height estimation and prediction models, mainly using well-known classical modeling techniques, such as non-linear regression models (Colbert et al. 2002, Kearsley et al. 2017) and Bayesian modeling (Zhang et al. 2014), using in most cases allometric models. Further research evaluating more recent and promising modeling methods such as techniques based on the artificial intelligence are warranted to improve efficacy. Previous research on height-diameter (h - d) modeling of cedar plantations in Turkey was conducted by Catal (2012), although it was based on a limited number of traditional h - d models, without further investigation of the potential efficiency of other modeling approaches.

Generally, h - d models may be simple or generalized. Simple models utilize d to calculate h , as the relationship between the two variables is well-known (Bronisz & Mehtätalo 2020). The limitation with simple models, however, is the large sampling effort required (Gómez-García et al. 2014). In contrast, the correlation between d and h (total height) varies among species and can be influenced by stand-specific variables such as stand structure and site productivity (Lhotka 2012). Models that include measures of stand structure, site productivity and tree social position are commonly referred to as generalized h - d (GM)

□ (1) Faculty of Forestry, Isparta University of Applied Sciences, East Campus, 32260, Isparta (Turkey); (2) Faculty of Agriculture, Forestry and Natural Environment, School of Forestry and Natural Environment, Aristotle University of Thessaloniki, GR-54124 Thessaloniki (Greece); (3) Ministry of Agriculture and Forestry, VI. Regional Directorate, 27002, Burdur (Turkey)

@ Maria J Diamantopoulou (mardi.diamant@gmail.com)

Received: Oct 04, 2021 - Accepted: Jan 11, 2022

Citation: Karatepe Y, Diamantopoulou MJ, Özçelik R, Sürücü Z (2022). Total tree height predictions *via* parametric and artificial neural network modeling approaches. *iForest* 15: 95-105. - doi: [10.3832/ifor3990-015](https://doi.org/10.3832/ifor3990-015) [online 2022-03-21]

Communicated by: Angelo Rita

models. Mehtätalo et al. (2015) used fixed-effects (FE) and generalized mixed-effects (ME) modeling techniques to show differences between marginal and plot-specific *h-d* relationships, in a range of different geographical/ecological regions and tree species. The use of additional stand-level variables was previously shown to improve tree height estimates (Vargas-Larreta et al. 2009, Lhotka 2012). For example, Newton & Amponsah (2007) used dominant height (H_0) within a stand to estimate stand-level competition.

Since the estimation and prediction of tree heights are crucial in tree growth and tree volume prediction, alternative and accepted modeling techniques exist such as the generalized, FE and ME models, and artificial neural network modeling techniques, that may be useful in modeling the relationship between *h-d* and reducing the measurement efforts on the ground. The most basic (Huang et al. 1992, Peng et al. 2001, Soares & Tomé 2002) and generalized *h-d* models (Gómez-García et al. 2015, Adame et al. 2008, Vargas-Larreta et al. 2009, Crecente-Campo et al. 2010, Lhotka 2012, Zang et al. 2016, Ozçelik et al. 2018) are available for many tree species. Most research has focused on mixed-models to understand *h-d* relationships (Calama & Montero 2004, Sharma & Parton 2007, Lhotka 2012, Ozçelik et al. 2013, 2018, Gómez-García et al. 2014, 2015, Mehtätalo et al. 2015, Adamec & Drápela 2016, Bronisz & Mehtätalo 2020). In Turkey, *h-d* equations were developed for some tree species at regional scale (Diamantopoulou & Ozçelik 2012, Ozçelik et al. 2014, Catal & Carus 2018).

Taurus cedar (*Cedrus libani* A. Rich.) natural range includes the Taurus Mountains of Turkey, where it is an ecologically and economically important component of the forest ecosystems (Boydak 2003). The most recent inventory data show that cedar covers an area of approximately 463,521 ha in Turkey, with a total growing stock of 27.4 million m³ (GDF 2015). It is not only an important resource of raw material for the

forest product industry, but it also fulfills critical ecological tasks including the reduction of soil erosion and the conservation of water resources, the mitigation of the adverse impacts of climate change, and is essential in the maintenance of biological diversity in Turkey. It has moderate soil requirements and is tolerant to extremes of temperature and drought in summer and to cold temperatures in winter (Saatçioğlu 1979); therefore, *C. libani* has great potential for afforestation, especially in areas with semi-arid climates. Given that 35% of Turkey's territory has semi-arid climate, cedar is a critical tree species for use in afforestation. As of 2000, an area of approximately 110,000 ha was re-afforested with cedar (Konukçu 2001). However, there is still relatively little information concerning the growth of Taurus cedar plantation compared with the quantity of information available for natural cedar stands. *H-d* relationship is one of the most important components of growth and yield models for sustainable management of Taurus cedar plantations. Moreover, to our best knowledge, *h-d* generalized mixed-effects model has not been developed for Taurus cedar plantations.

Until now, the use of artificial neural networks (ANNs) in forest modeling has led to the conclusion that they can be considered as significant alternative techniques for many characteristics of trees growth compared to classical modeling methods, both in classification tasks (Schmoldt et al. 1997, Liu et al. 2003, Cosenza et al. 2017), and for estimation and prediction problems (Leite et al. 2011, Soares et al. 2013, Reis et al. 2016, Ozçelik et al. 2017, Ercanli et al. 2018, Monteiro Da Silva et al. 2018, Vacchiano et al. 2018, Wu et al. 2019, Sun et al. 2021). Specifically, artificial intelligence was successfully implemented to the development of total tree height models as well (Li & Jiang 2010, Diamantopoulou & Ozçelik 2012, Ozçelik et al. 2013, Vieira et al. 2018, Thanh et al. 2019, Ercanli 2020). The majority of the ANN modeling studies employed multi-

layer perceptron architecture combined by the standard backpropagation algorithm.

The incentive for developing an artificial neural network modeling approach in the present work was that ANNs can discover and thus automatically model the relationships underlying input and output variables, a property reflected in the connection weights of the network. Instead, traditional approaches, such as the widely accepted generalized, fixed and mixed-effects modeling methods, require certain assumptions about the form of a fitting function which must be specified in advance, introducing limitations that mainly impact model applicability to different scenarios. In contrast, an ANN model is trained to find this relationship, thus fitting complex nonlinear models. In addition, given that ANNs provide resilience to outliers, the techniques generally perform well where there are missing or inaccurate data; this problem is typically found in forest-data measurements which require substantial effort in-the-field, increasing proneness to errors. Furthermore, this generalization ability enables ANN models to generate reliable predictions with new data sets. ANN modeling techniques, therefore, can overcome problems in basic data gathered in the forest, including nonlinear relationships, multicollinearity, heteroscedasticity, outliers and noise. Although ANNs may suffer from over-fitting of data, this problem can be avoided by selecting a suitable training architecture followed by testing and validation.

The objective of this work was to compare the effectiveness of different modeling approaches in providing accurate predictions of total stem height of *C. libani* plantations of Turkey. After developing the ANN models, validation was carried out using simple measurements from the forest environment, such as *d*. Variability between sampling plots was included in the models by recording dominant heights (H_0) and diameters (D_0) of each sample plot. The adjusted fixed-and mixed-effects models, after being localized using calibration data from one to three sample cedar trees were compared against the prediction abilities of the novel ANN models. For this purpose, two different learning algorithms, that have shown significant ability to cope the known disadvantages and limitations of the most used standard backpropagation learning algorithm (Wu & Ji 2015), i.e., the resilient propagation (RPANN) and the Levenberg-Marquardt (LMANN), were applied as innovative modeling strategies in forest model building, with the aim of optimizing the learning process and better exploring the information in real data measurements originating from the forests.

Materials and methods

Site description

Data used for development of models were gathered in cedar plantations. Sam-

Tab. 1 - Summary statistics for the fitting and the validation data sets. (*d*): diameter at breast height (1.3 m above the ground); (*h*): total tree height; (H_0): dominant height; (D_0): dominant diameter; (*G*): basal area per hectare; (d_m): mean diameter; (h_m): mean height; (D_g): quadratic mean diameter; (SD): standard deviation.

Variable	Fitting data (70%) (630 trees in 21 plots)				Validation data (30%) (270 trees in 9 plots)			
	Mean	Min	Max	SD	Mean	Min	Max	SD
<i>d</i> (cm)	10.31	3.20	20.00	3.12	11.37	5.40	20.00	3.50
<i>h</i> (m)	6.12	3.00	11.75	1.42	6.66	3.20	11.00	1.81
Age (years)	26.00	17.00	35.00	6.30	31.00	23.00	35.00	4.70
D_0 (cm)	13.96	9.72	18.30	2.38	14.86	7.82	19.24	3.49
H_0 (m)	7.51	5.44	10.02	1.21	8.30	5.66	10.65	1.67
<i>G</i> (m ² ha ⁻¹)	30.30	12.75	50.70	10.86	36.96	12.01	58.25	15.02
d_m (cm)	10.29	6.68	13.53	2.02	11.35	6.74	14.64	2.51
h_m (m)	6.10	4.31	7.93	1.03	6.66	4.50	8.65	1.37
D_g (cm)	10.57	6.98	13.92	2.08	11.60	6.77	14.92	2.59

ple plots were in the Isparta Regional Directorate of Forestry, which includes the Isparta and Burdur provinces (south-western Turkey), at an elevation between 1000 to 1600 m a.s.l. Climate in this region is characterized as transitional from the Mediterranean to continental climate. Annual precipitation varies between 426 to 814 mm, while annual mean temperature ranges from 11.9 to 13.2 °C (MGM 2013). The bed-rock of the study site is composed of sedimentary rock including limestone, sandstone, and claystone.

Data collection

To determine diameter-height associations, data were collected from 30 sampling plots randomly distributed among Taurus cedar plantations in Isparta and Burdur Forest Regions. Sample plots were selected to represent a range of age classes, site conditions, and densities for the species in the Mediterranean Region of Turkey. Plot size varied between 270 to 540 m², depending on density, with a minimum of 30 trees per plot. The age of the sampling plots varied from 17 to 35. Two perpendicular over-bark diameters (d_i) were measured with a precision of 0.1 cm and averaged to find actual diameter (d , cm) for each tree. A Blume-Leiss hypsometer was used to obtain total heights with precision of 0.5 m. Depending on plot size, dominant heights (H_o) and diameters (D_o) were obtained as the average heights and diameters of the 100 trees with the thickest diameter at breast height (d) per hectare.

The available Cedar data set was randomly split into two groups; some 21 sample plots (70% of all plots) were randomly selected for the model fitting while the remaining 9 sample plots (30% of all plots) were used for evaluating the models' performance on this new data set (Tab. 1). The basic descriptive statistics for the dataset used has been given in Tab. 1 and illustrated in Fig. 1.

Generalized height-diameter models (GM)

Twenty-one generalized height-diameter models selected from earlier studies were evaluated using non-linear least squares (NLS – Crecente-Campo et al. 2010, Corral-Rivas et al. 2014). These models included different stand-level variables such as H_o , D_o , and quadratic mean diameter, to represent the variation between forest stands.

Fixed-effects model (FE)

A large number of nonlinear model forms with two and three parameters were evaluated for cedar plantations, including those noticed by Huang et al. (1992) and Lei et al. (2009) as Curtis, Weibull, Exponential, Chapman-Richards, Gompertz, Schnute, and Korf-Lundqvist. The NLS were fitted to the datasets to determine the most appropriate model for cedar plantations. To evaluate the prediction accuracy of the models, statistical and graphical evaluations

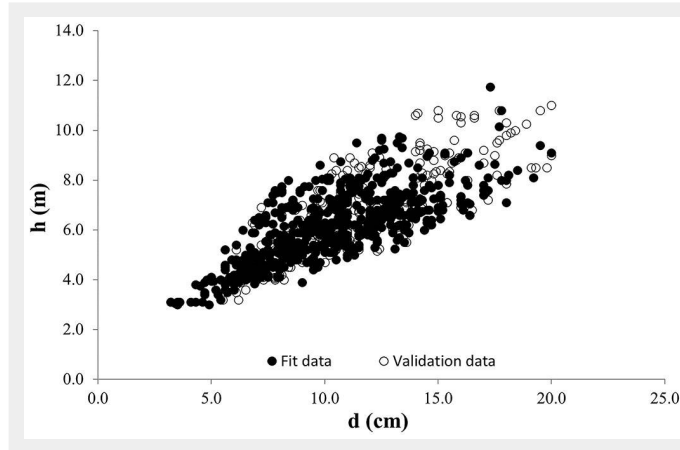


Fig. 1 - Scatter plot of total height (h) against diameter at breast height (d) for cedar pine trees for both fitting and validation data sets.

were used. Of the models assessed, predictive capability of the Chapman-Richards appeared best to model height-diameter relationships in cedar plantations (eqn. 1):

$$h_{ij} = 1.3 + \beta_1 (1 - \exp(-\beta_2 d_{ij}))^{\beta_3} + \varepsilon_{ij} \quad (1)$$

where h_{ij} is the total tree height (m) of the j -th tree in the i -th plot, d_{ij} is the diameter (cm), and β_1 , β_2 , β_3 are model parameters.

To calibrate the FE model, a correction factor k^* (eqn. 2) suggested by Temesgen et al. (2008) was used. Temesgen et al. (2008) indicated that when the heights of a subsample of n_{im} trees from the i -th stand is known, the predicted heights of the remaining trees from the same stand can be calibrated (eqn. 2):

$$k^* = \frac{\sum_{j=1}^{n_{im}} [(\hat{h}_{ni} - 1.30)(h_{ij} - 1.30)]}{\sum_{j=1}^{n_{im}} (\hat{h}_{ij} - 1.30)} \quad (2)$$

where \hat{h}_{ni} is the predicted height from the FE model (eqn. 1), and h_{ij} is the observed value.

The predicted values of height (\tilde{h}) for the remaining trees in the same stand can be calculated as follows (eqn. 3):

$$\tilde{h}_{ij} = 1.3 + k^* \beta_1 [1 - \exp(-\beta_2 d_{ij})]^{\beta_3} \quad (3)$$

The influence of the correction factor on the prediction accuracy of a nonlinear FE model was also investigated.

Nonlinear mixed-effects model (ME)

In the ME model structure, parameters of eqn. 1 comprise the plot-specific random-effect parameters and the population level fixed-effect parameters that are common to all trees. In matrix notation, eqn. 1 can be represented as (eqn. 4):

$$y_i = f(\mathbf{b}, \mathbf{u}_i, \mathbf{d}_i) + \varepsilon_i \quad (4)$$

where $\mathbf{y}_i = [h_{i1}, h_{i2}, \dots, h_{in}]^T$, $\mathbf{d}_i = [d_{i1}, d_{i2}, \dots, d_{in}]^T$, $\varepsilon_i = [\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{in}]^T$, n_i is the sample size for tree height values in plot i , and \mathbf{u}_i and \mathbf{b} are column vectors representing random- and fixed-effects parameters, respectively; $\varepsilon_i \sim N(0, \mathbf{R})$ and $\mathbf{u}_i \sim N(0, \mathbf{D})$, where \mathbf{R} and \mathbf{D} are diagonal matrices. Additionally, it is assumed that \mathbf{u}_i and ε_i are independent.

SAS procedure NLMIXED (SAS Institute 2010) was used to obtain the parameters of nonlinear ME models (eqn. 4).

The first-order Taylor series expansion can be used to obtain the random-effect parameters \mathbf{u}_i for the i -th plot (Hall & Clutter 2004, Meng & Huang 2009 – eqn. 5):

$$\hat{\mathbf{u}}_i^{k+1} = \hat{\mathbf{D}} \mathbf{Z}_i^T (\mathbf{Z}_i \hat{\mathbf{D}} \mathbf{Z}_i^T + \hat{\mathbf{R}})^{-1} [\mathbf{y}_i - f(\hat{\mathbf{b}}, \hat{\mathbf{u}}_i^k, \mathbf{d}_i) + \mathbf{Z}_i \hat{\mathbf{u}}_i^k] \quad (5)$$

where $\hat{\mathbf{u}}_i^k$ is an estimate of the random-effect parameters for plot i at the k -th iteration, $\hat{\mathbf{D}}$ is an estimate of \mathbf{D} , the variance-covariance matrix for \mathbf{u}_i , and \mathbf{Z}_i is (eqn. 6):

$$\mathbf{Z}_i = \left. \frac{\partial f(\mathbf{b}, \mathbf{u}_i, \mathbf{d}_i)}{\partial \mathbf{u}_i} \right|_{\hat{\mathbf{b}}, \hat{\mathbf{u}}_i^k} \quad (6)$$

$\hat{\mathbf{R}}$ is an estimate of \mathbf{R} , the variance-covariance matrix for ε_i , \mathbf{y}_i is the $m \times 1$ vector of observed tree heights, and m is the number of tree height measurements used to localize the height growth model.

Estimating \mathbf{u}_i requires an iterative process. In eqn. 5, a null starting value ($\hat{\mathbf{u}}_i^0 = 0$) was used and repeatedly updated until the absolute difference between $\hat{\mathbf{u}}_i^k$ and $\hat{\mathbf{u}}_i^{k+1}$ was smaller than a predetermined tolerance limit. For the random effects, the outcome is the approximated empirical best linear unbiased predictor (EBLUP).

ANN models

The most significant advantage of an ANN model is that it is trained from measured data without the need to add further information. That is, after the fitting of the system with the available ground-truth data, the ANN model automatically discovers the existing dependencies, and finally produces a trained model. On the other hand, the net topology, the learning algorithms used, and the learning constraints are specified by the modeler.

Levenberg-Marquardt artificial neural network model (LMANN)

The Levenberg-Marquardt (LM) algorithm was selected for application in the multilayer perceptron learning of ANN (Levenberg 1944, Marquardt 1963). Past studies showed that the LM algorithm has the ability, on the one hand, to cope with

known disadvantages and limitations of the standard backpropagation learning algorithm, such as slow convergence, much off-line training requirements, instability, trapped at local minima (Adeloye & De Munari 2006, Wilamowski & Yu 2010, Scanlan & Mulvaney 2013, Wu & Ji 2015); on the other hand, LM algorithm have showed significant ability on the prediction of environmental variables (Wu 2014, Diamantopoulou et al. 2015, Zhou et al. 2018, Ozçelik et al. 2019).

A detailed description of how the Levenberg-Marquardt algorithm is embedded to neural network training can be found in Hagan & Menhaj (1994). Concisely, the Hessian matrix (H_m) approximation is introduced by the LM algorithm as (Ozçelik et al. 2019 – eqn. 7):

$$H_m = J^T J + \mu I \tag{7}$$

where I is the identity matrix, J is the Jacobian matrix, and μ is the combination coefficient.

The update rule of the LM algorithm is $w_{t+1} = w_t - (H_m)^{-1} J^T e$, where w are the weights, and e are the biases. The efficiency and convergence of the LMANN models are very sensitive to the configuration of the coefficient value (μ) of eqn. 7. As the value becomes larger, the more weight is given to gradient descent learning with a small step size. The smaller it is, the more weight is given to large step sizes. For μ equal to zero, the algorithm turns to the Gauss-Newton method.

Resilient propagation artificial neural network model (RPANN)

Resilient propagation, known as Rprob, is a neural network algorithm which was fully described by Riedmiller & Braun (1993) and successfully used (Saputra et al. 2017, Florescu & Igel 2018), so that known disadvantages of the standard back-propagation algorithm can be overcome. The back-propagation algorithm has known drawbacks, due to elementary learning processes. The main problems of the learning phase include slow convergence resulting from

small changes in weights, and the neural network weights trapping around local optima.

Rprob mainly refers to the gradient direction. The algorithm calculates a delta value (Δ_{ij}) for each weight that increases when the gradient does not change sign under the prerequisite that the system works on the right direction, or decreases when the gradient does change sign. The learning rule used to calculate delta value which improves the change of the network weights is the following (Riedmiller & Braun 1993 – eqn. 8):

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \cdot \Delta_{ij}^{(t-1)} & \text{if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \cdot \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ \eta^- \cdot \Delta_{ij}^{(t-1)} & \text{if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \cdot \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)} & \text{otherwise} \end{cases} \tag{8}$$

where $0 < \eta^- < 1 < \eta^+$.

Finally, the new weight value w_{ij}^t between i and j neurons in two consecutive layers on the $(t-1)$ is given by (eqn. 9):

$$w_{ij}^t = w_{ij} + \Delta w_{ij}^{t-1} \tag{9}$$

Aspects of ANN model training

In geotechnical engineering, the selection of the proper input variables for ANN model building is usually based on *a priori* knowledge of the physical problem (Maier & Dandy 2000). In order to construct a reliable model for total stem height prediction, the sensitivity analysis was used to choose the proper input layer nodes of the ANNs. The analysis was performed on the available variables measured in the field (Tab. 1), that were selected in advance due to the physical problem (Cariboni et al. 2007).

To avoid overfitting and undertraining of the network in reaching the best possible training efficiency of the model, the training process was initiated with the smallest possible network and no hidden nodes. Then the construction of the model continued by adding any necessary hidden nodes

in order the network learning to be completed. Finally, the architecture is reached to an end, when the agreement obtained between ANN model predictions and targets can be considered as satisfactory.

Overfitting avoidance of the ANN models led to acceptable generalization ability of the models. For this reason, the k -fold cross validation resampling method with $k=10$ was applied (Olson & Delen 2008, May et al. 2010) to both training and testing data sets (Leahy 1994), due to its efficiency and easiness to apply. This procedure was not applied in regression model building, where there is little to be gained by separating data into parts (Hursch 1991). As described above, the whole data set was randomly divided into two parts. The fitting data set was used for the construction of all different types of models. The validation data set was used only for the evaluation of the predictive capability of all constructed models.

Due to the fact that the ANN model error value is strongly affected by the number of the weights, the credibility and the generalization capability of the model was determined through the mean error rate on the cross-validation examples, and the validation data set, respectively (Fig. 2).

As mentioned above, the efficiency and convergence of training of the LMANN models significantly depends on the proper value selection for the combination coefficient (μ in eqn. 7). The initial value of μ was set to 0.001 (Diamantopoulou et al. 2015) and the final value that led to the lowest sum of square errors model value was found equal to 1.00e-07.

Similar to the LMANN models, training of RPANN models is very sensitive to the value of delta rule $\Delta_{ij}^{(t)}$ (eqn. 8), which is responsible for the sign (positive or negative) of the gradient to show the direction of the adjustment weight. Through this process the iteration of learning stops when the error target is reached. The initial delta value was 0.07, with increment steps of 1.2 and decrease steps of 0.5.

The architecture of both the LMANN and RPANN algorithms finalized after the detection of the correct number of nodes in the hidden layer. This final number was obtained during a trial-and-error process: the process began with one hidden node and was repeated until the desired target error value was reached. During the trial-and-error process, selection of the correct combination of transfer functions was made. The combination of functions that provided the optimal behavior of both the ANN models was: (i) the hyperbolic tangent transfer function (*tansig*) to transfer knowledge from the input layer to the hidden layer; and (ii) the linear activation function (*purelin*), transferring information from the hidden layer to the output layer (Fausett 1994). The ANN models were constructed using the neural network toolbox of the MATLAB software ver. R2017a (Beale et al. 2017) developed by The Math Works™ Inc.

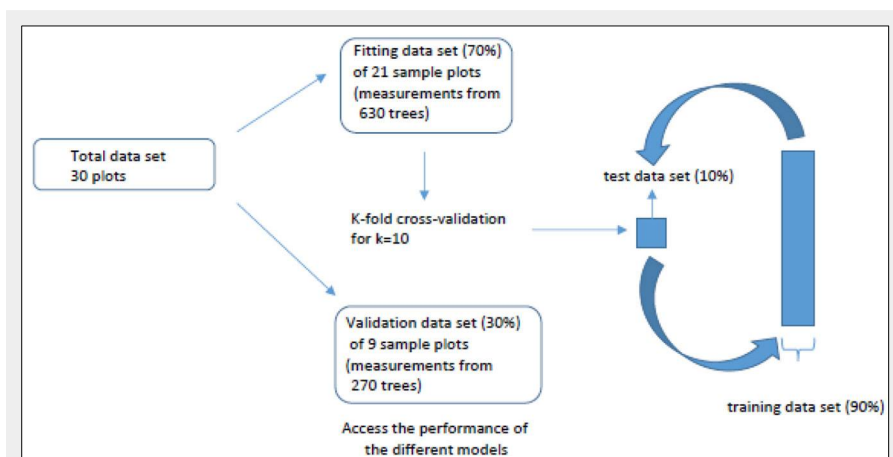


Fig. 2 - The 10-fold cross validation division for the ANN models construction.

Evaluation

The evaluation of the different modeling approaches (i) Generalized Model (GM), (ii) non-calibrated fixed-effects model [FE(0)], (iii) calibrated fixed-effect model [FE(1-3)], (iv) calibrated mixed-effects model [ME (1-3)], (v) LMANN, and (vi) RPANN techniques, was made using the mean difference (MD, m), the mean absolute difference (MAD, m), the fit index (FI) and the root square mean error (RSME), which were calculated as follows:

$$MD = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})}{\sum_{i=1}^n n_i} \tag{10}$$

$$MAD = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} |h_{ij} - \hat{h}_{ij}|}{\sum_{i=1}^n n_i} \tag{11}$$

$$FI = 1 - \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})^2}{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \bar{h}_i)^2} \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})^2}{\sum_{i=1}^n n_i}} \tag{13}$$

where n is number of plots, n_i is number of tree height values for plot i , h_{ij} and \hat{h}_{ij} are measured and predicted heights, respectively, and \bar{h}_i is mean height of h_{ij} for plot i .

It is worth noting, that the adjusted FE and the calibrated ME models' parameters were "localized" by the use of a number of sampled heights in each plot. Trincado et al. (2007) recommended that the number of tree heights measured in each plot ranged from one to three. Finally, to compare and evaluate the generalization ability of the constructed models by the different modeling approaches, the above metrics were computed for the validation data set as well (Ozçelik et al. 2018, Bronisz & Mehtälö 2020).

Results

Parametric modeling approaches (generalized models, fixed-and mixed-effects models)

Twenty-one GM models were fitted using the NLS method with the fitting data set, while the predictive capability of the models was evaluated using the validation data set. The final GM model (Krumland & Wensel 1988) including D_o and H_o as additional independent variables, was (eqn. 14):

$$h = 1.3 + (H_o - 1.3) \frac{\exp[\beta_1 d^{\beta_2 (H_o - 1.3)}]}{\exp[\beta_1 D_o^{\beta_2 (H_o - 1.3)}} \tag{14}$$

The model given in eqn. 14 and its modified forms was successfully used by many authors (e.g., Crecente-Campo et al. 2010). H_o was the most important contributor among stand level variables since it can allow a plot-level adaptation related to many factors, such as genetics, topography, silvicultural regime, and environment (Huang et al. 2009). Further, the accuracy of h - d models can be improved by the incorpora-

Tab. 2 - Fit statistics for eqn. 1 for different combinations of random parameters.

Random Parameters	AIC (smaller is better)	BIC (smaller is better)
None (Fixed-effects)	2503	2522
β_1	1617	1624
β_2	1821	1828
β_3	1605	1612
β_1 and β_2	1540	1550
β_1 and β_3	1539	1548
β_2 and β_3	1714	1724
β_1, β_2 and β_3	1656	1670

Tab. 3 - Parameter estimates and standard errors (in parentheses) for the fixed- and mixed-effects regression models. (σ^2): residual variance of the model; ($\sigma^2_{u_1}$): variance of the random-effect u_1 ; ($\sigma^2_{u_2}$): variance of the random-effect u_2 ; ($\sigma^2_{u_1 u_2}$): covariance of the random-effects u_1 and u_2 .

Parameter estimates	Generalized Model	Fixed-effects	Mixed-effects
β_1	-6.9087 (0.5670)	7.8237 (0.6187)	24.7062 (11.3829)
β_2	-0.0294 (0.0045)	0.1348 (0.0315)	0.0159 (0.0113)
β_3	-	1.5604 (0.2998)	0.8643 (0.0750)
σ^2	-	0.8311 (0.0468)	0.2216 (0.0129)
$\sigma^2_{u_1}$	-	-	73.9740 (113.22)
$\sigma^2_{u_2}$	-	-	0.0448 (0.0176)
$\sigma^2_{u_1 u_2}$	-	-	1.6671 (1.3800)

tion of D_o and H_o as covariates (Crecente-Campo et al. 2010, Raptis et al. 2021).

A series of parameters from eqn. 1 was implemented (Tab. 2). Model (1) with random components β_1 and β_3 resulted in the lowest values for both Akaike's (AIC) and Bayesian (BIC) information criteria. The final ME model was (eqn. 15):

$$h_{ij} = 1.3 + (\beta_1 + u_1) [1 - \exp(-\beta_2 d_{ij})]^{\beta_3 + u_2} \tag{15}$$

where u_1 and u_2 are random parameters. Parameter estimates were obtained using the fitting data set for the best models (GM, FE and ME models) and are given in Tab. 3.

According to the predictions, the marginal non-calibrated fixed-effects (FE) response model gave more unsuccessful outcomes, when random parameters of the mixed models were zero and prior height information was not available for model calibration for both datasets (Tab. 4).

Artificial neural network modeling approaches

ANN models used a three-layer architecture that included: (i) one input layer with three variables (diameter at breast height, d ; dominant height, H_o ; and dominant diameter, D_o). Sensitivity analysis was performed using the available variables listed in Tab. 1. The three variables given above had significant entry on the target variable higher than one; (ii) one output layer with one target variable (total tree stem height,

h) specified by the physical problem; and (iii) one hidden layer, comprising the proper number of hidden nodes specified by the network learning.

Using the LM algorithm for ANN model learning, the LMANN showing the best adaptation to the fitting data set was 3-2-1/0.5512, i.e., the specific model included a single input layer with the input nodes d , H_o and D_o , one hidden layer with two hidden nodes and an output layer with one node (h), with RMSE equal to 0.5512. Using the RP algorithm for ANN model learning, the RPANN showing the best adaptation to the fitting data set was the 3-2-1/0.5387 model, i.e., the specific model comprised one input layer with input variables d , H_o and D_o , a single output layer with output node h as in LMANN models building and, between these layers, a single hidden layer with two hidden nodes resulted from the trial-and-error process. The RMSE was 0.5387. For both ANN algorithms, the optimum number of hidden nodes was determined following examination of different numbers of nodes. The optimum values were selected as those leading to the best learning of the ANN models, with the smallest RMSE value (Fig. 3).

The $k=10$ cross validation method applied during the learning using the fitting data set for both ANN algorithms for training and testing in the construction phase, ensured the absence of overfitting and increased the ability of the ANNs to generalize. Both networks were trained for as many epochs as required for the models to

Tab. 4 - Fit statistics for the constructed basic *h-d* mixed-model, generalized *h-d* model and artificial neural network models (LMANN and RPANN) for (a) the fitting dataset, (b) the validation data sets, and (c) the calibrated adjusted fixed and mixed models, using 1, 2, and 3 sampled trees for calibration. (§): denote the best method for cedar plantations; (*): for p-value <0.05, the null hypothesis ($H_0: \mu_1 = \mu_2$) was rejected.

Dataset	Model	Type	RMSE	FI	MD/two tailed p-value*	MAD
(a) Fitting data set (630 trees in 21 plots)	LMANN: 3-2-1	ANN	0.5512	0.8241	-0.0407494	0.4306
	RPANN: 3-2-1	ANN	0.5387 §	0.8353	0.0074/0.729	0.4155 §
	Chapman-Richards (eqn. 1)	Fixed	0.9116	0.5912	-0.0019833	0.7088
	Chapman-Richards (eqn. 15)	Mixed	0.6316	0.8059	0.1905/0.000	0.4684
	Krumland & Wensel (eqn. 14)	Generalized	0.5677	0.8422	-0.7526316	0.4299
(b) Validation data set (270 trees in 9 plots)	LMANN: 3-2-1	ANN	0.6839	0.8	-0.1138	0.5673
	RPANN: 3-2-1	ANN	0.6671 §	0.8176	-0.0995 §	0.55 §
	Chapman-Richards (eqn. 1)	Fixed [F(0)]	1.1401	0.6017	0.2101	0.8943
	Krumland & Wensel (eqn. 14)	Generalized (GM)	0.721	0.7825	-0.2272	0.5985
(c) Validation data (270 trees in 9 plots) after using localization	Chapman-Richards (eqn. 3)	Fixed-Calibrated, FE(1)	0.8759	0.7649	0.3114	0.6417
	Chapman-Richards (eqn. 3)	Fixed-Calibrated, FE(2)	0.7427	0.831	-0.1683	0.5937
	Chapman-Richards (eqn. 3)	Fixed-Calibrated, FE(3)	0.7348	0.8346	-0.1778	0.5809
	Chapman-Richards (eqn. 15)	Mixed-Calibrated ME(1)	0.7573	0.8243	0.2129	0.5798
	Chapman-Richards (eqn. 15)	Mixed-Calibrated ME(2)	0.6799	0.8585	-0.1335 §	0.5377
	Chapman-Richards (eqn. 15)	Mixed-Calibrated ME(3)	0.6773 §	0.8594 §	-0.1523	0.5242 §

reach the minimum MSE value (Fig. 4). As shown in Fig. 4, the estimated errors for the training and the test data sets were not significantly different for either ANN model, which indicated the absence of overfitting. The LMANN was stopped after 200 training interactions (epochs) as no influential reduction of the MSE value occurred after 28 epochs, while the RPANN was trained for 500 interactions as there was no influential reduction in MSE after 167 epochs (Fig. 4).

Evaluation of the different modeling approaches

A synopsis of the evaluation statistics calculated with the fitting and validation data sets using the different modeling approaches is given in Tab. 4. The best outcomes among the GM models are reported in Tab. 4. Comparing the MD, MAD, RMSE, and FI values for the fitting data set, both the RPANN and the LMANN models gave

the best results (Tab. 4). Further, in order to enrich the conclusions derived from the results for the constructed models using the fitting data set, where the total height approximately normally distributed, the t-paired test was applied. The two-tailed p-values of the test were <0.05 for both the LMANN and the RPANN constructed models, suggesting that there is no mean difference between the observed and the estimated height values. The same conclusion can be derived for both the fixed-effects (FE) Chapman-Richards model (eqn. 1) and the generalized Krumland & Wensel (eqn. 14) model, while the mean error of mixed-effects model (eqn. 14) was found significant ($p < 0.05$) for the fitting data set.

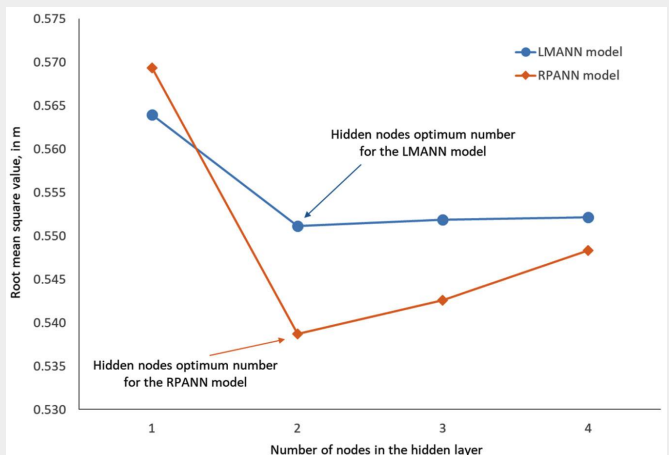
The accuracy of the estimated results from the ANN models was examined further by analyzing the correlation between the observed *h* values and their predictions from the models (Freedman et al. 2007). The Pearson's correlation coefficient (*r*)

was computed as the ratio between the covariance of the observed heights and the height predictions by the constructed models. The LMANN model's estimations gave $r = 0.9223$, while for the RPANN model gave $r = 0.9259$, indicating a high accordance between observed and predicted values. The variability of the estimation errors ($CV\% = (RMSE/h \cdot 100)$) in relation to the mean of the total tree height for the fitting data set was 9.00% and 8.80% of the mean height for the LMANN and the RPANN models, respectively. The reliability and the predictive capabilities of both ANN models was also examined based on the validation data set taken from 270 trees in 9 plots in the same forest areas. The LMANN model predicted an *r* value of 0.9318 with a $CV\%$ of 10.27%, while the *r* value for the RPANN model was 0.9336, with a $CV\%$ of 10.02%. According to the validation data set, the Chapman-Richards mixed-calibrated model using three trees for calibration showed the next best predictive capabilities, with $CV\%$ equals to 10.17% and $r = 0.9309$, while the non-calibrated generalized model of Krumland & Wensel (1988) gave similar results as compared to Chapman-Richards mixed-calibrated model, with $CV\%$ equals to 10.83% and $r = 0.9269$. Finally, the Chapman-Richards fixed-model produced the poorest results for the validation data set with $CV\%$ equals to 17.13 and $r = 0.8140$.

Discussion

Important ecosystem services including usable forest products and biodiversity conservation can be obtained through sustainably managed forests. The development of a reliable model that could estimate the total height of trees using easily measurable characteristics of trees, can

Fig. 3 - Optimum number of nodes in the hidden layer for the LMANN and RPANN models.



contribute efficiently to the above goal. Our approach attempts to compare, suggest and finally provide an alternative new methodology to the field of optimal forest management design, which would be able to give reliable predictions of tree heights, as this tree characteristic has great influence on the stand description, site quality determination, and tree and stand wood volume estimation.

Parametric modeling techniques assessment

The adjusted fixed-effects model developed for Taurus cedar stem heights in the present work had a much lower mean deviation (MD) for all calibration alternatives compared to the fixed-effects model. With all calibration parameters tested, the fit index increased at rates varying from 27% to 39%, with a mean increase of 35%. Similar findings were also reported in the literature (Lhotka 2012, VanderSchaaf 2013, Ozçelik et al. 2018). The most effective improvement in predictions, among the tested calibration alternatives, was obtained through the use of 3 trees for calibration in each sampling plot. Different sampling alternatives tested for both adjusted fixed-effects model and ME model, corresponding to number of trees measured in each plot. These tree heights were used to estimate the random parameters of the ME model, to calibrate heights and to calculate the assessment statistics for the validation data set (Tab. 4).

According to the ME model, lower mean deviations were obtained following calibration with 3 trees, when compared with the adjusted FE model. Comparing all calibration alternatives, ME model gave a decrease of 32% and 39% in the mean deviation and mean absolute difference, respectively. Previous studies (Trincado et al. 2007, Lhotka 2012, Huang et al. 2009, Gómez-García et al. 2014, VanderSchaaf 2013) also demonstrated that calibration can significantly improve height predictions. There were no large differences between the calibrations performed with 2 and 3 trees for cedar plantations (Tab. 4) through testing different calibration alternatives with the use of the mixed-effects model, while there were similar trends between different calibration alternatives for MAD and FI values. When the calibrated fixed effect model and the calibrated mixed effect models were compared, for all calibration alternatives, the ME model was more successful than the calibrated FE model in accurately estimating heights in cedar plantations.

The results obtained using the Krumland & Wensel (1988) model, compared with the outcomes of the FE model, suggest that the GM model gave more accurate results than the FE model. The generalized $h-d$ model gave an average reduction of 17% in MAD values, when compared to the FE model. The GM models exhibited better results than the adjusted FE models for all

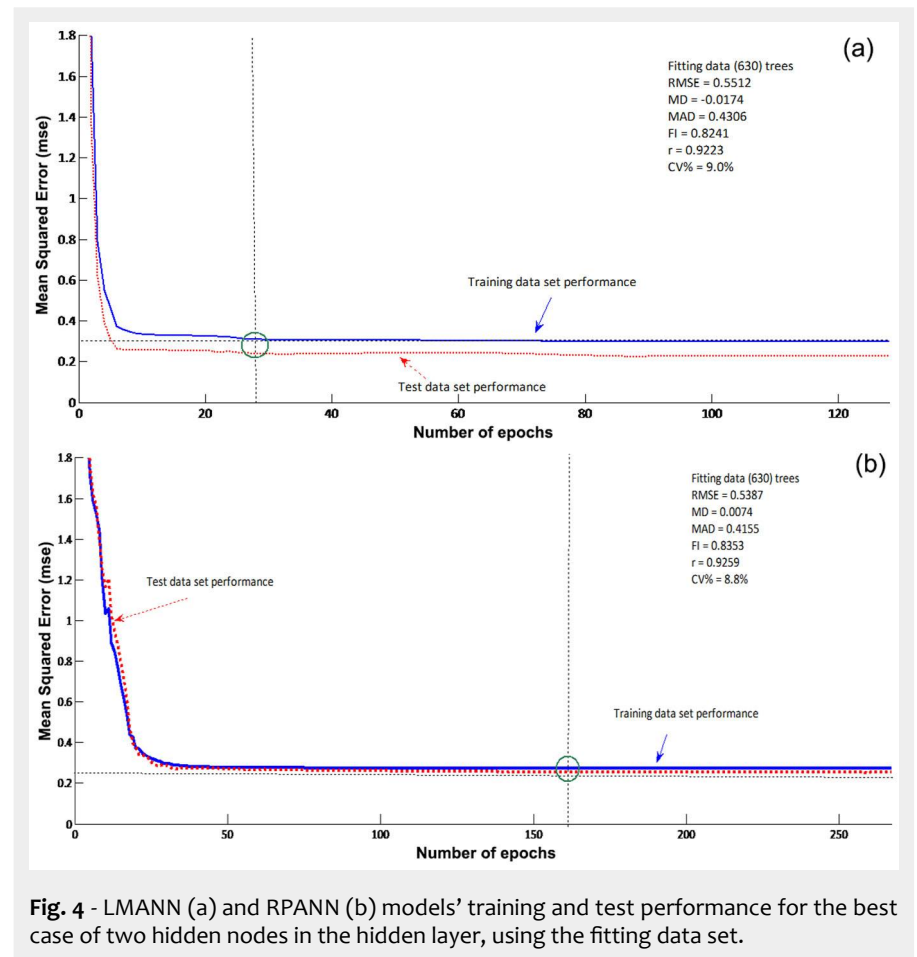


Fig. 4 - LMANN (a) and RPANN (b) models' training and test performance for the best case of two hidden nodes in the hidden layer, using the fitting data set.

evaluation statistics, except for MD values. The model of Krumland & Wensel (1988) includes H_0 and D_0 as fixed effects at plot level. H_0 is preferred to other parameters, as measurement of mean height across plots involves more sampling effort than H_0 (Huang et al. 2009). However, Bronisz & Mehtätalo (2020) did not suggest H_0 to be included in the model as a fixed effect predictor when a small sample of H_0 values based on a certain number of dominant height measurements is used in the model fitting process. In this case, the inclusion of a similar number of height values could improve model accuracy; otherwise, model outcomes may suffer from large and unpredictable measurement errors impacts on the estimates (Bronisz & Mehtätalo 2020). The calibrated basic mixed $h-d$ model produced more accurate tree height predictions than the generalized $h-d$ model for Turkish cedar plantations without the inclusion of other measurements as covariates. This is an important issue with practical implications in forest economics, such as the reduction of time, effort, and cost for collecting additional covariates. Similar conclusions have been reported by Huang et al. (2009).

The sample size used in model calibration had an impact on prediction accuracy, as height predictions improved depending on the number of trees used in the calibration process (Trincado et al. 2007, Huang et al. 2009, VanderSchaaf 2013, Ozçelik et al.

2018). The accuracy of prediction, however, decreased substantially beyond a certain number of trees. Therefore, the number of trees to be used for calibration should be chosen so as to strike a balance between the costs of inventory and the prediction accuracy. Calama & Montero (2004) suggested that calibration should be performed with additional trees, thus increasing the success of height predictions. Trincado et al. (2007) also argued that height prediction accuracy increases with increasing the numbers of trees used for calibration. According to these authors, the prediction accuracy reached to maximum using three calibration trees. Similar outcomes were also reported by Temesgen et al. (2008). Adame et al. (2008) tested different calibration alternatives and suggested that the use 2 or 3 trees was necessary and adequate for calibration. Huang et al. (2009) used between 1 to 9 trees for calibration, obtaining the most accurate results in height predictions when the calibration was performed with a single tree. The results of the present study suggested that there was a minimal improvement in prediction capability of the model when the number of calibration trees was increased above three.

Assessment of artificial neural network modeling techniques

We examined the performances of the LM and RP algorithms in tree height predic-

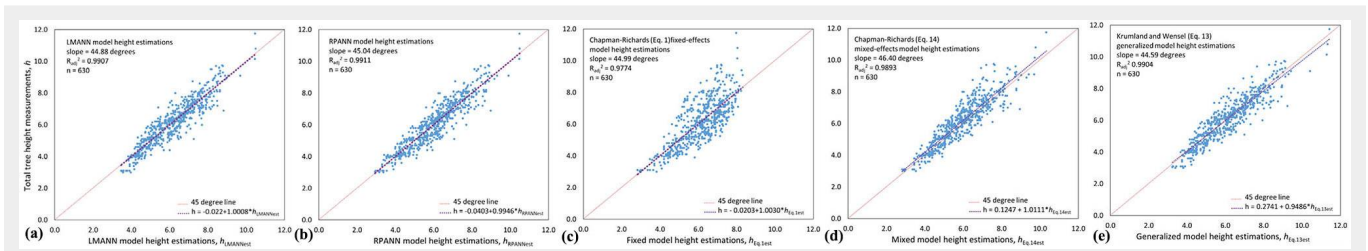


Fig. 5 - Forty-five-degree plots with the linear trendline of height estimations obtained by (a) LMANN, (b) RPANN, (c) fixed, (d) mixed, and (e) generalized models.

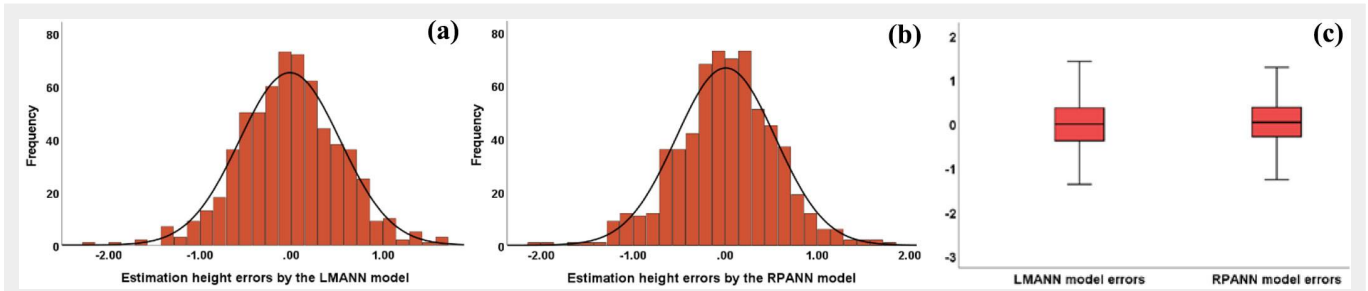


Fig. 6 - Errors histogram for height estimations obtained by the (a) LMANN, (b) RPANN models, and (c) errors box plots for both ANN models using the training data set.

tion, with the disadvantages and restrictions of the standard back-propagation algorithm (Fausett 1994, Patterson 1996, Wilamowski & Yu 2010). The low dispersion of predicted values around the 1:1 line throughout the observed range of values (Fig. 5a, Fig. 5b) indicates that both ANN

models had similar adjustments to the data used in model construction. Moreover, both ANN models showed better adaptation to the fitting data set than the fixed (Fig. 5c), mixed (Fig. 5d) and generalized (Fig. 5e) models. Furthermore, the statistical evaluation of the null hypothesis (H_0 : in-

tercept = 0 and H_0 :slope = 1) for the linear relationship $h_{obs} = intercept + slope \cdot h_{pred}$ was performed for all models (Fig. 5). The null hypotheses were not rejected at $\alpha=0.05$ for either ANN models, indicating no significant differences between the 45 degree line and the linear trendline derived by the observed and the estimated height values (Fig. 5a, Fig. 5b). Finally, the RPANN model (Fig. 5b) appeared to provide slightly more accurate fitting to the observed data as compared to the LMANN model (Fig. 5a). The distribution of prediction errors from ANN models (Fig. 6a, Fig. 6b) revealed a frequency peak at zero and a rapid decline of the frequency at larger error values, thus corroborating that both ANN models are well-trained networks. Nonetheless, it is worth noting that the most accurate predictions were obtained by the best RPANN model, with a RMSE value of 2.25% lower than the error obtained from the best LMANN model.

The prediction errors histograms and the relative errors box plots (Fig. 7, Fig. 8), along with the validation statistics presented in Tab. 4, clearly show that the RPANN model provided the most accurate and reliable prediction results as compared to the other models tested. The variability of the prediction errors (CV%) of the RPANN model was approximately 2.43%, which was lower than that of the LMANN model, and 1.47% lower than the variability derived from the Chapman-Richards mixed effect calibrated model [ME(3)]. The distribution analysis of prediction errors for all tested models suggested that the performances of the RPANN model are superior to those of all the other un-calibrated and calibrated models. Therefore, we conclude that the RPANN model provides the best performances in the estimation and predic-

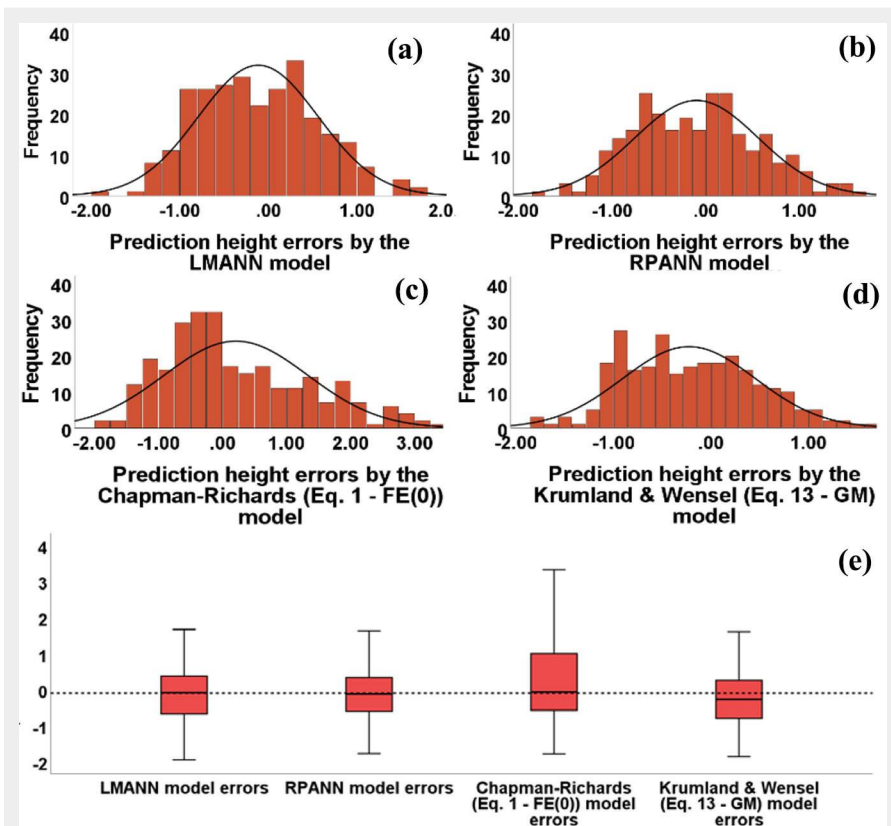
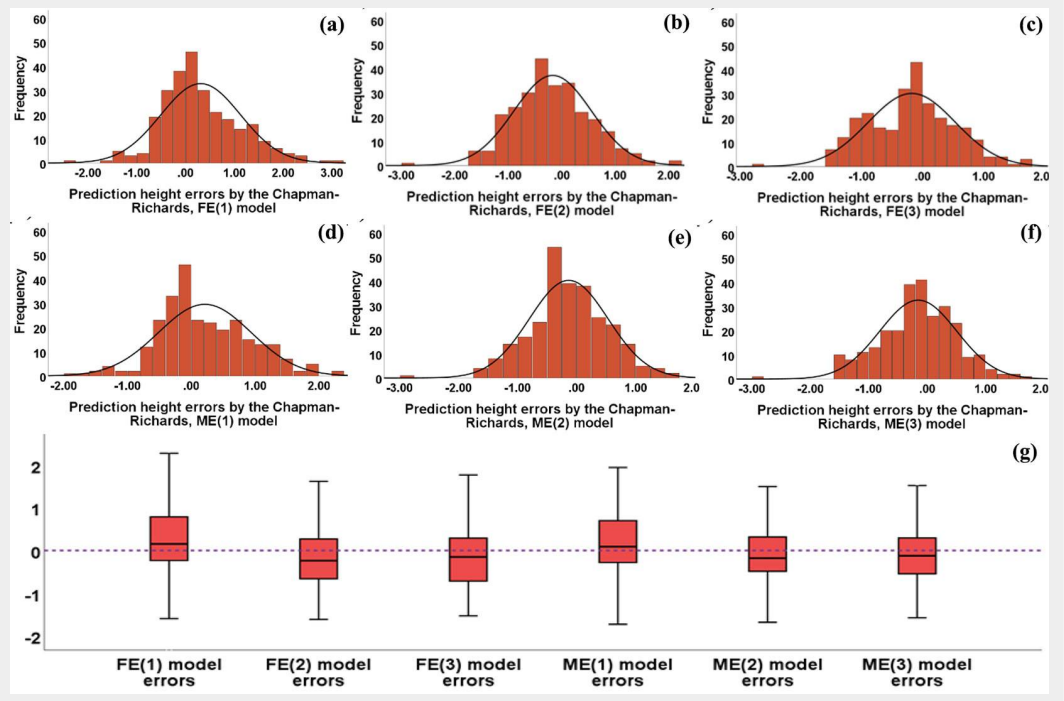


Fig. 7 - Errors histograms for height predictions obtained by (a) LMANN, (b) RPANN, (c) Chapman-Richards [eqn. 1 - FE(0)], (d) Krumland & Wensel (eqn. 13 - GM) models, and (e) errors' box plots for all models using the validation data set..

Fig. 8 - Errors histograms for height predictions obtained by the Chapman-Richards (a) FE(1), (b) FE(2), (c) FE(3), (d) ME(1), (e) ME(2), and (f) ME(3) models, and (g) errors box plots for all models using the validation data set.



tion of tree height, as compared with fixed, generalized, fixed-calibrated and mixed calibrated models.

Conclusions

In this study, equations for height prediction were developed for *Cedrus libani* plantations, which involved the evaluation of five alternative modeling methods: (i) the Generalized (GM); (ii) the non-calibrated fixed-effects [FE(0)]; (iii) the calibrated fixed-effect, using 1 to 3 calibration trees per plot [FE(1-3)]; (iv) the calibrated mixed-effects using 1 to 3 calibration trees per plot [ME(1-3)]; and (v) the ANNs using two different learning algorithms, the RP and the LM. As compared to the FE, all the constructed models (generalized, adjusted FE, ME, LMANN and BPANN) gave better performance, providing accurate and reliable tree height predictions. As expected, the predictive ability of the adjusted and ME models improved with increasing the calibration sample size. However, increasing the number of trees used for calibration had a minimal effect on the accuracy of the models.

The LMANN and RPANN models constructed in the present work showed reliable abilities to generalize from the input data and gave accurate predictions of the total tree height variable. Both alternatives exhibited similar performance in the construction phases, in addition to the ability to generalize. Based on the cumulative results from all approaches, the resilient propagation algorithm was more effective in predicting total stem height of *C. libani* growing in Turkey. The effectiveness and efficiency of ANN modeling in general and, specifically of the resilient propagation algorithm, along with the fit to the current dataset, the potential of this modeling ap-

proach to forest management, are clear. This direction is being followed in future research.

Author contribution

Y.K.: conceptualization, editing, results evaluation; M.D.: methodology, investigation, formal analysis, writing, review, editing; R.Ö.: methodology, investigation, formal analysis, review, writing, editing; Z.S.: data collection, data curation.

Acknowledgements

We thank the Turkish General Directorate of Forestry for its contribution to field work. We also thank Prof. Steve Woodward from University of Aberdeen for his valuable comments and suggestions for revising the English grammar of the text.

Funding

This work was financially supported by the Scientific Research Project Coordination Unit of the Suleyman Demirel University, Isparta, Turkey (project no. 2262-YL-10).

References

- Adame P, Del Río M, Canellas I (2008). A mixed nonlinear height-diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecology and Management* 256: 88-98. - doi: [10.1016/j.foreco.2008.04.006](https://doi.org/10.1016/j.foreco.2008.04.006)
- Adamec Z, Drápela K (2016). Comparison of parametric and nonparametric methods for modeling height-diameter relationships. *iForest - Biogeosciences and Forestry* 10 (1): 1-8. - doi: [10.3832/ifor1928-009](https://doi.org/10.3832/ifor1928-009)
- Adeloye A, De Munari A (2006). Artificial neural network based generalized storage-yield-reliability models using the Levenberg-Marquardt algorithm. *Journal of Hydrology* 326 (1-4): 215-230. - doi: [10.1016/j.jhydrol.2005.10.033](https://doi.org/10.1016/j.jhydrol.2005.10.033)

- Beale MH, Hagan M, Demuth H (2017). *Neural network toolbox: user's guide, v. R2017a*. Mathworks Inc, Natick, MA, USA, pp. 446.
- Boydak M (2003). Regeneration of Lebanon cedar (*Cedrus libani* A. Rich.) on karstic lands in Turkey. *Forest Ecology and Management* 178: 231-243. - doi: [10.1016/S0378-1127\(02\)00539-X](https://doi.org/10.1016/S0378-1127(02)00539-X)
- Bronisz K, Mehtätalo L (2020). Mixed-effects generalized height-diameter model for young silver birch stands on post-agricultural lands. *Forest Ecology and Management* 460, 117901. - doi: [10.1016/j.foreco.2020.117901](https://doi.org/10.1016/j.foreco.2020.117901)
- Calama R, Montero G (2004). Interregional nonlinear height diameter model with random coefficients for stone pine in Spain. *Canadian Journal of Forest Research* 34: 150-163. - doi: [10.1139/x03-199](https://doi.org/10.1139/x03-199)
- Cariboni J, Gatelli D, Liska R, Saltelli A (2007). The role of sensitivity analysis in ecological modeling. *Ecological Modelling* 203: 167-182. - doi: [10.1016/j.ecolmodel.2005.10.045](https://doi.org/10.1016/j.ecolmodel.2005.10.045)
- Catal Y (2012). Height-diameter model for black locust, Anatolian black pine and Taurus cedar tree species in Lakes Region. *Turkish Journal of Forestry* 13: 92-96. - doi: [10.18182/tjf.62652](https://doi.org/10.18182/tjf.62652)
- Catal Y, Carus S (2018). A height-diameter model for brutian pine (*Pinus Brutia* Ten.) plantations in southwestern Turkey. *Applied Ecology and Environmental Research* 16 (2): 1445-1459. - doi: [10.1566/aeer](https://doi.org/10.1566/aeer)
- Colbert KC, Larsen DR, Lootens JR (2002). Height-diameter equations for thirteen mid-western bottomland hardwood species. *Northern Journal of Applied Forestry* 19 (4): 171-176. - doi: [10.1093/njaf/19.4.171](https://doi.org/10.1093/njaf/19.4.171)
- Corral-Rivas S, Alvarez-González JG, Crecente-Campo F, Corral-Rivas JJ (2014). Local and generalized height-diameter models with random parameters for mixed, uneven-aged forests in Northwestern Durango, Mexico. *Forest Ecosystems* 1: 1-9. - doi: [10.1186/2197-5620-1-6](https://doi.org/10.1186/2197-5620-1-6)
- Cosenza DN, Soares AAV, De Alcântara AEM, Da Silva AAL, Rode R, Soares VP, Leite HG (2017).

- Site classification for eucalypt stands using artificial neural network based on environmental and management features. *Cerne* 23: 310-320. - doi: [10.1590/01047760201723032352](https://doi.org/10.1590/01047760201723032352)
- Crecente-Campo F, Tomé M, Soares P, Diéguez-Aranda U (2010). A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *Forest Ecology and Management* 259: 943-952. - doi: [10.1016/j.foreco.2009.11.036](https://doi.org/10.1016/j.foreco.2009.11.036)
- Curtis RO (1967). Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *Forest Science* 13: 365-375. - doi: [10.1093/forestscience/13.4.365](https://doi.org/10.1093/forestscience/13.4.365)
- Diamantopoulou MJ, Ozçelik R (2012). Evaluation of different modeling approaches for total tree height estimation in Mediterranean Region of Turkey. *Forest Systems* 21: 383-397. - doi: [10.5424/fs/2012213-02338](https://doi.org/10.5424/fs/2012213-02338)
- Diamantopoulou MJ, Ozçelik R, Crecente-Campo F, Eler U (2015). Estimation of Weibull function parameters for modeling tree diameter distribution using least squares and artificial neural networks methods. *Biosystems Engineering* 133: 33-45. - doi: [10.1016/j.biosystemseng.2015.02.013](https://doi.org/10.1016/j.biosystemseng.2015.02.013)
- Ercanlı I, Günlü A, Senyurt M, Keles S (2018). Artificial neural network models predicting the leaf area index: a case study in pure even-aged Crimean pine forests from Turkey. *Forest Ecosystems* 5: 1-12. - doi: [10.1186/s40663-018-0149-8](https://doi.org/10.1186/s40663-018-0149-8)
- Ercanlı I (2020). Innovative deep learning artificial intelligence applications for predicting relationships between individual tree height and diameter at breast height. *Forest Ecosystems* 7: 1-18. - doi: [10.1186/s40663-020-00226-3](https://doi.org/10.1186/s40663-020-00226-3)
- Fausett L (1994). *Fundamentals of neural networks architectures*. Algorithms and applications. Prentice Hall, Englewood Cliffs, NJ, USA, pp. 461.
- Florescu C, Igel C (2018). Resilient backpropagation (RPROP) for batch-learning in tensorflow. In: *Proceedings of the "6th International Conference on Learning Representations - ICLR 2018"*. OpenReview.net 2018, web site. [online] URL: <http://dblp.uni-trier.de/db/conf/iclr/iclr2018w.html#Florescu18>
- Freedman D, Pisani R, Purves R (2007). *Statistics - International student edition (4th edn)* (Pisani Purves R eds). WW Norton and Company, New York, USA, pp. 720.
- GDF (2015). *Forest resources*. General Directorate of Forests - GDF, Ankara, Turkey, pp. 162.
- Gómez-García E, Diéguez-Aranda U, Castedo-Dorado F, Crecente-Campo F (2014). A comparison of model forms for the development of height-diameter relationships in even-aged stands. *Forest Science* 60: 560-568. - doi: [10.5849/forsci.12-099](https://doi.org/10.5849/forsci.12-099)
- Gómez-García E, Fonseca TF, Crecente-Campo F, Almeida LR, Diéguez-Aranda U, Huang S, Marques CP (2015). Height-diameter models for maritime pine in Portugal: a comparison of basic, generalized and mixed-effects models. *iForest - Biogeosciences and Forestry* 9 (1): 72-78. - doi: [10.3832/ifo1520-008](https://doi.org/10.3832/ifo1520-008)
- Hagan MT, Menhaj MB (1994). Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks* 5 (6): 989-993. - doi: [10.1109/72.329697](https://doi.org/10.1109/72.329697)
- Hall DB, Clutter M (2004). Multivariate multilevel nonlinear mixed effects models for timber yield predictions. *Biometrics* 60: 16-24. - doi: [10.1111/j.0006-341X.2004.00163.x](https://doi.org/10.1111/j.0006-341X.2004.00163.x)
- Huang S, Titus SJ, Wiens DP (1992). Comparison of nonlinear height-diameter functions for major Alberta tree species. *Canadian Journal of Forest Research* 22 (9): 1297-1304. - doi: [10.1139/x92-172](https://doi.org/10.1139/x92-172)
- Huang S, Wiens DP, Yang Y, Meng SX, Vander-schaaf CL (2009). Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods. *Forest Ecology and Management* 258: 1235-1247. - doi: [10.1016/j.foreco.2009.06.017](https://doi.org/10.1016/j.foreco.2009.06.017)
- Hursch R (1991). Validation samples. *Biometrics* 47: 1193-1194.
- Kearsley E, Moonen PC, Hufkens K, Doetterl S, Lisingo J, Bosela FB, Boeckx P, Beeckman H, Verbeeck H (2017). Model performance of tree height-diameter relationships in the central Congo Basin. *Annals of Forest Science* 74 (1): 1-13. - doi: [10.1007/s13595-016-0611-0](https://doi.org/10.1007/s13595-016-0611-0)
- Konukçu M (2001). *Forests and forestry*. State Planning Organization publications, Ankara, Turkey, pp. 238. [in Turkish]
- Krumland BE, Wensel LC (1988). A generalized height-diameter equation for coastal California species. *Western Journal of Applied Forestry* 3: 113-115. - doi: [10.1093/wjaf/3.4.113](https://doi.org/10.1093/wjaf/3.4.113)
- Leahy K (1994). The overfitting problem in perspective. *AI Expert: The Magazine of Artificial Intelligence* 9 (10): 35-36. [ISSN 0888-3785]
- Lei X, Peng C, Wang H, Zhou X (2009). Individual height-diameter models for young black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*) plantations in New Brunswick, Canada. *Forestry Chronicle* 85: 43-56. - doi: [10.5558/tfc85043-1](https://doi.org/10.5558/tfc85043-1)
- Leite HG, Da Silva MLM, Binoti DHB, Fardin L, Takizawa FH (2011). Estimation of inside-bark diameter and heartwood diameter for *Tectona grandis* Linn. trees using artificial neural networks. *European Journal of Forest Research* 130 (2): 263-269. - doi: [10.1007/s10342-010-0427-7](https://doi.org/10.1007/s10342-010-0427-7)
- Levenberg K (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* 2: 164-168. - doi: [10.1090/qam/10666](https://doi.org/10.1090/qam/10666)
- Lhotka JM (2012). Height-diameter relationships in Sweetgum (*Liquidambar styraciflua*)-dominated stands. *Southern Journal of Applied Forestry* 36 (2): 98-106. - doi: [10.5849/sjaf.10-039](https://doi.org/10.5849/sjaf.10-039)
- Li YX, Jiang LC (2010). Application of ANN algorithm in tree height modeling. *Applied Mechanics and Materials* 20 (23): 756-761. - doi: [10.4028/www.scientific.net/AMM.20-23.756](https://doi.org/10.4028/www.scientific.net/AMM.20-23.756)
- Liu C, Zhang L, Davis CJ, Solomon DS, Brann TB, Caldwell LE (2003). Comparison of neural networks and statistical methods in classification of ecological habitats using FIA data. *Forest Science* 49 (4): 619-631. - doi: [10.1093/forestscience/49.4.619](https://doi.org/10.1093/forestscience/49.4.619)
- Maier HR, Dandy GC (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental Modeling and Software* 15: 101-124. - doi: [10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- Marquardt DW (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics* 11 (2): 431-441. - [online] URL: <http://www.jstor.org/stable/2098941>
- May RJ, Maier HR, Dandy GC (2010). Data splitting for artificial neural networks using SOM-based stratified sampling. *Neural Networks* 23 (2): 283-294. - doi: [10.1016/j.neunet.2009.11.009](https://doi.org/10.1016/j.neunet.2009.11.009)
- Mehtätalo L, De-Miguel S, Gregoire TG (2015). Modeling height-diameter curves for prediction. *Canadian Journal of Forest Research* 45: 826-837. - doi: [10.1139/cjfr-2015-0054](https://doi.org/10.1139/cjfr-2015-0054)
- Meng SX, Huang S (2009). Improved calibration of nonlinear mixed-effects models demonstrated on a height growth function. *Forest Science* 55 (3): 238-248. - doi: [10.1093/forestscience/55.3.238](https://doi.org/10.1093/forestscience/55.3.238)
- MGM (2013). *Official statistics*. Statistical data of Isparta Province and Districts. Turkish State Meteorological Service - MGM, Ankara, Turkey, web site. [In Turkish] [online] URL: <http://mgm.gov.tr/eng/Yearly-Climate>
- Monteiro Da Silva EJ, Maia RD, Cabacinha CD (2018). Bee-inspired RBF network for volume estimation of individual trees. *Computers and Electronics in Agriculture* 152: 401-408. - doi: [10.1016/j.compag.2018.07.036](https://doi.org/10.1016/j.compag.2018.07.036)
- Newton P, Amponsah I (2007). Comparative evaluation of five height-diameter models developed for black spruce and jack pine stand-types in terms of goodness-of-fit, lack-of-fit and predictive ability. *Forest Ecology and Management* 247: 149-166. - doi: [10.1016/j.foreco.2007.04.029](https://doi.org/10.1016/j.foreco.2007.04.029)
- Olson DL, Delen D (2008). *Advanced data mining techniques*. Springer-Verlag, Berlin, Heidelberg, Germany, pp. 180. - doi: [10.1007/978-3-540-76917-0](https://doi.org/10.1007/978-3-540-76917-0)
- Ozçelik R, Diamantopoulou MJ, Crecente-Campo F, Eler U (2013). Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. *Forest Ecology and Management* 306: 52-60. - doi: [10.1016/j.foreco.2013.06.009](https://doi.org/10.1016/j.foreco.2013.06.009)
- Ozçelik R, Diamantopoulou MJ, Brooks JR (2014). The use of tree crown variables in over-bark diameter and volume prediction models. *iForest - Biogeosciences and Forestry* 7 (3): 132-139. - doi: [10.3832/ifo1520-008](https://doi.org/10.3832/ifo1520-008)
- Ozçelik R, Diamantopoulou MJ, Eker M, Gürlevik N (2017). Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. *Forest Science* 63 (3): 291-302. - doi: [10.5849/FS-16-006](https://doi.org/10.5849/FS-16-006)
- Ozçelik R, Cao QV, Trincado G, Göçer N (2018). Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. *Forest Ecology and Management* 419: 240-248. - doi: [10.1016/j.foreco.2018.03.051](https://doi.org/10.1016/j.foreco.2018.03.051)
- Ozçelik R, Diamantopoulou MJ, Trincado G (2019). Evaluation of potential modeling approaches for Scots pine stem diameter prediction in north-eastern Turkey. *Computers and Electronics in Agriculture* 162: 773-782. - doi: [10.1016/j.compag.2019.05.033](https://doi.org/10.1016/j.compag.2019.05.033)
- Patterson DW (1996). *Artificial neural networks, theory and applications*. Prentice Hall, Singapore, pp. 477.
- Peng C, Zhang L, Liu J (2001). Developing and

- validating nonlinear height-diameter models for major tree species of Ontario's boreal forests. *Northern Journal of Applied Forestry* 18 (3): 87-94. - doi: [10.1093/njaf/18.3.87](https://doi.org/10.1093/njaf/18.3.87)
- Raptis DI, Kazana V, Kazaklis A, Stamatiou C (2021). Mixed-effects height-diameter models for black pine (*Pinus nigra* Arn.) forest management. *Trees* 35: 1167-1183. - doi: [10.1007/s00468-021-02106-x](https://doi.org/10.1007/s00468-021-02106-x)
- Reis LP, De Souza AL, Mazzei L, Dos Reis PCM, Leite HG, Soares CPB, Torres CMME, Da Silva LF, Ruschel AR (2016). Prognosis on the diameter of individual trees on the eastern region of the amazon using artificial neural networks. *Forest Ecology and Management* 382: 161-167. - doi: [10.1016/j.foreco.2016.10.022](https://doi.org/10.1016/j.foreco.2016.10.022)
- Riedmiller M, Braun H (1993). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. In: "IEEE International Conference on Neural Networks", vol. 1, pp. 586-591. - doi: [10.1109/ICNN.1993.298623](https://doi.org/10.1109/ICNN.1993.298623)
- Robinson AP, Wykoff WR (2004). Imputing missing height measures using a mixed-effects modeling strategy. *Canadian Journal of Forest Research* 34 (12): 2492-2500. - doi: [10.1139/X04-137](https://doi.org/10.1139/X04-137)
- Saatçioğlu A (1979). *Silviculture II*. Istanbul University Faculty of Forestry publications, Istanbul, Turkey, pp. 562. [In Turkish]
- Saputra W, Tulus Zarlis M, Sembring RW, Hartama D (2017). Analysis resilient algorithm on artificial neural network backpropagation. *Journal of Physics: Conference Series* 930: 012035. - doi: [10.1088/1742-6596/930/1/012035](https://doi.org/10.1088/1742-6596/930/1/012035)
- SAS Institute (2010). *SAS/OR® 9.22 user's guide: mathematical programming*. SAS Institute Inc., Cary, NC, USA, pp. 1264.
- Scanlan D, Mulvaney D (2013). Graphics processor unit hardware acceleration of Levenberg-Marquardt artificial neural network training. *Research Inventory: International Journal of Engineering and Science* 2 (7): 7. [online] URL: <http://hdl.handle.net/2134/13092>
- Schmoldt DL, Li P, Abbott AL (1997). Machine vision using artificial neural networks with local 3D neighborhoods. *Computers and Electronics in Agriculture* 16 (3): 255-271. - doi: [10.1016/S0168-1699\(97\)00002-1](https://doi.org/10.1016/S0168-1699(97)00002-1)
- Sharma M, Parton J (2007). Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. *Forest Ecology and Management* 249 (3): 187-198. - doi: [10.1016/j.foreco.2007.05.006](https://doi.org/10.1016/j.foreco.2007.05.006)
- Soares FAA, Flores EL, Cabacinha CD, Carrijo GA, Veiga ACP (2013). Recursive diameter prediction for calculating merchantable volume of eucalyptus clones using Multilayer Perceptron. *Neural Computing and Applications* 22 (7): 1407-1418. - doi: [10.1007/s00521-012-0823-7](https://doi.org/10.1007/s00521-012-0823-7)
- Soares P, Tomé M (2002). Height-diameter equation for first rotation eucalypt plantations in Portugal. *Forest Ecology and Management* 166: 99-109. - doi: [10.1016/S0378-1127\(01\)00674-0](https://doi.org/10.1016/S0378-1127(01)00674-0)
- Sun Y, Ao Z, Jia W, Chen Y, Xu K (2021). A geographically weighted deep neural network model for research on the spatial distribution of the down dead wood volume in Liangshui National Nature Reserve (China). *iForest - Biogeosciences and Forestry* 14 (4): 353-361. - doi: [10.3832/IFOR3705-014](https://doi.org/10.3832/IFOR3705-014)
- Temesgen H, Monleon V, Hann D (2008). Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Canadian Journal of Forest Research* 38 (3): 553-565. - doi: [10.1139/X07-104](https://doi.org/10.1139/X07-104)
- Thanh TN, Tien TD, Shen HL (2019). Height-diameter relationship for *Pinus koraiensis* in Mengjiagang Forest Farm of Northeast China using nonlinear regressions and artificial neural network models. *Journal of Forest Science* 65 (4): 134-143. - doi: [10.17221/5/2019-JFS](https://doi.org/10.17221/5/2019-JFS)
- Trincado G, VanderSchaaf CL, Burkhart HE (2007). Regional mixed-effects height-diameter models for loblolly pine (*Pinus taeda* L.) plantations. *European Journal of Forest Research* 126 (2): 253-262. - doi: [10.1007/s10342-006-0141-7](https://doi.org/10.1007/s10342-006-0141-7)
- Vacchiano G, Berretti R, Motta R, Mondino Borogno E (2018). Assessing the availability of forest biomass for bioenergy by publicly available satellite imagery. *iForest - Biogeosciences and Forestry* 11 (4): 459-468. - doi: [10.3832/IFOR2655-011](https://doi.org/10.3832/IFOR2655-011)
- VanderSchaaf CL (2013). Mixed-effects height-diameter models for commercially and ecologically important hardwoods in Minnesota. *Northern Journal of Applied Forestry* 30 (1): 37-42. - doi: [10.5849/njaf.12-018](https://doi.org/10.5849/njaf.12-018)
- Vargas-Larreta B, Castedo-Dorado F, Alvarez-González JG, Barrio-Anta M, Cruz-Cobos F (2009). A generalized height-diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico). *Forestry* 82 (4): 445-462. - doi: [10.1093/forestry/cppo16](https://doi.org/10.1093/forestry/cppo16)
- Vieira GC, De Mendonça R, Da Silva GF, Zanetti SS, Da Silva MM, Dos Santos AR (2018). Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of the Total Environment* 619: 1473-1481. - doi: [10.1016/j.scitotenv.2017.11.138](https://doi.org/10.1016/j.scitotenv.2017.11.138)
- Wilamowski BM, Yu H (2010). Improved computation for Levenberg-Marquardt training. *IEEE Transactions on Neural Networks* 21 (6): 930-937. - doi: [10.1109/TNN.2010.2045657](https://doi.org/10.1109/TNN.2010.2045657)
- Wu D (2014). Estimation of forest volume based on LM-BP neural network model. *Computer Modelling and New Technologies* 18 (4): 131-137. [online] URL: http://www.cmnt.lv/upload-files/ns_51art21_CMNT1804-38_Wu.pdf
- Wu D, Ji Y (2015). Dynamic estimation of forest volume based on multi-source data and neural network model. *Journal of Agricultural Science* 7 (3): 18-31. - doi: [10.5539/jas.v7n3p18](https://doi.org/10.5539/jas.v7n3p18)
- Wu C, Chen Y, Peng C, Li Z, Hong X (2019). Modeling and estimating aboveground biomass of *Dacrydium pierrei* in China using machine learning with climate change. *Journal of Environmental Management* 234: 167-179. - doi: [10.1016/j.jenvman.2018.12.090](https://doi.org/10.1016/j.jenvman.2018.12.090)
- Zhang X, Duan A, Zhang J, Xiang C (2014). Estimating tree height-diameter models with the Bayesian method. *The Scientific World Journal* 2014, 683691. - doi: [10.1155/2014/683691](https://doi.org/10.1155/2014/683691)
- Zang H, Lei X, Zeng W (2016). Height-diameter equations for larch plantations in northern and northeastern China: a comparison of the mixed-effects, quantile regression and generalized additive models. *Forestry* 89 (4): 434-445. - doi: [10.1093/forestry/cpw022](https://doi.org/10.1093/forestry/cpw022)
- Zhou R, Wu D, Fang L, Xu A, Lou X (2018). A Levenberg-Marquardt backpropagation neural network for predicting forest growing stock based on the least-squares equation fitting parameters. *Forests* 9 (12): 757. - doi: [10.3390/f9120757](https://doi.org/10.3390/f9120757)