Investigation of parametric and artificial neural network modeling approaches for total tree height prediction in cedar plantations

Yasin Karatepe

Isparta University of Applied Sciences

Maria J. Diamantopoulou

Aristotle University of Thessaloniki

Ramazan Özçelik (✉ ramazanozcelik@isparta.edu.tr)

Isparta University of Applied Sciences <https://orcid.org/0000-0003-2132-2589>

Zerrin Sürücü

Ministry of Agriculture Forestry, IV. Regional Directorate

Research

Keywords: Height model prediction, mixed-effects models, Levenberg-Marquardt algorithm, resilient propagation

Posted Date: October 27th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-96662/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Investigation of parametric and artificial neural network modeling approaches for total tree height prediction in cedar plantations

by

Yasin Karatepe¹, Maria J. Diamantopoulou², Ramazan Özçelik^{1*}, and Zerrin Sürücü³

¹Faculty of Forestry, Isparta University of Applied Science, East Campus, 32260, Isparta, Turkey.

yasinkaratepe@isparta.edu.tr; ramazanozcelik@isparta.edu.tr

^{1*}Corresponding author.

²Faculty of Forestry and Natural Environment, Aristotle University of Thessaloniki, GR-54124

Thessaloniki, Greece

mardi.diamant@gmail.com

³Ministry of Agriculture and Forestry, VI. Regional Directorate, 27002, Burdur, Turkey.

zerrin.surucu@tarimorman.gov.tr

Abstract:

Background: Height-diameter relationships are of critical importance in tree and stand volume estimation. Stand description, site quality determination and proper forest management decisions originate from reliable stem height predictions.

Methods: In the context of this work, the prediction ability of the developed height-diameter models was investigated for cedar (*Cedrus libani* A. Rich.) plantations in Western Mediterranean Region of Turkey. Towards this direction, parametric modeling methods such as fixed-effects, generalized models, and mixed-effects were evaluated. Furthermore, in an effort to come up with the construction of more reliable stem-height prediction models, artificial neural networks were employed using two different modeling algorithms: the Levenberg-Marquardt and the resilient back-propagation.

Results: Taking into account the prediction behavior of each respective modelling strategy while using a new validation data set, the mixed-effects model with calibration using 3 trees for each plot seems to be a reliable alternative to the rest standard modelling approaches given the evaluation statistics regarding the predictions of tree heights.

Conclusion: Finally, as for providing the most reliable results as compared to the remaining, the resilient propagation algorithm showed its capability of providing more accurate predictions of the tree stem height and thus it can be a reliable alternative to pre-existing modelling methodologies.

Keywords: Height model prediction; mixed-effects models; Levenberg-Marquardt algorithm; resilient propagation

Background

Cedar (*Cedrus libani* A. Rich.) is an important tree species both ecologically and economically and also it is presently found primarily in the Taurus Mountains of Turkey (Boydak 2004). The most recent inventory data show that cedar covers an area of around 463,521 hectares in Turkey, of which a total growing stock is 27.4 million m³ (GDF 2015). It is not only an important raw material resource of forest products industry, but also fulfils critical ecological tasks such as preserving soil and water resources, mitigation of and adaptation to the adverse impacts of climate change, and provides a vital for maintaining biological diversity in Turkey. It has moderate soil requirements and is tolerant to temperature and drought in summer and to the cold temperatures in winter (Saatçioğlu 1979); therefore, it has a big afforestation potential especially in areas with semi-arid climate. Given that 35% of Turkey's territory has semi-arid climate, cedar is a critical tree species that can be used for afforestation. As of 2000, an area of around 110.000 ha was afforested with cedar (Konukçu 2001).

Globally forest areas continue to decline in the last 30 years while plantations have increased in the same timeframe and they now comprise 7% of the total global forest area. These plantations are of great importance in economic and ecological terms. Thus, it is essential to have robust and reliable information about the growth and development characteristics of these species for sustainable management of plantations and to measure the success of the afforestation efforts. Diameter at breast height (d , 1.3 m above ground level) and total stem height (h) are the two basic parameters for the inventory, planning and management of plantations. These parameters are used as the fundamental variables in several forest practices such as prediction of standing yield, growth projection, carbon accounting, and identification of stand structural diversity, determination of various damages in forest or stands and to predict missing tree heights (Curtis 1967; Peng et al. 2001; Newton and Amponsah 2007; Adame et al. 2008).

Although, there are published papers dealing with the tree height estimation and prediction models using mainly standard modeling techniques (Colbert et al. 2002; Newton and Amponsah 2007; Zhang et al. 2014; Kearsley et al. 2017), there is still space for further research about topic when evaluating more recent modeling methods such as artificial intelligence. Existing research on height-diameter ($h-d$) models for cedar plantations in Turkey was conducted by Çatal (2012). However, the aforementioned research was based on limited number of traditional $h-d$ models, without any further investigation of the efficiency of other modeling approaches in the same topic.

In forest management applications or field inventories, d in trees can be measured rapidly, easily, and precisely while h is more time consuming, challenging and costly to measure tree height (Sharma and Parton 2007; Meng and Huang 2009; Corral-Rivas et al. 2014; 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 situation, 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. Finally, these created models are used to predict the height of any tree in the stand with unknown height (Huang et al. 1992; Robinson and Wykoff 2004; Lei et al. 2009).

Generally, $h-d$ models can be defined as simple and generalized. Simple models express height as a function of tree diameter at breast height (Bronisz and Mehtätalo 2020). This modeling approach represents the best equation that can be obtained for specific stand and point in time. The main problem with this approach is the large sampling effort required (Gómez-García et al. 2014). On the other hand, as stated by Lhotka (2012), research evaluating the correlation between tree diameter and total height has found that this relationship varies among species and can be influenced by stand-specific variables such as stand structure and site productivity. Models that include measures of stand structure, site productivity and tree social position are commonly referred to as generalized $h-d$ models. In several studies, the inclusion

of additional stand-level variables improved tree height estimates (López Sánchez et al. 2003; Sharma and Zhang 2004; Castedo-Dorado et al. 2006; Vargas-Larreta et al. 2009). For example, Newton and Amponsah (2007) utilized stand dominant height as a measure of stand-level competition. Temesgen et al. (2008) stated that stand-level basal area per hectare and crown competition was useful variables for predicting tree height.

Nowadays, there are alternative modeling approaches such as the generalized, fixed and mixed-effects models, artificial neural network techniques, etc that can help to this direction by modeling the relationship between $h-d$ and finally reducing the measurement effort on the ground.

Most basic (e.g., Zhang 1997; Fang and Bailey 1998; Huang et al. 1992; Peng et al. 2001; Soares and Tomé 2002) and generalized $h-d$ models (e.g., López-Sánchez et al. 2003; Adame et al. 2008; Vargas-Larreta et al. 2009; Huang et al. 2009; Crecente-Campo et al. 2010; Lhotka 2012; Gómez-García et al. 2015; Zang et al. 2016; Özçelik et al. 2018) have been developed for many tree species. Many studies have used mixed-models to develop $h-d$ relationships (e.g., Lappi 1997; Calama and Montero 2004; Trincado et al. 2007; Sharma and Parton 2007; Lhotka 2012; Özçelik et al. 2013; VanderSchaaf 2013; Gómez-García et al. 2014; Gómez-García et al. 2015; Zang et al. 2016; Özçelik et al. 2018; Bronisz and Mehtätalo 2020). In Turkey, $h-d$ models were developed for some tree species at regional scale (Diamantopoulou and Özçelik 2012; Özçelik et al. 2014; Çatal and Carus 2018). However, research on the development of $h-d$ model using mixed-effects approach for cedar plantations are very limited (Ercanlı 2015; Özçelik et al. 2013; Özçelik et al. 2018).

Until now, in the field of forest modeling, the implemented research led to the conclusion that artificial neural network models (ANNs) can be considered as a significant alternative modeling technique for many characteristics of trees against to standard modeling methods (Liu et al. 2003; Özçelik et al. 2008; Soares et al. 2013; Imada 2014; Ercanlı et al.

2018; Socha et al. 2020). Moreover, according to the same research field, ANN models have gained publicity because they have successfully been applied for classification (Schmoldt et al. 1997; Sarigul et al. 2003; Cosenza et al. 2017) and estimation and prediction problems (Pertsen et al. 2010; Leite et al. 2011; Reis et al. 2016; Özçelik et al. 2017; Monteiro da Silva et al. 2018; Zhou et al. 2018; Socha et al. 2020). Despite the fact that ANNs need effort for avoiding over or under-fitting the data during their training phase, artificial intelligence has been successfully used for total tree height models construction (Li and Jiang 2010; Diamantopoulou and Özçelik 2012; Özçelik et al. 2013; Vieira et al. 2018; Thanh et al. 2019; Ercanli 2020), as well.

The majority of the research efforts of the literature employ multilayer perceptron architecture combined by the standard backpropagation algorithm for determining the optimal weights. The objective of this study was to investigate the effectiveness of different modeling approaches in terms of providing accurate predictions regarding the total stem height of the cedar tree plantations in Western Mediterranean Region of Turkey. Upon fitting the constructed models and in an effort to enhance the feasibility of our approach, we use easy to obtain measurements from the forest environment, such as the breast height tree stem diameter. Parametric modeling methods such as the fixed-effects, generalized, adjusted fixed-effects and mixed-effects models after being localized using calibration data of one to three sample cedar pine trees were compared against the prediction ability derived by the constructed artificial neural network models. Two different learning algorithms, the Levenberg-Marquardt (LMANN) and the resilient propagation (RPANN) were examined towards optimizing the learning procedure and thus better harness the information residing in real data measurements originating from the forest ecosystem.

Material and Methods

Data

This study was carried out in the Taurus cedar plantations in Western Mediterranean Region of Turkey. The sampling plots are within the Isparta Regional Directorate of Forestry, which includes Isparta and Burdur provinces. The elevation of the sampling plots from the sea level ranges from 1000 to 1600 m. The climate of the study site is characterized as transition climate from the Mediterranean climate to continental climate. It has a variable precipitation regime, while the annual precipitation ranges from 426 to 814 mm, the annual mean temperature in distribution areas of ranges from 11.9 to 13.2 °C (MGM 2013). The bedrock of the study site is composed of sedimentary rock such as limestone, sandstone and clay stone.

The data was collected from 30 sampling plots randomly distributed in the site in order to identify the diameter-height associations of the plantations in the Western Mediterranean Region. The sample plots were selected to represent the existing range of ages, stand densities and sites for the tree species in Turkey. Plot size ranged from 270 to 540 m², depending on stand density, in order to achieve a minimum of 30 trees per plot. The age of the sampling plots varied from 17 to 35. For each tree, two perpendicular diameters (outside-bark 1.3 m above ground level) were measured to the nearest 0.1 cm and were then averaged to obtain diameter at breast height (d , cm). Total heights of these trees were measured to the nearest 0.5 m with a Blume-Leiss hypsometer. The dominant height (H_0) and dominant diameter (D_0) were defined as the average height of the 100 largest-diameter trees per hectare, depending on plot size.

The data were randomly divided into two groups; totally of 21 sample plots (70% of all sample plots) were randomly selected as the model fitting dataset. The remaining 9 sample plots were used as the model validation dataset for evaluating model performance when predictions were made (Table 1). Summary statistics for the fitting and the validation data sets are given in Table 1. Moreover, both the data sets are graphically depicted in Figure 1.

Methods

Generalized h-d models

One practical alternative is to develop a generalized h - d relationship that describes the relationship between d and h and includes stand variables as predictors (e.g. dominant height, quadratic mean diameter, dominant diameter, number of trees per hectare, stand basal area, etc.).

In this study, 21 generalized h - d models selected and evaluated from previous studies (López-Sánchez et al. 2003; Sharma and Zhang 2004; Castedo-Dorado et al. 2006; Sharma and Parton 2007; Crecente-Campo et al. 2010) were fitted using non-linear least squares. These models include generalizations of the most flexible equations for h - d relationships (Zhang 1997; Peng et al. 2001).

Fixed-effects model

A large number of nonlinear model forms were evaluated for cedar plantations; including those reported by Huang et al. (1992) and Lei et al. (2009) as Weibull, Exponential, Chapman-Richards, Gompertz, Schnute, and Korf-Lundgvist. The nonlinear models were fitted by the use of ordinary nonlinear least-squares (ONLS) to the fitting data set. Among the candidate models, the Chapman-Richards produced the best performance to model the h - d relationship for the cedar plantations.

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

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

Calibration of the fixed-effects model

Temesgen et al. (2008) showed that when the heights of a subsample of n_{im} trees from the i^{th} stand are known, the predicted heights of the remaining trees from the same stand can be calibrated by use of the following correction factor:

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

where k^* is correction factor, $\hat{h}_{hi}$ is predicted height from equation (1), and h_{ij} is observed height. Then the predicted height for a tree from the same stand can be calibrated as follows:

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

The effect of the correction factor on the predictive performance of a nonlinear fixed-effects model was also investigated in this study.

Nonlinear mixed-effects model (NLME)

In the mixed-effects framework, all parameters of Eq.(1) can be expressed as fixed-effects parameters (common to all trees), and some or all parameters contain additional random components, which are specific to individual plots. Eq.(1) can be written in matrix form as follows;

$$\mathbf{y}_i = \mathbf{f}(\mathbf{b}, \mathbf{u}_i, \mathbf{d}_i) + \boldsymbol{\varepsilon}_i \quad (4)$$

where $\mathbf{y}_i = [h_{i1}, h_{i2}, \dots, h_{in_i}]^T$, $\mathbf{d}_i = [d_{i1}, d_{i2}, d_{i3}, \dots, d_{in_i}]^T$, $\boldsymbol{\varepsilon}_i = [\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{in_i}]^T$, n_i is number of height measurements for plot i , and $\mathbf{b}$ and $\mathbf{u}_i$ are column vectors of fixed- and random effects parameters, respectively. The assumptions are: $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{R})$, and $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{D})$, where $\mathbf{R}$ and $\mathbf{D}$ are diagonal matrices, assuming that the $\boldsymbol{\varepsilon}_i$ and $\mathbf{u}_i$ are independent. Procedure NLMIXED from SAS (SAS Institute Inc. 2010) was used to obtain fixed- and random-effects parameters of Eq. (4).

The random parameters $\mathbf{u}_i$ for plot i can be predicted by the use of the first-order Taylor series expansion (Meng and Huang 2009):

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

where $\hat{\mathbf{u}}_i^k$ is estimate of the random parameters for tree i at the k th iteration, $\hat{\mathbf{D}}$ is estimate of $\mathbf{D}$, the variance-covariance matrix for $\mathbf{u}_i$, $\mathbf{Z}_i = \left. \frac{\partial \mathbf{f}(\mathbf{b}, \mathbf{u}_i, \mathbf{d}_i)}{\partial \mathbf{u}_i} \right|_{\hat{\mathbf{b}}, \hat{\mathbf{u}}_i}$, $\mathbf{R}$ is the variance-covariance matrix

for $\mathbf{\epsilon}_i$, $\mathbf{y}_i$ is the $m \times 1$ vector of observed heights, and m is number of tree measurements used in localizing the height growth model.

An iterative procedure was needed to estimate $\mathbf{u}_i$. Using a null starting value ($\hat{\mathbf{u}}_i^0 = \mathbf{0}$), Eq.(5) was repeatedly updated until the absolute difference between two successive iterations was smaller than a predetermined tolerance limit. The result was an approximation of the empirical best linear unbiased predictor (EBLUP) for random effects.

ANN models

The most significant advantage of an artificial neural network model is the fact that it is trained from the measured data in hand without any need to add further information. After the system fitting with the real data, the ANN model discovers automatically the existing dependencies, and finally produces the trained model. For this purpose, it is very important the net topology and the learning algorithms, under specific rules, to be specified by the modeler.

Levenberg-Marquardt artificial neural network model (LMANN)

As an intermediate algorithm between the steepest descent method and the Gauss-Newton algorithm, the Levenberg-Marquardt (LM) algorithm (Levenberg 1944; Marquardt 1963) was selected to be applied in the multilayer perceptron learning of ANN. Past studies showed that the (LM) algorithm has the ability on the one hand to overcome the 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 and De Munari 2006; Wilamowski and Yu 2010; Scanlan and Mulvaney 2013; Wu and Ji 2015), on the other hand this learning algorithm showed significant ability on the prediction of environmental variables (Wu 2014; Diamantopoulou et al. 2015; Zu et al. 2018; Özçelik et al. 2019).

The original description of the Levenberg-Marquardt algorithm is given in Levenberg (1944) and Marquardt (1963) papers. A detailed description of the application of Levenberg-

Marquardt algorithm to neural network training is given in Hagan and Menhaj (1994) paper. In a few words LM algorithm introduces the approximation to Hessian matrix (H_m) as (Özçelik et al. 2019):

$$H_m = J^T J + \mu I \quad (6)$$

where J : is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, μ : is the combination coefficient that is always positive and I : the identity matrix.

The update rule of the LM algorithm is expressed by the function $w_{t+1} = w_t - (H_m)^{-1} J^T e$, where w : are the weights, and e : are the biases. The effectiveness and convergence of the LMANN models are very sensitive to the adjustment of the combination coefficient (μ) of equation (6). 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. Given the fact that the Gauss-Newton method's more accurate behavior near an error minimum, the aim was to shift toward the Gauss-Newton method as fast as possible (Diamantopoulou et al. 2015).

Resilient propagation artificial neural network model (RPANN)

The resilient propagation, known as Rprob, is an interesting neural network algorithm, which has been completely described (Riedmiller and Braun 1993) and successfully used (Prasad et al. 2013; Saputra et al. 2017; Florescu and Igel 2018) in order for known disadvantages of the standard back-propagation algorithm to be overcome. Back propagation algorithm has known drawbacks. As the learning of the standard back propagation is more or less elementary, the main problems of its learning face are the slow convergence because of small changes in weights, and the neural network weights trapping around local optima.

The resilient propagation algorithm mainly refers to the direction of the gradient. It calculates a delta value (Δ_{ij}) for each weight that increases when the gradient doesn't change

sign under the prerequisite that the system works on the right direction or decreases when the gradient does change sign. The following learning rule is applied to calculate delta value which improves the change of the network weights (Riedmiller and Braun 1993; Prasad et al. 2013):

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \times \Delta_{ij}^{(t-1)} & , \text{ if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \times \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ \eta^- \times \Delta_{ij}^{(t-1)} & , \text{ if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \times \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)} & , \text{ else} \end{cases} \quad (7)$$

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:

$$w_{ij}^t = w_{ij} + \Delta w_{ij}^{t-1} \quad (8)$$

Some aspects of ANN models' 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 and Dandy 2000). In our case, in order to construct a reliable model for total tree stem height prediction, the choice of the input variables introduced to the input layer of the ANNs was based to sensitivity analysis (Cariboni et al. 2007). Sensitivity analysis was performed to the available variables were selected in advance due to the physical problem to be measured in the field.

In an effort to avoid overfitting and undertraining of the network reaching the best possible training efficiency of the model, the number of hidden nodes in the hidden layer was investigated by starting the training with the smallest possible network, without any hidden nodes. Then the construction of the model continued by adding any necessary hidden node in order the network learning to be deemed complete and the architecture finalized when satisfactory agreement was obtained between network predictions and target outputs.

Overfitting avoidance of the ANN models led to acceptable generalization ability of the models. For this reason, in order to have both training and testing data sets (Leahy 1994), due to its efficiency and easiness to apply, the k-fold cross validation resampling method with $k=10$, was used (Olson and Delen 2008; May et al. 2010). As described in “Data” section, the whole data set was randomly divided into fitting and validation data sets. The fitting data set was used for the construction of all different type models, while the models in their construction phase never saw the validation data set. Validation data set used only for the access of the performance of all models. Especially for the ANN models construction, the 10-fold cross validation division was used for the fitting data set that was further repeatedly divided into training (90% of the fitting data) and testing (10% of the fitting data) data sets as follows (Figure 2).

Due to the fact that the number of weights is strongly related to the ANN model error value, the reliability of the ANN model and its ability to generalize was measured through the error that was estimated as the average error rate on these cross-validation examples (Fig. 2).

As mentioned, the effectiveness and convergence of the training of the LMANN models significantly depends on the proper value selection for the combination coefficient (μ) (Eq. 6). Following the procedure described in Diamantopoulou et al. (2015), the initial value of (μ) was set to 0.001 and the final value of (μ) that led to the lowest sum of square errors model value was found equal to 1.0000e-07.

Similarly to the LMANN models, the training of the RPANN models is very sensitive to the value of delta rule ($\Delta_{ij}^{(t)}$) (Eq. 7). The value of this rule 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 step 1.2 and decrease step 0.5.

The architecture of both different algorithms LMANN and RPANN, finalized when the proper number of hidden nodes in the hidden layer was obtained by the trial and error process, starting from one hidden node until the desired target error value was reached. During the application of trial and error methodology, the selection of the proper combination of the transfer functions was made. The combination that boosted the models' learning to the most trustworthy behavior was a) the hyperbolic tangent transfer function (tansig) that transfers the information from the input layer to the hidden layer and b) the linear activation function (purelin) that transfers the information from the hidden layer to the output layer (Fausett 1994; Beale et al. 2014).

The neural network toolbox of the Matlab software (Beale et al. 2014) developed by The Math Works TM Inc, (R2011b version) that provided by the Aristotle University of Thessaloniki, Greece, was used for the artificial neural network models building.

Evaluation

Different modeling approaches evaluated in this study: (1) fixed-effects, (2) generalized, (3) ANN, (4) adjusted fixed-effects, and (5) calibrated mixed-effects techniques. According to the last two methods, parameters of the adjusted fixed-effects and mixed-effects models were “localized” by use of sampled heights in each plot, and then applied to predict all tree heights in the plots. As recommended by Trincado et al. (2007), three sampling scenarios were considered, corresponding to number of tree heights (ranging from one to three) measured in each plot. In order the generalization ability of the constructed models by the different modeling approaches to be compared and evaluated, the following statistics computed for the validation data set (Özçelik et al. 2018; Bronisz and Mehtätalo 2020):

$$\text{Mean difference (m)} : MD = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})}{\sum_{i=1}^n n_i} \quad (9)$$

$$\text{Mean absolute difference (m)} : MAD = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} |h_{ij} - \hat{h}_{ij}|}{\sum_{i=1}^n n_i} \quad (10)$$

$$\text{Fit index} : FI = 1 - \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})^2}{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \bar{h}_i)^2} \quad (11)$$

$$\text{Root mean square error} : RMSE = \sqrt{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (h_{ij} - \hat{h}_{ij})^2}{\sum_{i=1}^n n_i} \right)} \quad (12)$$

where n is number of plots, n_i is number of tree measurements for plot i , h_{ij} and $\hat{h}_{ij}$ are observed and predicted values of tree height, respectively, and $\bar{h}_i$ is average value of h_{ij} for plot i .

Results

Nonlinear regression models (Generalized models, fixed-and mixed-effects models)

As stated earlier, twenty-one generalized h - d models selected from previous studies and then the models were fitted using *ONLS* method using the fitting data set while the predictive capability of the models were evaluated using the validation dataset. The final generalized- h - d model (Krumland and Wensel 1988) including dominant diameter and dominant height as additional independent variables, resulting in:

$$h = 1.3 + (H_0 - 1.3) \left(\exp(\beta_1 d^{\beta_2 (H_0 - 1.3)}) / \exp(\beta_1 D_0^{\beta_2 (H_0 - 1.3)}) \right) \quad (13)$$

Different combinations of mixed parameters from equation (1) were implemented (Table 2). The results of Table 2 shows that model (1) with random components β_1 and β_3 produced the lowest values of Akaike's information criterion (AIC) (Akaike 1974) and Bayesian information criterion (BIC) among various combinations of mixed parameters.

The final mixed-effects model was:

$$h_{ij} = h_{ij} = 1.3 + (\beta_1 + u_1) (1 - \exp(-\beta_2 d_{ij}))^{(\beta_3 + u_2)} \quad (14)$$

where u_1 and u_2 are random parameters.

Parameter estimates for the best models (generalized fixed-effects, and mixed-effects) were obtained using the fitting data set (Table 3).

Compared to fixed-effects models, mixed-effects models require prior height information, which is not necessary for the case of the fixed-effects models. The latest models consider all observations as independent and the residuals are measured against the population average. On the contrary, the residuals of the model including random parameters take into consideration tree-specific predictions (Bronisz and Mehtätalo 2020).

Artificial neural network modeling approaches

The aspect of the ANN models was the three-layer architecture that includes a) one input layer with three variables (diameter at breast height, d ; dominant height, H_0 ; and dominant diameter, D_0). Sensitivity analysis performed to the available variables. The above three variables were found to have significant entry on the target variable configuration, with value of error ratio higher to one; b) one output layer with one target variable (total tree stem height, h) specified by the physical problem, and c) one hidden layer, consisting of the proper number of hidden nodes specified by the network learning.

Using the Levenberg-Marquardt (LM) algorithm for the ANN models learning, the Levenberg-Marquardt Artificial Neural Network (LMANN) model showed the best adaptation to the fitting data set, was the 3-2-1/0.5512 model. This means that the specific LMANN model consisted of one input layer with three input nodes (d , H_0 and D_0), one hidden layer with two hidden nodes and one output layer with one output node (h), with value of the root mean square error equals to 0.5512. Using the Resilient Propagation (RP) algorithm for the ANN models learning, the Resilient Propagation Artificial Neural Network (RPANN) model showed the best adaptation to the fitting data set, was the 3-2-1/0.5387 model. The specific RPANN model consisted of one input layer with the same input variables (d , H_0 and D_0), one output layer with the same output node (h) as in LMANN models building and between them one hidden layer with two hidden nodes as resulted by the error and trial procedure that has been followed. The value of the root mean square error was found equal to 0.5387. For both ANN algorithms, the

optimum number of hidden nodes in their hidden layers was determined after the examination of different number of nodes. The best values were selected as the values that led to the best learning of the ANN models, with the smallest root mean square error value (Figure 3).

Moreover, the $k=10$ cross validation method applied in the fitting data set for both ANN algorithms for the models training and testing in their construction phase using the fitting data set, ensures the absence of overfitting and their ability to generalize. Both networks were trained for as many epochs as it needed in order the models to reach to the minimum mean square error value (Figure 4). As can be seen (Figure 4), the estimation errors for the training and the test data sets were not significantly different for both ANN models, fact that indicated the absence of overfitting in the modeling systems. The LMANN was stopped after 200 training interactions (epochs) as there was no influential reduction of the mean square error value after the 28 epochs, while the RPANN was trained for 500 interactions as there was no influential reduction of the mean square error value after the 167 epochs (Fig. 4).

Evaluation of the different modeling approaches

According to the fitting and the validation data set, the evaluation statistics calculated for the different modeling approaches and their values are shown in Table 4. The best results among the generalized $h-d$ models were obtained from Krumland and Wensel (1988) are presented in Table 4. Comparing the MD, MAD, RMSE, and FI values (Table 4), Krumland and Wensel (1988) model gave relatively smaller MD and MAD and bigger FI values than the other generalized models.

On top of the results of Table 4, the accuracy of the estimation results of the ANN models was examined by the correlation coefficient (r) between the h values and their estimations by both ANN models. According to the estimations obtained by the LMANN model, the r value was equal to 0.9223, while the r value for the estimations obtained by the RPANN model was equal to 0.9259. The variability of the estimation errors in relation to the

mean of the total tree heights ($CV\%$) in the fitting data set was 9.00% and 8.80% for the LMANN and the RPANN models, respectively. The reliability and the predictive capabilities of both ANN models was also examined based on the new validation data set consisting of 270 trees in 9 plots in the same area, never seen by the constructed models before. In order to further investigate the ability of our models to generalize, r , and the coefficient of variation ($CV\%$) between the h values and their predictions by both ANN models were calculated. According to the predictions obtained by the LMANN model, the r value was equal to 0.9318 and the $CV\%$ was 10.27%, while the r value for the RPANN model was equal to 0.9336, with $CV\%$ equals to 10.02%.

Discussion

Standard modeling techniques assessment

The adjusted fixed-effects model had a much lower mean deviation (MD) for all calibration alternatives compared to the fixed-effects model. Considering all calibration alternatives, the fit index increased at variable rates from 27% to 39%. The mean increase was 35%. Similar findings were also reported by Lhotka (2012), VanderSchaaf (2013) and Özçelik et al. (2018). In this study, the most remarkable prediction improvement among the tested calibration alternatives was obtained through the use of three trees for calibration in each sampling plot. Different sampling alternatives tested for both adjusted fixed-effects model and NLME model, corresponding to number of trees measured in each plot. These tree heights were used to estimate random parameters of NLME model and also to calibrate heights and calculate the evaluation statistics for the validation data set (Table 4).

The mixed-effects model had lower mean deviation values obtained through the calibration with 3 trees compared to the adjusted fixed-effects model. Considering all calibration alternatives, there was a decrease of 32% and 39% on average in the mean deviation

and mean absolute difference with the use of mixed-effects model, respectively. Previous studies (Trincado et al. 2007; Temesgen et al. 2008; Huang et al. 2009; VanderSchaaf 2013; Gómez-García et al. 2014; Özçelik et al. 2018) also demonstrated that the calibration improved height predictions significantly. Table 4 reveals that there were no big differences between the calibrations performed with 2 and 3 trees for cedar plantations through testing different calibration alternatives with the use of the mixed-effects model, while there were similar trends between different calibration alternatives as regards MAD and FI values. When the calibrated fixed effect model and the calibrated mixed effect model are compared, it is seen that for all calibration alternatives, the mixed effect model is more successful for height estimates in cedar plantations than calibrated fixed-effects model.

The results obtained with the model developed by Krumland and Wensel (1988); compared to the results obtained for the fixed effect model according to all used evaluation statistics, the generalized *h-d* model produced more successful results than the fixed effect model. If a general evaluation is made for all criterion values, the generalized *h-d* model achieved an average of 17% reduction in MAD values, when compared to the fixed effect model. Generally, the generalized *h-d* models produced better results than adjusted-fixed models for all evaluation statistics except for MD values.

According to the impact of calibration sample size on prediction accuracy, the findings demonstrated that the height predictions improved depending on the increase in the number of trees used for calibration as also reported by Trincado et al. (2007), Temesgen et al. (2008), Huang et al. (2009), VanderSchaaf (2013) and Özçelik et al. (2018). However, the success of prediction started decreasing substantially after a certain number of trees. Therefore, the number of trees to be used for calibration should be selected in a way to strike a balance between the inventory-taking cost and the prediction accuracy and success. Calama and Montero (2004) suggested that calibration should be performed with additional trees to increase

the success of height predictions. Trincado et al. (2007) argued that the accuracy of height predictions increased with the increased number of trees to be used for calibration but that increase was not remarkable. We also found a similar result in our study and the prediction success in the calibration with the use of three. Similar results were also reported by Temesgen et al. (2008) and Huang et al. (2009) used a series of 1 to 9 trees for calibration but the best result in height predictions was achieved in the calibration performed with 1 tree. Adame et al. (2008) tested different calibration alternatives and reported that it was necessary and adequate to use 2 or 3 trees for calibration. Generally, this study reported that there was a minimal improvement in prediction capability of the model depending on the increasing number of trees beyond the number of three trees used for calibration.

Artificial neural network modeling techniques assessment

Due to their ability to overcome known disadvantages and limitations of the standard back-propagation algorithm (Rumelhart et al. 1986; Fausett 1994; Patterson 1996; Wilamowski and Yu 2010) which is the most used optimisation technique applied for the multilayer-perceptron training of ANNs, the Levenberg-Marquardt (LM) and the Resilient Propagation (RP) algorithms were used. According to the proximity of each point to the 45-degree line throughout the range of the measured total heights (Fig. 5b) indicates that both constructed ANN models were showed similar adjustment to the measured data. In addition, the RPANN model appears to provide slightly more accurate adaptation as compared to the LMANN model (Fig. 5a).

Moreover, taking into account the statistical exploration of the calculated estimation errors derived by the ANN models, having the errors frequency peak at zero, quickly falling off at the two sides of smaller error values, both ANN models can be considered as well-trained healthy networks.

Finally, the results of Table 4 and the ones of Figures 4 and 5 regarding the fitting data set indicate that the constructed LMANN model as compared to the RPANN model, showed more or less similar performance, with similar results in both r and $CV\%$ values. As derived by all statistics given for both ANN models, it is worth noting that the most accurate estimation results are derived by the best RPANN model, with root mean square error value 2.25% lower than the error value obtained from the best LMANN model.

Given the statistics calculated for the validation data set of Table 4 along with the r and $CV\%$ values, RPANN model provides the most accurate and reliable prediction results as compared to the ones of the LMANN model. The variability of the RPANN model prediction errors ($CV\%$) was about 2.43% lower than the ones derived from the LMANN model. As shown in Figure 6, the prediction errors histogram and the relative errors box plots showed more or less similar effectiveness of the predictive ability of the LMANN and the RPANN models, with slightly less error variability of the stem height predictions as derived from the constructed RPANN model.

Finally, according to the results of Table 4, upon comparing all modeling approaches, the resilient propagation artificial neural network model showed a clear estimation and prediction superiority over fixed, generalized, fixed-calibrated and mixed calibrated models.

Conclusion

In the context of this work, height prediction models were developed for cedar pine plantations in Western Mediterranean Region of Turkey, which involve the evaluation of five alternative methods; (1) fixed-effects model, (2) generalized h-d models, (3) adjusted fixed-model, (4) mixed-effects model, and (5) ANN models. Except for the modelling process itself, this methodology in the cases of the adjusted and mixed-effects models also employs the calibration of the parameters of the fixed-effects and mixed-effects models using a subset of

height measurements, ranging from 1 to 3 sample trees per plot. As compared to the fixed-effects, all constructed models (generalized, adjusted fixed-effects, mixed-effects, LMANN and BPANN models) achieve better performance given the fact that they provide more accurate and reliable tree height predictions. Given the results, it is obvious that the performance of the adjusted and mixed-effects models improved with increasing sample size. However, increasing the number of trees used in the calibration had a minimal effect on improving the predictive ability of the models.

As for the constructed ANN models, it seems that both alternatives exhibit similar performance in their construction phase as well as in terms of their ability to generalize. In addition, based on the cumulative results regarding all evaluated approaches, it is obvious that the resilient propagation algorithm is more effective in terms of predicting the total stem height of the cedar trees originating from Western Mediterranean Region of Turkey. Conclusively, it is worth noting that the effectiveness and efficiency of ANN modelling in general and in specific the resilient propagation algorithm along with its fit to the current dataset provides raises the potential of its applicability to further modelling cases that lie in the area of forest management. Such directions constitute interesting future research endeavours.

Acknowledgements

Not applicable

Funding

Cedar data for the present study was provided from financially supported project by Suleyman Demirel University, Scientific Research Project Coordination Unit (Project No: 2262-YL-10).

Availability of data and materials

The dataset used and/or analysed during the current study are available from corresponding author on reasonable request.

Authors' contributions

YK and ZS were responsible for collecting the data. RÖ and MD performed the statistical analysis. The manuscript was prepared by YK, MD, RÖ and ZS. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Authors' information

¹Faculty of Forestry, Isparta University of Applied Science, East Campus, 32260, Isparta, Turkey.

²Faculty of Forestry and Natural Environment, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece, ³Ministry of Agriculture and Forestry, VI. Regional Directorate, 27002, Burdur, Turkey.

Literature Cited

- Adame P, del Río M, Canellas I (2008) A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). For Ecol Manag 256(1-2):88-98.
- Adeloye A, De Munari A (2006) Artificial neural network based generalized storage–yield–reliability models using the Levenberg–Marquardt algorithm. J Hydrol 326(1-4):215-230.
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19(6):716-723.
- Beale M, Hagan M, Demuth H (2014) Neural network Toolbox™ User's Guide. The MathWorks Inc., Natick, MA.
- Boydak M (2003) Regeneration of Lebanon cedar (*Cedrus libani* A. Rich.) on karstic lands in Turkey. For Ecol Manag 178(3):231-243.
- Bronisz K, Mehtätalo L (2020) Mixed-effects generalized height–diameter model for young silver birch stands on post-agricultural lands. For Ecol Manag 460:117901.
- Calama R, Montero G (2004) Interregional nonlinear height diameter model with random coefficients for stone pine in Spain. Can J For Res 34(1):150-163.
- Cariboni J, Gatelli D, Liska R, Saltelli A (2007) The role of sensitivity analysis in ecological modelling. Ecol Model 203(1-2):167-182.
- Catal Y, Carus S (2018) A height-diameter model for brutian pine (*Pinus Brutia* Ten.) plantations in southwestern Turkey. Appl Ecol Environ Res 16:1445-1459.
- Çatal Y (2012) Height-diameter model for black locust, Anatolian black pine and Taurus cedar tree species in Lakes Region. Turk J For 13(2):92-96.
- Colbert KC, Larsen DR, Lootens JR (2002) Height-diameter equations for thirteen midwestern bottomland hardwood species. North J Appl For 19(4):171-176.

- Corral-Rivas S, Álvarez-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. *For Ecosyst* 1(1):6.
- Cosenza DN, Soares AAV, Alcântara AEMd, Silva AALd, Rode R, Soares VP, Leite HG (2017) Site classification for eucalypt stands using artificial neural network based on environmental and management features. *Cerne* 23(3):310-320.
- 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. *For Ecol Manag* 259(5):943-952.
- Curtis RO (1967) Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *For Sci* 13(4):365-375.
- da Silva Jr EM, Maia RD, Cabacinha CD (2018) Bee-inspired RBF network for volume estimation of individual trees. *Comput Electron Agric* 152:401-408.
- Diamantopoulou M, Özçelik R (2012) Evaluation of different modeling approaches for total tree-height estimation in Mediterranean Region of Turkey. *For Ecosyst* 21(3):383-397.
- Diamantopoulou MJ, Özçelik R, Crecente-Campo F, Eler Ü (2015) Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. *Biosyst Eng* 133:33-45.
- Dorado FC, Diéguez-Aranda U, Anta MB, Rodríguez MS, von Gadow K (2006) A generalized height–diameter model including random components for radiata pine plantations in northwestern Spain. *For Ecol Manag* 229(1-3):202-213.
- Ercanlı I (2015) Nonlinear mixed effect models for predicting relationships between total height and diameter of Oriental beech trees in Kestel, Turkey. *Rev Chapingo Ser Cie* 21(2):185-202.

- Ercanlı İ, Günlü A, Şenyurt M, Keleş S (2018) Artificial neural network models predicting the leaf area index: a case study in pure even-aged Crimean pine forests from Turkey. For Ecosyst 5(1):29.
- Ercanlı, İ (2020) Innovative deep learning artificial intelligence applications for predicting relationship between individual tree height and diameter at breast height. For Ecosyst 7:12.
- Fang Z, Bailey R (1998) Height–diameter models for tropical forests on Hainan Island in southern China. For Ecol Manag 110(1-3):315-327.
- Fausett L (1994) Fundamentals of neural networks: architectures, algorithms, and applications. Prentice-Hall, NJ, United States.
- Florescu C, Igel C (2018) Resilient backpropagation (RPROP) for batch-learning in tensorflow. Paper presented at International Conference on Learning Representations, Vancouver, Canada, 30 April- 03 May 2018.
- GDF (2015) Forest Resources. The General Directorate of Forests, Ankara.
- Gomez-Garcia E, Fonseca TF, Crecente-Campo F, Almeida LR, Dieguez-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 9(1):72.
- 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. For Sci 60(3):560-568.
- Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw Learn Syst 5(6):989-993.
- Huang S, Titus SJ, Wiens DP (1992) Comparison of nonlinear height–diameter functions for major Alberta tree species. Can J For Res 22(9):1297-1304.

- Huang S, Wiens DP, Yang Y, Meng SX, Vanderschaaf CL (2009) Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods. *For Ecol Manag* 258(7):1235-1247.
- Imada AA (2014) Literature review: Forest management with neural network and artificial intelligence. In *International Conference on Neural Networks and Artificial Intelligence*. Springer, pp 9-21.
- 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. *Ann For Sci* 74(1):7.
- Konukçu M (2001) *Forests and Forestry*. State Planning Organization publications, Ankara. (In Turkish).
- Krumland BE Wensel LC (1988) A generalized height-diameter equation for coastal California species. *West J Appl For* 3(4):113-115.
- Lappi J (1997) A longitudinal analysis of height/diameter curves. *For Sci* 43(4):555-570.
- Leahy K (1994) The overfitting problem in perspective. *AI Expert* 9(10):35-36.
- 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. *For Chron* 85(1):43-56.
- 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. *Eur J For Res* 130(2):263-269.
- Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. *Q Appl Math* 2(2):164-168.
- Lhotka JM (2012) Height-diameter relationships in Sweetgum (*Liquidambar styraciflua*)-dominated stands. *South J Appl For* 36(2):98-106.

- Li YX, Jiang LC (2010) Application of ANN Algorithm in Tree Height Modeling. In: Applied Mechanics and Materials, vol 20. Trans Tech Publ. pp 756-761.
- Lindstrom MJ, Bates DM (1990) Nonlinear mixed effects models for repeated measures data. Biometrics 673-687.
- 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. For Sci 49(4):619-631.
- Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Modell Softw 15(1):101-124.
- Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 11(2):431-441.
- May RJ, Maier HR, Dandy GC (2010) Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Netw 23(2):283-294.
- McClelland JL, Rumelhart DE, Group PR (1986) Parallel distributed processing. Explorations in the Microstructure of Cognition 2:216-271.
- Meng SX, Huang S (2009) Improved calibration of nonlinear mixed-effects models demonstrated on a height growth function. For Sci 55(3):238-248.
- MGM (Turkish State Meteorological Service) (2013) Official statistics (Statistical Data of Isparta Province and Districts) (In Turkish).
- 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. For Ecol Manag 247(1-3):149-166.
- Olson D, Delen D (2008) Advanced data mining techniques. Springer Science & Business Media, Berlin.

- Özçelik R, Diamantopoulou MJ, Trincado G (2019) Evaluation of potential modeling approaches for Scots pine stem diameter prediction in north-eastern Turkey. *Comput Electron Agric* 162:773-782.
- Özçelik R, Diamantopoulou MJ, Wiant Jr HV, Brooks JR (2008) Comparative study of standard and modern methods for estimating tree bole volume of three species in Turkey. *For Prod J* 58(6):73.
- Özçelik R, Diamantopoulou MJ, Crecente-Campo F, Eler U (2013) Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. *For Ecol Manag* 306:52-60.
- Özçelik 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. *For Ecol Manag* 419:240-248.
- Özçelik R, Diamantopoulou MJ, Eker M, Gürlevik N (2017) Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. *For Sci* 63(3):291-302.
- Özçelik R, Yavuz H, Karatepe Y, Gürlevik N, Kırış R (2014) Development of ecoregion-based height-diameter models for 3 economically important tree species of southern Turkey. *Turk J Agric For* 38(3):399-412.
- Patterson DW (1996) *Artificial Neural Networks, Theory and Applications*. Prentice Hall, Singapore.
- Peng C, Zhang L, Liu J (2001) Developing and validating nonlinear height–diameter models for major Tree species of Ontario's boreal forests. *North J Appl For* 18(3):87-94.
- Prasad N, Singh R, Lal SP (2013) Comparison of back propagation and resilient propagation algorithm for spam classification. In *2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation*, pp 29-34.

- 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. *For Ecol Manag* 382:161-167.
- Riedmiller M, Braun HA (1993) Direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *IEEE international conference on neural networks*, pp 586-591.
- Robinson AP, Wykoff WR (2004) Imputing missing height measures using a mixed-effects modeling strategy. *Can J For Res* 34(12):2492-2500.
- Saatçioğlu F (1979) *Silviculture II*. Istanbul University Faculty of Forestry publications, Istanbul. (In Turkish).
- Sánchez CAL, Varela JG, Dorado FC, Alboreca AR, Soalleiro RR, González JGÁ, Rodríguez FS (2003) A height-diameter model for *Pinus radiata* D. Don in Galicia (Northwest Spain). *Ann For Sci* 60(3):237-245.
- Saputra W, Zarlis M, Sembiring RW, Hartama D (2017) Analysis Resilient Algorithm on Artificial Neural Network Backpropagation. In *Journal of Physics: Conference Series*, IOP Publishing, p 012035.
- Sarigul E, Abbott AL, Schmoldt DL (2003) Rule-driven defect detection in CT images of hardwood logs. *Comput Electron Agric* 41(1-3):101-119.
- SAS Institute Inc. (2010) *SAS/OR® 9.22 User's Guide: Mathematical programming*. SAS Institute Inc, Cary, NC.
- Scanla D, Mulvaney D (2013) Graphics processor unit hardware acceleration of Levenberg-Marquardt artificial neural network training. *Research Inventory: Int J Eng Sci* 2(7):1-7.
- Schmoldt DL, Li P, Abbott AL (1997) Machine vision using artificial neural networks with local 3D neighborhoods. *Comput Electron Agric* 16(3):255-271.

- Sharma M, Yin Zhang S (2004) Height–diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*. Scand J For Res 19(5):442-451.
- Sharma M, Parton J (2007) Height–diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. For Ecol Manag 249(3):187-198.
- 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 Comput Appl 22(7-8):1407-1418.
- Soares P, Tomé M (2002) Height–diameter equation for first rotation eucalypt plantations in Portugal. For Ecol Manag 166(1-3):99-109.
- Socha J, Netzel P, Cywicka D (2020) Stem Taper Approximation by Artificial Neural Network and a Regression Set Models. Forests 11(1):79.
- Temesgen H, Monleon V, Hann D (2008) Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Can J For Res 38(3):553-565.
- Trincado G, VanderSchaaf CL, Burkhart HE (2007) Regional mixed-effects height–diameter models for loblolly pine (*Pinus taeda* L.) plantations. Eur J For Res 126(2):253-262.
- VanderSchaaf CL (2012) Mixed-effects height-diameter models for commercially and ecologically important conifers in Minnesota. North J Appl For 29(1):15-20.
- Vargas-Larreta B, Castedo-Dorado F, Álvarez-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.
- Vieira GC, de Mendonça AR, da Silva GF, Zanetti SS, da Silva MM, dos Santos AR (2018) Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. Sci Total Environ 619:1473-1481.
- Wilamowski BM, Yu H (2010) Improved computation for Levenberg–Marquardt training. IEEE Trans Neural Netw Learn Syst 21(6):930-937.

- Wu D (2014) Estimation of forest volume based on LM-BP neural network model. *Comput Model New Technol* 18:131-137.
- Wu D, Ji Y (2015) Dynamic estimation of forest volume based on multi-source data and neural network model. *J Agric Sci* 7(3):18.
- 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.
- Zhang L (1997) Cross-validation of non-linear growth functions for modelling tree height–diameter relationships. *Ann Bot* 79(3):251-257.
- Zhang X, Duan A, Zhang J, Xiang C (2014) Estimating tree height-diameter models with the Bayesian method. *Sci World J* 2014:683691.
- 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.

Figures

Figure 1

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

Figure 2

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

Figure 3

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

Figure 4

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

Figure 5

45-degree plot for a) the LMANN and b) the RPANN model height estimations.

Figure 6

Errors histogram for a) the LMANN and b) the RPANN model height predictions, and c) errors box plots for both ANN models, for the validation data set.