

# Yield of forests in Ankara Regional Directory of Forestry in Turkey: comparison of regression and artificial neural network models based on statistical and biological behaviors

Ferhat Bolat,  
Ilker Ercanli,  
Alkan Günlü

Models of forest growth and yield provide important information on stand and tree developments and the interactions of these developments with silvicultural treatments. These models have been developed based on assumptions such as independence of observations, uncorrelated error terms, and error terms with constant variance; if these factors are absent, there may be problems with multicollinearity, autocorrelation, or heteroscedasticity, respectively. These problems, which have several adverse effects on parameter estimates, are statistical phenomena and must be avoided. In recent years, the artificial neural network (ANN) model, thanks to its superior features such as the ability to make successful predictions and the absence of the requirement for statistical assumptions, has been commonly used in forestry modeling. However, while goodness-of-fit measures were taken into consideration in the assessment of ANN models, the control of the biological characteristics of model predictions was ignored. In this study, variable-density yield models were developed using nonlinear regression and ANN techniques. These modeling techniques were compared based on some goodness-of-fit measures and the principles of forest yield. The results showed that ANN models were more successful in meeting expected biological patterns than regression models.

**Keywords:** Bayesian, Machine Learning, Gompertz, Overfitting

## Introduction

Regression models have been widely used to maximize the explained variation in forest measurements, such as individual tree heights, individual tree diameter increments, stand basal areas, stand volumes, and site indexes (Vanclay 1994). The reliability of regression models depends on the fulfillment of certain statistical assumptions (Zuur et al. 2010). Because forest inventory data are spatially or temporally correlated (Fox et al. 2001), the statistical assumptions are usually violated, and this leads to biased estimates (West et al. 1984). To deal with the statistical problems,

the autoregressive modeling technique (Hirigoyen et al. 2021) and mixed-effect modeling technique (Karatepe et al. 2022) can be used. However, some issues associated with these modeling techniques have remained in question (Wang et al. 2008). For instance, it may be problematic to determine the parameter of random effects in a mixed effects model. While some studies take goodness-of-fit measures into consideration (Ogana et al. 2020, Raptis et al. 2021), others consider the variance in the estimated parameters (Fang & Bailey 2001, Cheng 2021).

Artificial neural network (ANN) models that are independent of statistical assumptions have been widely used in forestry modeling. In earlier studies, ANN models were commonly employed to classify forest characteristics (Kanellopoulos et al. 1993, Foody et al. 1997), but recently they have been used for both the classification and prediction of forest characteristics such as site index (Lima et al. 2017) and stand basal area (Che et al. 2019).

Most researchers have paid attention to goodness-of-fit statistics in evaluating the performance of ANN models (Ashraf et al. 2013, Karatepe et al. 2022). However, growth and yield predictions for forests should also be evaluated based on the principles of forest growth and yield (Weiskittel et al. 2011, Seki & Sakici 2022). To our knowledge, no studies have explored the success of ANN models in meeting biological

expectations. In addition, many researchers have provided very little information about the setting of the network parameters. This study found that when the network parameters are not correctly set, the predictions may be implausible from a biological point of view even if the statistical scores are satisfactory.

The objectives of this study were to (i) develop variable-density yield models using a modified Gompertz model and an ANN model and (ii) compare their performances based on goodness-of-fit statistics and the principles of forest yield.

## Materials and methods

### Study area

In this study, natural monospecific Crimean pine stands (*Pinus nigra* Arnold subsp. *pallasiana* Lamb.) were sampled. They cover 44.8% (136,985.96 ha) of the total forest area in the Ankara Regional Directory of Forestry in Turkey (GDF 2020). Other common species in this area are various oak species (*Quercus cerris* L., *Quercus infectoria* Olivier, *Quercus frainetto* Ten., *Quercus robur* L., and *Quercus petraea* Matt.) and the Scots pine (*Pinus sylvestris* L.). Fig. 1 shows the sample plots and location of the study area (33° 48' 54" - 37° 10' 06" N; 39° 51' 28" - 41° 02' 47" E). The elevation of the study area varied between 239 m and 2541 m above mean sea level. The mean annual precipitation and mean an-

□ Çankiri Karatekin University, Faculty of Forestry, 18200, Çankiri (Turkey)

@ Ferhat Bolat (fbolat@karatekin.edu.tr)

Received: Apr 14, 2022 - Accepted: Nov 04, 2022

**Citation:** Bolat F, Ercanli I, Günlü A (2023). Yield of forests in Ankara Regional Directory of Forestry in Turkey: comparison of regression and artificial neural network models based on statistical and biological behaviors. *iForest* 16: 30-37. - doi: 10.3832/ifor4116-015 [online 2023-01-22]

Communicated by: Maurizio Marchi

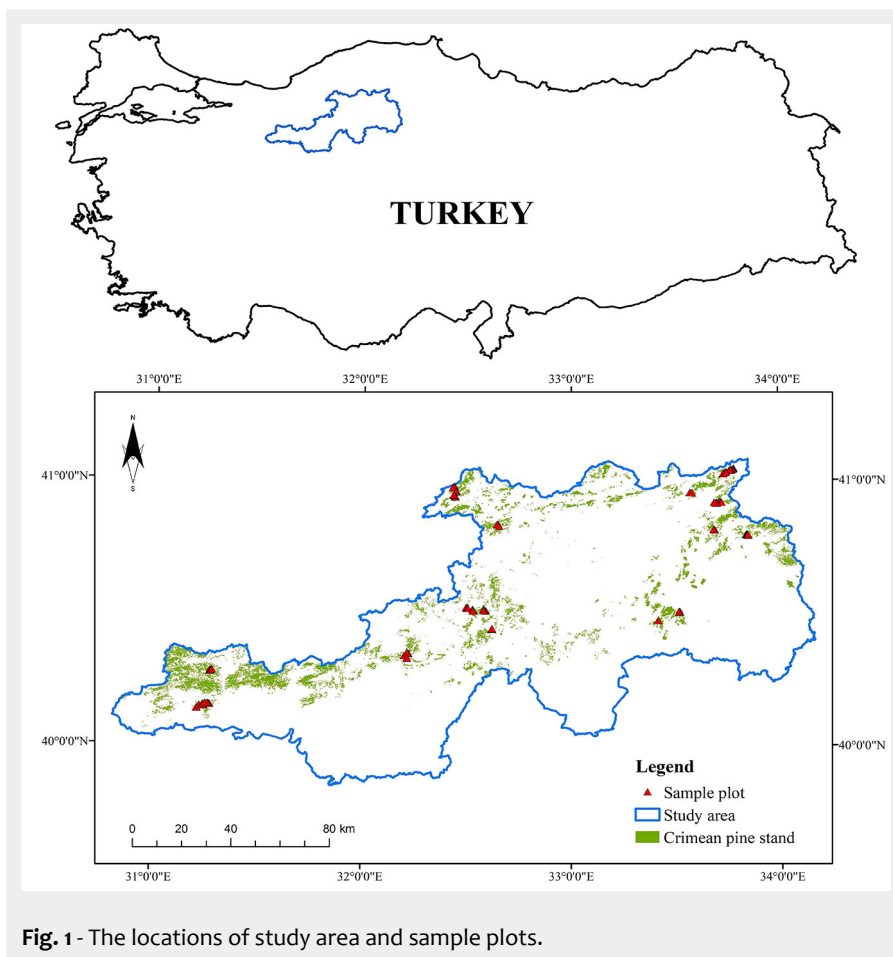


Fig. 1 - The locations of study area and sample plots.

were measured at 1.3 m above the ground; (ii) the mean dominant height of the sample plots was calculated by averaging the height of a certain number of trees corresponding to the 100 tallest trees per hectare; then, the site index value of the sample plots was assigned using the site index table developed by Kalipsiz (1963); (iii) the mean age of the sample plots was calculated by averaging the age of the representative four trees considering the quadratic mean diameter of the sample plots; (iv) to establish height-diameter relationships, the height of 10 trees having different diameters at breast height in each sample plot was subsampled.

Tab. 1 shows the descriptive statistics of the Crimean pine stands.

Regression-type variable-density yield models

The selection of a suitable yield function is essential for modeling forest development. In forestry modeling, sigmoid functions such as the Lundqvist-Korf, Richards, and Hossfeld IV/McDill-Amateis functions have been widely used (Burkhardt & Tomé 2012) owing to their flexible forms. They produce the most appropriate curves for any stand characteristic depending on the site index and stand density. In this study, the nonlinear variable-density yield models (eqn. 1 to eqn. 6) were developed using the Gompertz function, which is one of the Richards-type growth functions. To facilitate the selection of starting values for the parameters, a solver tool was developed using the Generalized Simulated Annealing package in the R environment (see Appendix 1 in Supplementary material).

$$\ln(G) = b_1 \exp(-\exp(b_2 - b_3 b_2(1/A))) \quad (1)$$

$$b_2 = a_1(1/SI) + a_2(1/SD) + a_3(SI)^2 + a_4(SD)^2 \quad (2)$$

$$\ln(V) = b_1 \exp(-\exp(b_2 - b_3 b_2(1/A))) \quad (3)$$

$$b_2 = a_1 \ln(SI) + a_2(1/SD) + a_3(SI)^2 + a_4(SD)^2 \quad (4)$$

$$\ln(N) = b_1 \exp(-\exp(b_2 - b_3 b_2(1/A))) \quad (5)$$

$$b_2 = a_1 \frac{1}{SI} + a_2 \frac{1}{SD} + a_3 \ln(SI) + a_4 \ln(SD) \quad (6)$$

where G is the stand basal area (m<sup>2</sup> ha<sup>-1</sup>), V is the stand volume (m<sup>3</sup> ha<sup>-1</sup>), N is the number of trees (ha), A is the stand age (years), SD is the stand density developed by Curtis (1982), SI is the site index (m), and b<sub>1</sub>, ..., b<sub>3</sub>, and a<sub>1</sub>, ..., a<sub>4</sub> are the regression coefficients to be estimated.

In general, the variable-density yield models include SI, SD, and A as explanatory variables. In addition, the different explanatory variables derived from SI, SD, and A (e.g., ln(SI), ln(SD), SI<sup>2</sup>) can be used to effectively describe the development of the stands. In this study, the prediction accuracy and the principles of forest yield were considered when deciding which variable to use. The principles of forest yield considered in this study are: (i) V and G increase with increas-

nual temperature were 393.2 mm and 11.9 °C, respectively (TSMS 2022).

Field measurements

Forest yield models should ideally be developed using data from permanent sample plots. However, this study used data from temporary sample plots because there were no permanent sample plots of natural monospecific Crimean pine stands grown in the study area. The sample plots inventoried were selected to be representative of the range of site indexes, stand densities, and stand ages. A total of 180 sample plots were randomly chosen; in each, a circular sampling scheme was

adopted because the stands, which have different stand densities in same site index and have different ages in same site index and stand density, should be adequately sampled. If the number of samples is insufficient in terms of site index, stand density, and age, a variable-stand density model may not meet the principles of forest yield. The area of the sample plots was 400, 600, or 800 m<sup>2</sup> according to the crown closure of the plots. This study followed the forest management guidelines of Turkey in determining the area of the sample plots (Anonymous 2017). The measurements in the sample plots were: (i) all trees (≥ 4 cm in diameter at breast height)

Tab. 1 - Descriptive statistics of Crimean pine stands studied in this study. (Dq): quadratic mean diameter; (Hq): quadratic mean height; (A): age; (G): basal area; (N): stem number; (V): volume derived from the double entry tree volume equation  $V = 0.000202 \cdot d^{1.602553} h^{0.969154}$  (Bolat 2021); (SI): site index; (SD): stand density (Curtis 1982).

Variables	n	Min	Max	Mean	STD
Dq (cm)	180	2.00	56.30	27.14	10.89
Hq (m)	180	2.00	31.30	15.25	6.31
A (year)	180	13.50	161.67	74.11	32.87
G (m <sup>2</sup> ha <sup>-1</sup> )	180	2.00	70.00	31.11	14.46
N (number ha <sup>-1</sup> )	180	137.50	1550.00	630.02	327.59
V (m <sup>3</sup> ha <sup>-1</sup> )	180	7.00	921.67	323.79	216.15
SI (m)	180	10.00	34.00	21.16	6.68
SD	180	1.25	16.92	6.24	2.70

ing *SI*, but *N* declines; (ii) *V*, *G*, and *N* increase with increasing *SD*; (iii) *V* and *G* increase with age; (iv) *N* decreases with age – this reduction is generally high in young stands due to strong inter-tree competition, which leads to a high probability of mortality (Meyer 1938); (v) *V*, *B*, and *N* may develop in totally different ways depending on the *SI* and *SD* in the managed forests – the stands of the same age but in different *SIs* and *SDs* may have a variable proportion of yield; therefore, the S-shape of the yield curves should be characterized by the polymorphic curves; (vi) the yield curves should reach an asymptote (Gilbert et al. 2010) – in the absence of an asymptote, the long-term forecasts of the stand variables are limited and may even be invalid.

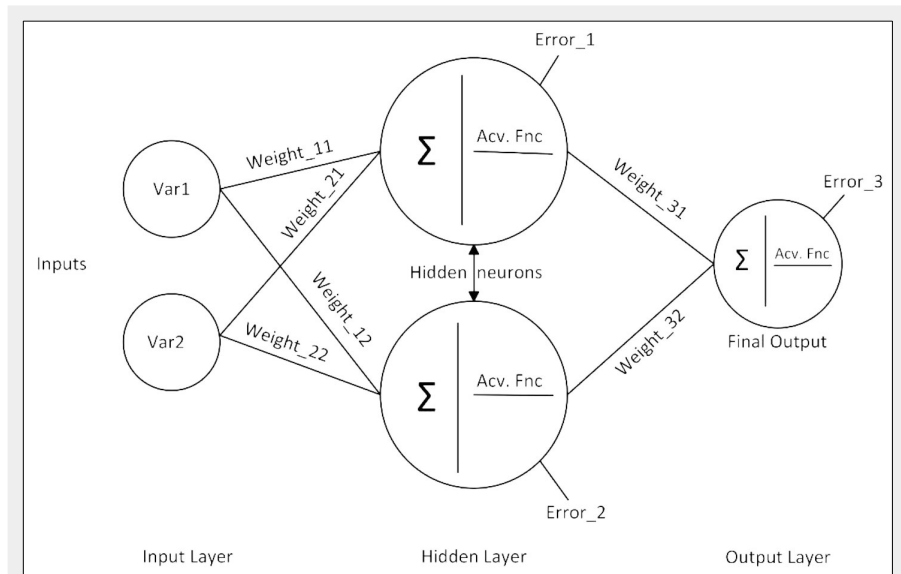
#### Artificial neural network type variable-density yield models

It is crucial to construct an ANN model that is consistent with the principles of forest yield. If ANN's parameters are not properly managed, the produced ANN model will lead, in general, to unreasonable results (Hamidi et al. 2021). Therefore, this study developed a promising modeling framework to meet the principles of forest yield.

An ANN model includes three fully connected layers: the input layer, hidden layer, and output layer. Each input with an appropriate weight is taken into the input layer without any data processing. The weighted inputs with added bias values are passed to the hidden layer through an activation function. Then, the resulting data are propagated forward to the next layer. This process is continued until the final output has been produced.

The various types of activation functions are selected according to the type of problem. The sigmoid (i.e., logistic and hyperbolic tangent functions) and linear activation functions seem to be suitable for biological relationships in the hidden layer and output layer, respectively (Ozçelik et al. 2010, Ashraf et al. 2013). The activation functions may significantly influence the performance of ANN models. The curve of the logistic function may reach the asymptote more quickly than the curve of the hyperbolic tangent function at higher and lower values (Dubey et al. 2021). This property of the logistic sigmoid function may cause the poor performance of ANN models due to the lack of sensitivity to variability in the observed data. In this study, the hyperbolic tangent and linear activation functions were selected at the hidden layer and output layer, respectively.

The inputs can be normalized to reduce the network complexity and improve the robustness of the network against outliers (Akilli & Hülya 2020). There are a number of normalization techniques, such as the z-score, min-max, and sigmoid techniques. In this study, the min-max technique was used to normalize inputs within a range of -1 to +1 because the regularization function



**Fig. 2** - A simple representation of ANN model with one hidden layer with two hidden neurons (or nodes). This network includes a total of nine parameters (six weight and three error parameters).

used worked well with this type of normalization (Foresee & Hagan 1997).

The cross-validation methods such as leave-one-out and *k*-fold cross-validation techniques are effective to evaluate the performance of an ANN model (Diamantopoulou & Ozçelik 2012). This study used the *k*-fold cross-validation method. In this method, dataset is randomly divided into *k* subsets having roughly equal size. The network is trained using *k*-1 subsets and tested using a single subset. This process is repeated *k* times and calculated mean bias of *k* subsets. When the bias is acceptable, the training will be stopped; otherwise, the training will continue until achieving minimum bias by simultaneously adjusting network parameters and *k* value. This study also considered the principles of forest yield when deciding the best combination of network parameters. In pre-analysis, *k* value was determined as five for this study.

Overfitting is a serious problem for ANN models. If an ANN model having good fit in the training dataset fails in the testing dataset, it means that this model suffers extensively from the overfitting problem. In general, this problem is inevitable for ANN models having a small number of observations. On the other hand, if the number of parameters in an ANN model is higher than the number of observations, the ANN model does not adequately fit the training dataset (Okut 2016). Because there was a total of 180 observations for each stand variable in this study, simple neural network models were implemented to prevent overfitting. The method of determining the number of parameters in an ANN model is shown in Fig. 2.

Another way to avoid overfitting is to select a suitable regularization training function, such as the Levenberg-Marquardt,

Bayesian, or variable learning rate gradient descent function. In this study, the Bayesian function was chosen because it can capture nonlinear patterns in the dataset and produce an S-shaped curve. This technique is also quite robust for overfitting (Skudnik & Jevšenak 2022).

As with the regression approach, the initial objective in an ANN model is to minimize errors (eqn. 7). This study aimed to minimize the mean of squared errors (also called the objective function). Different performance functions can be used as the objective function, such as the mean absolute error. In this study, the overall prediction performance of the ANN models was measured by the mean squared error, which is an indicator of the variation in the errors (eqn. 7):

$$E_D(D|Cw, M) = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2 \quad (7)$$

where  $O_i$  is the measured target values,  $P_i$  is the predicted outputs, and  $n$  is the number of observations;  $D$  is a training data set, and  $M$  is a given network architecture.

In the Bayesian regularization technique, an additional term is added to the objective function (eqn. 8), and this allows the penalization of large weights and thereby improves the generalization ability of the ANN models (Foresee & Hagan 1997).

$$F = \beta E_D(D|w, M) + \alpha E_w(w|CM) \quad (8)$$

where  $E_w$  is the sum of squared weights and  $\alpha$  and  $\beta$  are the parameters of objective function (also called the regularization parameters).

The regularization parameters  $\alpha$  and  $\beta$  are adaptively estimated for determining the best weights based on the Bayesian learn-

**Tab. 2** - Coefficient estimates (with *t*- and *p*-values, and standard error, SE) and goodness-of-fit statistics of nonlinear regression models developed for predicting basal area (*G*, m<sup>2</sup> ha<sup>-1</sup>), volume (*V*, m<sup>3</sup> ha<sup>-1</sup>), and the number of trees (*N*, number ha<sup>-1</sup>) of Crimean pine stands.

Model	Coefficient	Estimate	SE	<i>t</i> -value	<i>p</i> -value	<i>R</i> <sub>adj</sub> <sup>2</sup>	RMSE
<i>G</i>	<i>b</i> <sub>1</sub>	12.542	0.252	49.691	<0.0001	0.96	3.24
	<i>a</i> <sub>1</sub>	1.926	0.18	10.726	<0.0001		
	<i>a</i> <sub>2</sub>	0.527	0.04	13.257	<0.0001		
	<i>a</i> <sub>3</sub>	0.00004	0	4.0689	<0.0001		
	<i>a</i> <sub>4</sub>	-0.0005	0	9.7649	<0.0001		
<i>V</i>	<i>b</i> <sub>3</sub>	-21.846	2.471	8.8406	<0.0001	0.8	108.15
	<i>b</i> <sub>1</sub>	19.732	0.808	24.412	<0.0001		
	<i>a</i> <sub>1</sub>	2.214	0.382	5.79	<0.0001		
	<i>a</i> <sub>2</sub>	0.129	0.054	2.4	<0.0001		
	<i>a</i> <sub>3</sub>	0.001	0.001	1.832	<0.0001		
<i>N</i>	<i>a</i> <sub>4</sub>	-0.005	0.001	3.916	<0.0001	0.85	144.18
	<i>b</i> <sub>3</sub>	-54.144	16.15	3.352	<0.0001		
	<i>b</i> <sub>1</sub>	9.565	0.427	22.388	<0.0001		
	<i>a</i> <sub>1</sub>	1.111	0.325	3.415	<0.0001		
	<i>a</i> <sub>2</sub>	0.662	0.165	4.004	<0.0001		
	<i>a</i> <sub>3</sub>	0.111	0.026	4.266	<0.0001		
	<i>a</i> <sub>4</sub>	0.34	0.058	5.857	<0.0001		
<i>b</i> <sub>3</sub>	67.998	6.493	10.472	<0.0001			

ulation data divided into 10-year periods. Each site index class (i.e., 12, 22, and 32 m) included different stand density ratios (3, 5, 7, and 9; a stand density ratio of 9 means that the stand is fully stocked) and ages. After those new inputs were transformed based on the inputs of each ANN model, the testing process began.

**Model evaluation**

The statistical evaluation of the regression and ANN models was done using the adjusted coefficient of determination (*R*<sub>adj</sub><sup>2</sup>, eqn. 10) and root mean square error (RMSE, eqn. 11):

$$R_{adj}^2 = 1 - \left( 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - O_{mean})^2} \right) \left( \frac{n-1}{n+k-1} \right) \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n-k}} \quad (11)$$

where *O<sub>i</sub>* is a measured value, *P<sub>i</sub>* is a predicted value, *O<sub>mean</sub>* is a mean of the measured values, *n* is the number of measurements, and *k* is the number of coefficients (parameters for the ANN model). In calculating the *R*<sub>adj</sub><sup>2</sup> and RMSE, the inputs ln(*V*), ln(*G*), and ln(*N*) were back transformed to make the statistical evaluations more meaningful.

**Results**

The goodness-of-fit statistics, estimated coefficients, and *t*- and *p*-values of the nonlinear regression models developed for each stand variable are shown in Tab. 2. All of the coefficients estimated for *SI*, *SD*, and *A* and the other explanatory variables derived from *SI*, *SD*, and *A* were significant at the 5% level. The *R*<sub>adj</sub><sup>2</sup> values showed that the amounts of the variations explained by the nonlinear regression models (80% to 96%) were satisfactory for *G*, *V*, and *N*.

The goodness-of-fit statistics and network parameters of the best ANN models for each stand variable are shown in Tab. 3. The ANN models successfully described the variations in *G*, *V*, and *N* (the *R*<sub>adj</sub><sup>2</sup> value ranged from 0.86 to 0.97). Therefore, we concluded that simple ANN models having one hidden layer with a small number of hidden neurons (*n*<10) were adequate for developing the variable-density yield models.

The number of parameters in the ANN models (namely, the coefficients for the regression models) was higher than for the nonlinear regression models. However, both of them provided a similar statistical performance for *G* and *N*. The ANN model was better than the regression model in predicting *V*. Similarly, both of them yielded polymorphic curves with a variable asymptote (Fig. 3a-f, Fig. 4a-f). These results were compatible with the principles of forest yield. In addition, it appeared that the ANN models were better for reflecting the patterns of *G*, *V*, and *N* when relying on *SI* and *SD* than the nonlinear regression models.

ing rule and, hence, minimizing the regularized objective function (eqn. 9):

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)}$$

where *P*(*w* | *D*, *α*, *β*, *M*), *P*(*w* | *β*, *M*) and *P*(*w* | *α*, *M*) are the posterior probability, the likelihood function, and the prior distribution for network weights, respectively. *P*(*D* | *α*, *β*, *M*) is a normalization factor for regularization parameters.

Choosing the best number of hidden neurons, learning rate, and momentum factor is important for enhancing the learning capacity of ANN models. When many hidden neurons are used, ANN models tend to

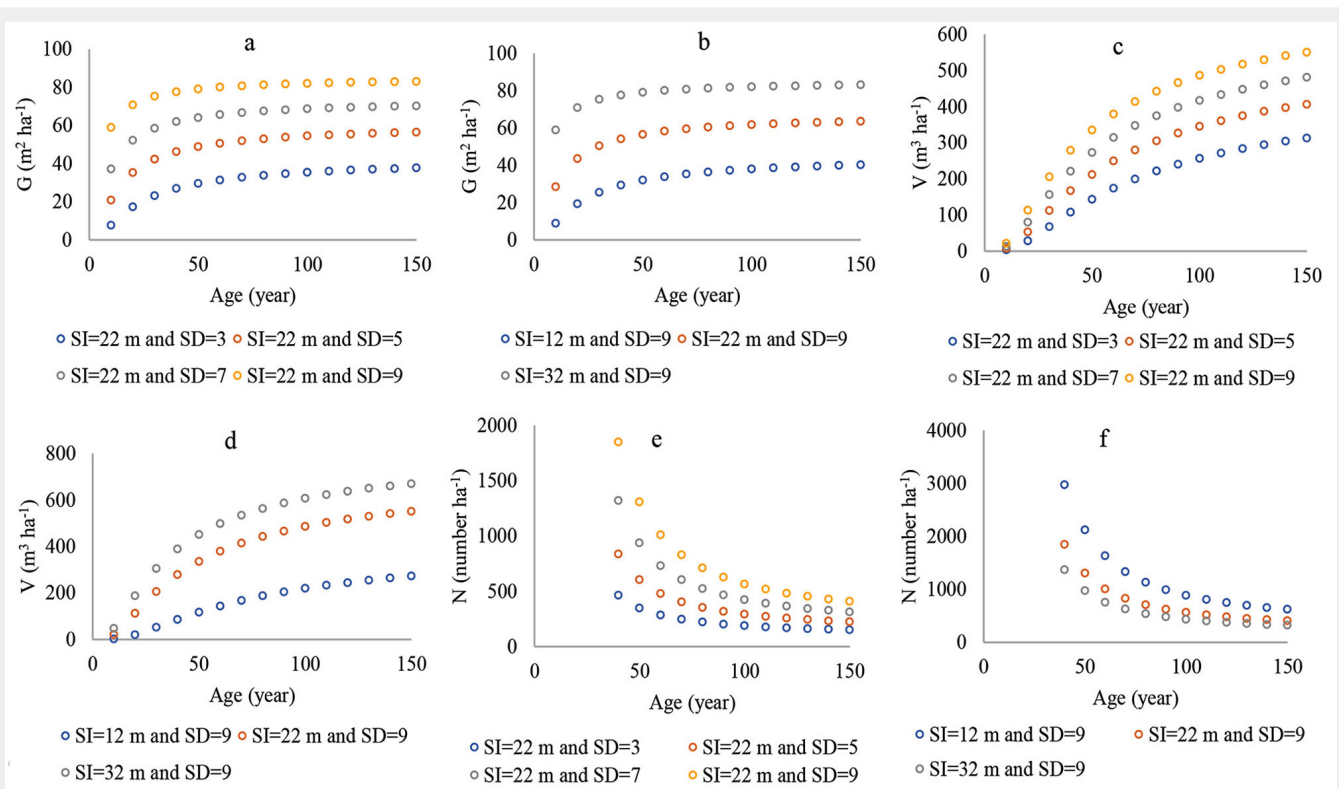
memorize the problem but do not generalize the relationships between inputs and output. Conversely, if only a few hidden neurons are used, ANN models do not reflect the patterns of the dataset (Ozcelik et al. 2013). While the learning rate controls the speed of learning, the momentum factor adjusts the current weight changes depending on the previous weight changes.

In order to achieve the optimum number of hidden neurons, learning rate, and momentum value, different combinations of these factors were studied. In this trial-and-error process, the number of hidden neurons varied from 2 to 10 and the learning rate and momentum factor varied from 0.1 to 0.9.

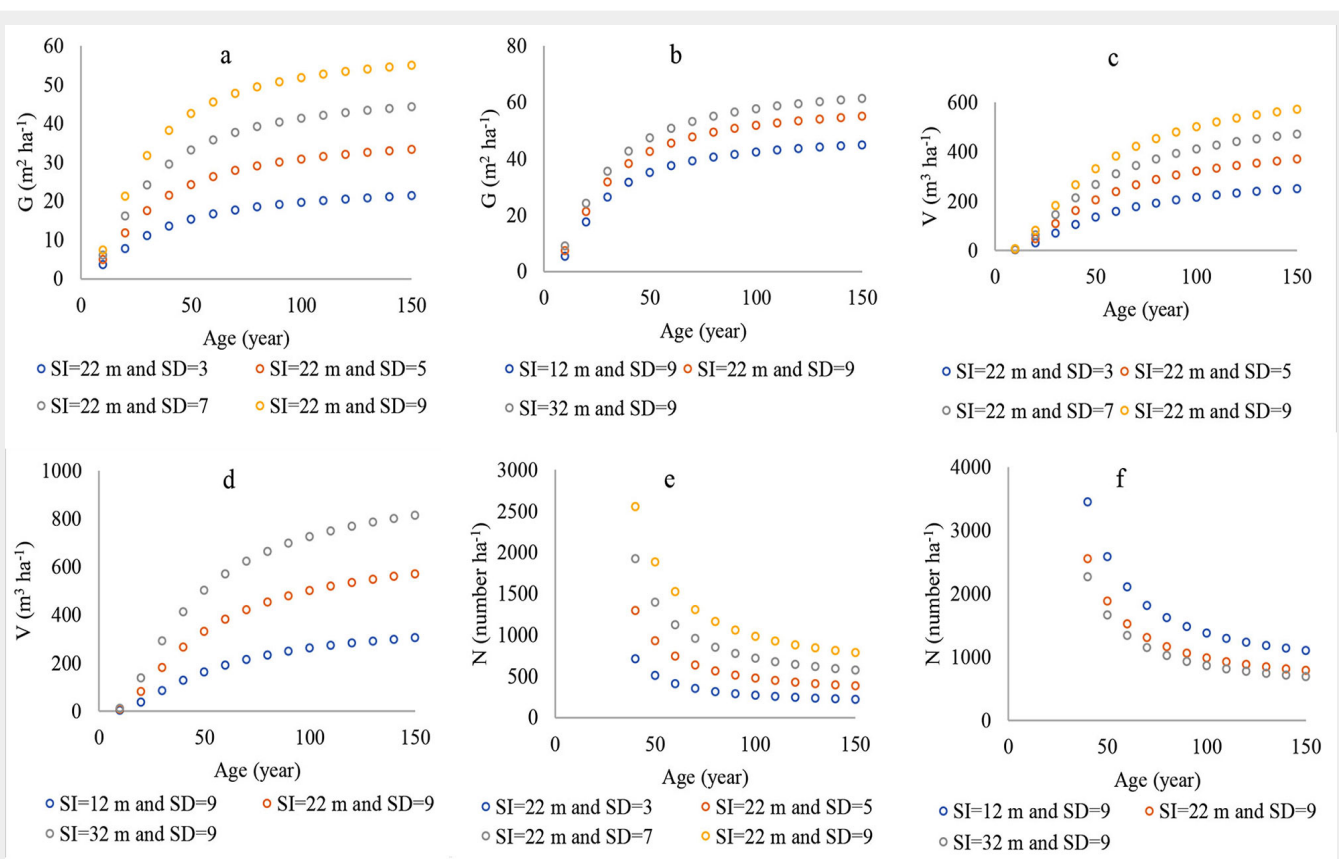
In this study, a total of 180 data items for each stand variable were used to evaluate the performance of the ANN models. To promote the performance of the ANN models and avoid overfitting, the five-fold cross-validation was employed using the entire dataset. The ANN models were conceived using the MATLAB® software (MathWorks 2020). Inputs and output were the independent and dependent variables of the modified Gompertz function, respectively. Specifically, the inputs were: [1/*SI*, 1/*SD*, *SI*<sup>2</sup>, *SD*<sup>2</sup>, and 1/*A*], [ln(*SI*), 1/*SD*, *SI*<sup>2</sup>, *SD*<sup>2</sup>, and 1/*A*], and [1/*SI*, 1/*SD*, ln(*SI*), ln(*SD*) and 1/*A*] for outputs ln(*G*), ln(*V*) and ln(*N*), respectively. Due to the small number of observations, the entire dataset was used for training the ANN models in this study. Therefore, the suitability of the ANN models for the forest yield principles was assessed using the artificially generated data, which consisted of 150 years' worth of sim-

**Tab. 3** - The number of parameters, learning rate, momentum, and goodness-of-fit statistics of ANN-based variable-density yield models developed for predicting basal area (*G*, m<sup>2</sup> ha<sup>-1</sup>), volume (*V*, m<sup>3</sup> ha<sup>-1</sup>), and stem number (*N*, number ha<sup>-1</sup>) of Crimean pine stands.

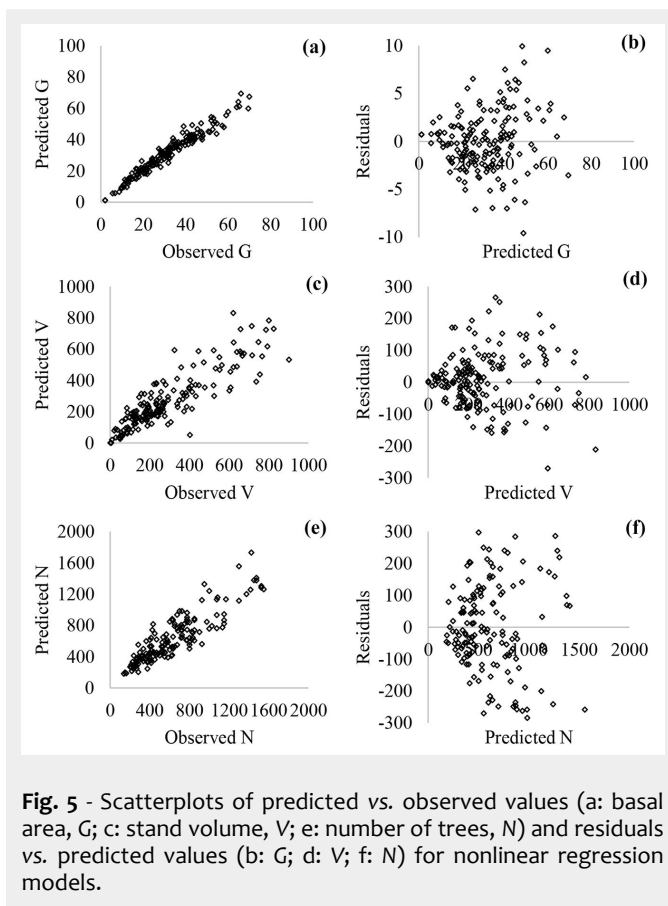
Parameter	<i>G</i>	<i>V</i>	<i>N</i>
Number of inputs	5	5	5
Number of network parameters	57	36	50
Number of hidden neurons	8	5	7
Learning rate	0.1	0.1	0.1
Momentum value	0.5	0.5	0.5
<i>R</i> <sub>adj</sub> <sup>2</sup>	0.97	0.97	0.86
RMSE	2.77	35.79	143.87



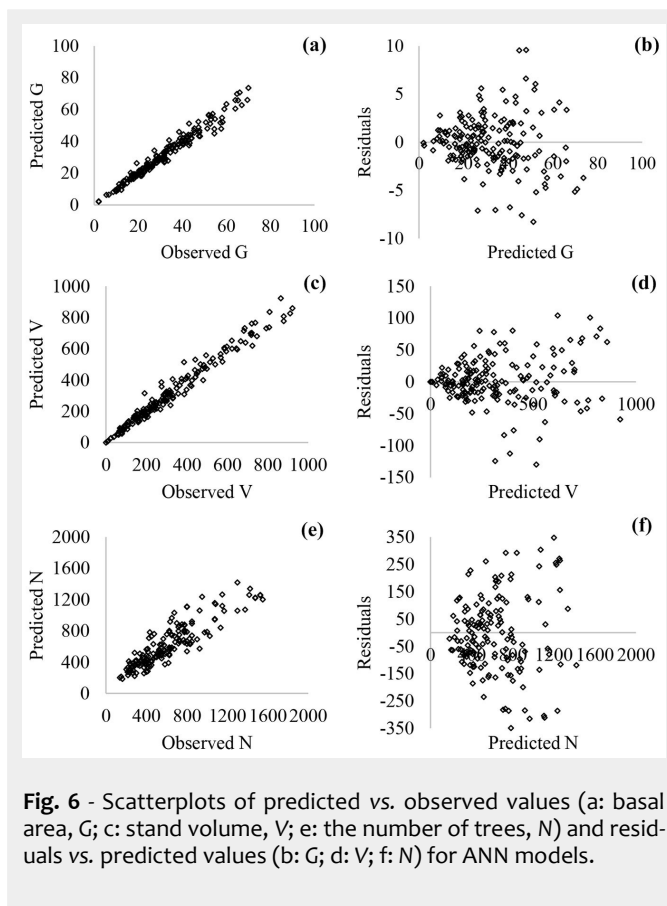
**Fig. 3** - Curves of nonlinear regression models for (a) G, basal area, (c) V, stand volume, and (e) N, number of trees, based on different stand densities within same site index; (b) basal area, (d) stand volume, and (f) number of trees based on different site indexes within same stand density.



**Fig. 4** - Curves of ANN models for (a) G, basal area, (c) V, stand volume, and (e) N, number of trees, based on different stand densities within same site index; (b) basal area, (d) volume and (f) the number of trees based on different site indexes within same stand density.



**Fig. 5** - Scatterplots of predicted vs. observed values (a: basal area, G; c: stand volume, V; e: number of trees, N) and residuals vs. predicted values (b: G; d: V; f: N) for nonlinear regression models.



**Fig. 6** - Scatterplots of predicted vs. observed values (a: basal area, G; c: stand volume, V; e: the number of trees, N) and residuals vs. predicted values (b: G; d: V; f: N) for ANN models.

According to Fig. 5 and Fig. 6, the non-constant variance in residuals of the nonlinear regression models was more prominent compared with the ANN models. In addition, both of the models had similar relationships between observed and predicted values (Fig. 5a, Fig. 5c, Fig. 5e and Fig. 6a, Fig. 6c, Fig. 6e).

### Discussion

In most studies using ANN models, the findings have been discussed only with the consideration of the goodness-of-fit statistics (Ashraf et al. 2013, Karatepe et al. 2022). In forestry modeling, the principles of forest growth and yield are extremely important (Vanclay & Skovsgaard 1997). Therefore, this study took into account both the goodness-of-fit statistics and the principles of forest yield and contrasted the ANN models accordingly. The ANN models developed showed reasonable statistical and biological behaviors in predicting G, V, and N. With this in mind, this study presented important findings for ANN modeling in forestry.

The Gompertz function was modified to develop the dynamic variable-density yield models in this study. The variables *SI* and *SD* were added to the growth-rate parameter of the Gompertz function, thereby the principles of forest yield could be met. A similar approach was used in the study of Temesgen et al. (2014). In their study, the Chapman-Richards function was expanded to incorporate stand and tree variables to improve height predictions in forests with

multiple species and multiple layers. Likewise, Fortin et al. (2007) modified the von Bertalanffy-Richards function to satisfy the expected biological patterns of basal area growth of red spruce and balsam fir after logging treatments.

ANN models were more dynamic than nonlinear regression models in describing patterns of G, V, and N. This result is attributable to the nonlinear transfer function in the hidden layer of ANN models (Jensen et al. 1999). Zhang et al. (2000) demonstrated that ANN models facilitated the modeling of complex nonlinear relationships between tree growth and climate variables. Guan et al. (1997) showed that the ANN model had great potential for constructing a dynamic model requiring a higher-order approximation.

Developing a dynamic nonlinear regression model can be difficult due to the difficulty in determining the starting values of the parameters (Fekedulegn et al. 1999). Developing an ANN model with many explanatory variables can be easier than developing a nonlinear regression model. On the other hand, the correlations among explanatory variables may affect the sign and magnitude of the regression parameters to be estimated and may result in unreasonable statistical and biological behaviors. Similarly, some irrelevant explanatory variables (inputs) may decrease the accuracy of ANN models or may induce small changes in their performance (Maier et al. 1998). Therefore, the selection of the vari-

ables is essential for both modeling approaches.

The accuracy of ANN models is highly dependent on the selection of network parameters (Maier & Dandy 2000). The choice of network parameters was reported in only a few studies. In the study of Liu & Zhang (2005), while the learning rate was 0.3, the momentum was 0.7 in predicting the lumber volume, lumber value, chip volume, and total product value. In the study of Ozçelik et al. (2013), the learning rate was 0.1 and the momentum was 0.3 in predicting the total tree heights. Castaño-Santamaría et al. (2013) found that the best values for the learning rate and the momentum were 0.1 and 0.2, respectively, in predicting tree heights. Ozçelik et al. (2014) found that the optimal values for the learning rate and the momentum were 0.1 and 0.3, respectively, in predicting tree stem volumes. In our study, the optimal learning rate and momentum were 0.1 and 0.5, respectively, in predicting the G, V, and N.

This study demonstrated that simple ANN models with one hidden layer and only a few hidden neurons ( $n < 10$ ) were adequate for accurately predicting the G, V, and N. These findings are compatible with the results of Castaño-Santamaría et al. (2013), Liu & Zhang (2005), and Ozçelik et al. (2013), who reported that one hidden layer and a small number of hidden neurons were adequate for the prediction of forest attributes.

## Conclusions

Many researchers paid attention to the goodness-of-fit statistics in evaluating ANN models and comparing ANN models with other modeling approaches. However, the biological rationales are often overlooked. This study showed that two different modeling approaches with similar error statistics may provide different curves depending on the *SI* and *SD* for *G*, *V*, and *N*. Therefore, it is recommended that both statistical and biological behaviors should be examined together when comparing different modeling approaches. In addition, we analyzed how the choice of network parameters affected the performance of ANN models. The best results were achieved using ANN models with one hidden layer and only a few hidden neurons. The Bayesian regularization training function with the tangent-sigmoid and linear activation functions at the hidden layer and the output layer, respectively, was suitable for meeting the principles of forest yield. The best learning rate and momentum were 0.1 and 0.5, respectively, for the prediction of *G*, *V*, and *N*. The knowledge obtained from this study can be a guideline for future similar studies.

## List of abbreviations

*A*: Stand age; *ANN*: Artificial Neural Network; *G*: Basal area per hectare; *N*: The number of trees per hectare; *SD*: Stand density; *SI*: Site index; *V*: Stand volume per hectare

## Acknowledgements

The authors thank the General Directorate of Forestry for the legal permission of field studies, as well as Dr. Sinan Bulut, forest engineer Ömer Bulut and Kutay Kaya for their help in collecting data.

## Funding

This study was a part of a PhD thesis supported financially by the Scientific Research Project Unit of Çankiri Karatekin University, Turkey (grant: OFo61218D05) and the Turkish Scientific and Technical Research Institute (TUBITAK, grant: TOVAG-119O061).

## Author contributions

FB completed the statistical analyses and wrote the manuscript draft. IE and AG contributed to the idea, reviewed and edited the manuscript. All authors have read and approved the final manuscript.

## References

Akilli A, Hülya A (2020). Evaluation of normalization techniques on neural networks for the prediction of 305-day milk yield. *Turkish Journal of Agricultural Engineering Research* 1: 354-367. - doi: [10.46592/turkager.2020.v0i02.011](https://doi.org/10.46592/turkager.2020.v0i02.011)

Anonymous (2017). Ekosistem tabanlı fonksiyonel orman amenajman planlarının düzenlenmesine ait usul ve esaslar [Principles and procedures for the regulation of ecosystem-based functional forest management plans]. Orman

ve Su İşleri Bakanlığı, Orman Genel Müdürlüğü, Orman Amenajmanı ve Planlama Dairesi Başkanlığı, Ankara, Teblig 299: 214 [in Turkish]

Ashraf MI, Zhao Z, Bourque CP-A, MacLean DA, Meng F-R (2013). Integrating biophysical controls in forest growth and yield predictions with artificial intelligence technology. *Canadian Journal of Forest Research* 43: 1162-1171. - doi: [10.1139/cjfr-2013-0090](https://doi.org/10.1139/cjfr-2013-0090)

Bolat F (2021). Ankara Orman Bölge Müdürlüğü Anadolu karaçami meşcerelerinde artım ve büyümenin yapay sinir ağları ile modellenmesi [Modeling forest growth and yield in Anatolian black pine stands located Ankara Forest District Directorate by artificial neural network]. PhD Thesis, Çankiri Karatekin Üniversitesi, Fen Bilimleri Enstitüsü, Turkey, pp. 213.

Burkhardt HE, Tomé M (2012). Modeling forest trees and stands. Springer Science+Business Media, Dordrecht, Netherlands, pp. 458. [online] URL: <http://books.google.com/books?id=XmcNcfEarsoC>

Castaña-Santamaría J, Crecente-Campo F, Fernández-Martínez JL, Barrio-Anta M, Obeso JR (2013). Tree height prediction approaches for uneven-aged beech forests in northwestern Spain. *Forest Ecology and Management* 307: 63-73. - doi: [10.1016/j.foreco.2013.07.014](https://doi.org/10.1016/j.foreco.2013.07.014)

Che S, Tan X, Xiang C, Sun J, Hu X, Zhang X, Duan A, Zhang J (2019). Stand basal area modelling for Chinese fir plantations using an artificial neural network model. *Journal of Forestry Research* 30: 1641-1649. - doi: [10.1007/s11676-018-0711-9](https://doi.org/10.1007/s11676-018-0711-9)

Chenge IB (2021). Height-diameter relationship of trees in Omo strict nature forest reserve, Nigeria. *Trees, Forests and People* 3 (2019): 100051. - doi: [10.1016/j.tfp.2020.100051](https://doi.org/10.1016/j.tfp.2020.100051)

Curtis RO (1982). A simple index of stand density for Douglas-fir. *Forest Science* 28: 92-94. [online] URL: <http://academic.oup.com/forests/science/article-abstract/28/1/92/4656511>

Diamantopoulou MJ, Özçelik 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)

Dubey SR, Singh SK, Chaudhuri BB (2021). A comprehensive survey and performance analysis of activation functions in deep learning. *ArXiv Preprint* 2109: 14545. - doi: [10.48550/arXiv.2109.14545](https://doi.org/10.48550/arXiv.2109.14545)

Fang Z, Bailey RL (2001). Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *Forest Science* 47: 287-300. [online] URL: <http://academic.oup.com/forests/science/article/47/3/287/4617175>

Fekedulegn D, Mac Siúrtáin MP, Colbert JJ (1999). Parameter estimation of nonlinear models in forestry. *Silva Fennica* 33: 327-336. - doi: [10.14214/sf.653](https://doi.org/10.14214/sf.653)

Foody GM, Lucas RM, Curran PJ, Honzak M (1997). Mapping tropical forest fractional cover from coarse spatial resolution remote sensing imagery. *Plant Ecology* 131: 143-154. - doi: [10.1023/A:1009775619936](https://doi.org/10.1023/A:1009775619936)

Fooresee FD, Hagan MT (1997). Gauss-Newton approximation to Bayesian learning. In: Proceedings of "International Conference on Neural Networks (ICNN'97)". Houston (TX, USA) 12

June 1997. IEEEXplore, pp. 1930-1935. - doi: [10.1109/ICNN.1997.614194](https://doi.org/10.1109/ICNN.1997.614194)

Fortin M, Daigle G, Ung C-H, Bégin J, Archambault L (2007). A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling. *European Journal of Forest Research* 126: 573-585. - doi: [10.1007/s10342-007-0179-1](https://doi.org/10.1007/s10342-007-0179-1)

Fox JC, Ades PK, Bi H (2001). Stochastic structure and individual-tree growth models. *Forest Ecology and Management* 154: 261-276. - doi: [10.1016/S0378-1127\(00\)00632-0](https://doi.org/10.1016/S0378-1127(00)00632-0)

GDF (2020). Forestry statistics in Turkey. Ministry of Agriculture and Forestry, General Directorate of Forestry, Forest Management and Planning Department, Ankara, Turkey. [In Turkish] [online] URL: <http://www.ogm.gov.tr/tr/e-kutuphane/resmi-istatistikler>

Gilbert H, Manning PJ, McDill ME, Sterner S (2010). Sawtimber yield tables for Pennsylvania forest management planning. *Northern Journal of Applied Forestry* 27: 140-150. - doi: [10.1093/njaf/27.4.140](https://doi.org/10.1093/njaf/27.4.140)

Guan BT, Gertner GZ, Parysow P (1997). A framework for uncertainty assessment of mechanistic forest growth models: a neural network example. *Ecological Modelling* 98: 47-58. - doi: [10.1016/S0304-3800\(96\)01936-9](https://doi.org/10.1016/S0304-3800(96)01936-9)

Hamidi SK, Weiskittel A, Bayat M, Fallah A (2021). Development of individual tree growth and yield model across multiple contrasting species using nonparametric and parametric methods in the Hyrcanian forests of northern Iran. *European Journal of Forest Research* 140: 421-434. - doi: [10.1007/s10342-020-01340-1](https://doi.org/10.1007/s10342-020-01340-1)

Hirigoyen A, Navarro-Cerrillo R, Bagnara M, Franco J, Requin F, Rachid-Casnati C (2021). Modeling taper and stem volume considering stand density in *Eucalyptus grandis* and *Eucalyptus dunnii*. *iForest - Biogeosciences and Forestry* 14 (2): 127-136. - doi: [10.3832/ifor3604-014](https://doi.org/10.3832/ifor3604-014)

Jensen JR, Qiu F, Ji M (1999). Predictive modeling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *International Journal of Remote Sensing* 20: 2805-2822. - doi: [10.1080/014311699211804](https://doi.org/10.1080/014311699211804)

Kalipsız A (1963). Türkiye'de Karaçam mescerelerinin tabii bünyesi ve verim kudreti üzerine araştırmalar [Studies on the natural structure and yield potential of Black pine stands in Turkey]. Tarım Bakanlığı/Orman Genel Müdürlüğü, İstanbul, Turkey, vol. 8, pp. 138. [In Turkish]

Kanellopoulos I, Wilkinson GG, Megier J (1993). Integration of neural network and statistical image classification for land cover mapping. In: Proceedings of "IEEE International Geoscience and Remote Sensing Symposium - IGARSS '93". Tokyo (Japan) 18-21 Aug 1993. IEEEXplore, vol. 2, pp. 511-513. - doi: [10.1109/IGARSS.1993.322597](https://doi.org/10.1109/IGARSS.1993.322597)

Karatepe Y, Diamantopoulou MJ, Özçelik R, Sürücü Z (2022). Total tree height predictions via parametric and artificial neural network modeling approaches. *iForest - Biogeosciences and Forestry* 15 (2): 95-105. - doi: [10.3832/ifor3990-015](https://doi.org/10.3832/ifor3990-015)

Lima MBO, Lustosa Junior IM, Oliveira EM, Ferreira JCB, Soares KL, Miguel EP (2017). Artificial neural networks in whole-stand level modeling of *Eucalyptus* plants. *African Journal of Agricul-*

- tural Research 12: 535-541. - doi: [10.5897/AJAR2016.11607](https://doi.org/10.5897/AJAR2016.11607)
- Liu C, Zhang SY (2005). Models for predicting product recovery using selected tree characteristics of black spruce. *Canadian Journal of Forest Research* 35: 930-937. - doi: [10.1139/x05-025](https://doi.org/10.1139/x05-025)
- Maier HR, Dandy GC, Burch MD (1998). Use of artificial neural networks for modelling cyanobacteria *Anabaena* spp. in the River Murray, South Australia. *Ecological Modelling* 105: 257-272. - doi: [10.1016/S0304-3800\(97\)00161-0](https://doi.org/10.1016/S0304-3800(97)00161-0)
- Maier HR, Dandy GC (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling and Software* 15: 101-124. - doi: [10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- MathWorks (2020). MATLAB User's guide, Version 2020b. The MathWorks, Inc., Natick, MA, USA.
- Meyer WH (1938). Yield of even-aged stands of Ponderosa pine. Technical Bulletin no. 630, USDA, Washington, DC, USA, pp. 60. [online] URL: <http://books.google.com/books?id=jjXoGnAaB4C>
- Ogana FN, Corral-Rivas S, Gorgoso-Varela JJ (2020). Nonlinear mixed-effect height-diameter model for *Pinus pinaster* Ait. and *Pinus radiata* D. Don. *Cerne* 26: 150-161. - doi: [10.1590/01047760202026012695](https://doi.org/10.1590/01047760202026012695)
- Okut H (2016). Bayesian regularized neural networks for small n big p data. In: "Artificial Neural Networks - Models and Applications". In: tech, London, UK, 28-48. - doi: [10.5772/63256](https://doi.org/10.5772/63256)
- Ozçelik R, Diamantopoulou MJ, Brooks JR, Wiant HV (2010). Estimating tree bole volume using artificial neural network models for four species in Turkey. *Journal of Environmental Management* 91: 742-753. - doi: [10.1016/j.jenvman.2009.10.002](https://doi.org/10.1016/j.jenvman.2009.10.002)
- 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/ifor0878-007](https://doi.org/10.3832/ifor0878-007)
- 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)
- Seki M, Sakici OE (2022). Ecoregional variation of Crimean Pine (*Pinus nigra* subspecies *pallasiana* [Lamb.] Holmboe) stand growth. *Forest Science* 68 (5-6): 452-463. - doi: [10.1093/forsci/afx030](https://doi.org/10.1093/forsci/afx030)
- Skudnik M, Jevšenak J (2022). Artificial neural networks as an alternative method to nonlinear mixed-effects models for tree height predictions. *Forest Ecology and Management* 507 (1): 120017. - doi: [10.1016/j.foreco.2022.120017](https://doi.org/10.1016/j.foreco.2022.120017)
- Temesgen H, Zhang CH, Zhao XH (2014). Modelling tree height-diameter relationships in multi-species and multi-layered forests: a large observational study from Northeast China. *Forest Ecology and Management* 316: 78-89. - doi: [10.1016/j.foreco.2013.07.035](https://doi.org/10.1016/j.foreco.2013.07.035)
- TSMS (2022). Turkish State Meteorological Service. Web site. [online] URL: <http://www.mgm.gov.tr/eng/forecast-cities.aspx>
- Vanclay JK (1994). Modelling forest growth and yield: applications to mixed tropical forests. CAB International, Wallingford, UK, pp. 329.
- Vanclay JK, Skovsgaard JP (1997). Evaluating forest growth models. *Ecological Modelling* 98: 1-12. - doi: [10.1016/S0304-3800\(96\)01932-1](https://doi.org/10.1016/S0304-3800(96)01932-1)
- Wang M, Borders BE, Zhao D (2008). An empirical comparison of two subject-specific approaches to dominant heights modeling: The dummy variable method and the mixed model method. *Forest Ecology and Management* 255: 2659-2669. - doi: [10.1016/j.foreco.2008.01.030](https://doi.org/10.1016/j.foreco.2008.01.030)
- Weiskittel AR, Hann DW, Jr JAK, Vanclay JK (2011). *Forest growth and yield modeling*. John Wiley and Sons, USA, pp. 431. - doi: [10.1002/9781119998518](https://doi.org/10.1002/9781119998518)
- West PW, Ratkowsky DA, Davis AW (1984). Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. *Forest Ecology and Management* 7: 207-224. - doi: [10.1016/0378-1127\(84\)90068-9](https://doi.org/10.1016/0378-1127(84)90068-9)
- Zhang Q-B, Hebda RJ, Zhang Q-J, Alfaro RI (2000). Modeling tree-ring growth responses to climatic variables using artificial neural networks. *Forest Science* 46: 229-239. [online] URL: <http://academic.oup.com/forestscience/article/46/2/229/4617154>
- Zuur AF, Ieno EN, Elphick CS (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1 (1): 3-14. - doi: [10.1111/j.2041-210X.2009.00001.x](https://doi.org/10.1111/j.2041-210X.2009.00001.x)

## Supplementary Material

**Appendix 1** - Syntax to facilitate the determination of initial values of a nonlinear function and the statistical significance of estimated parameters, based on the GenSA package in R.

Link: [Bolat\\_4116@suppl001.pdf](mailto: Bolat_4116@suppl001.pdf)