RESEARCH ARTICLE
Journal of Biometry Studies (2021) 1(1): 1-7
DOI: 10.29329/JofBS.2021.348.01

www.biometrystudies.com

Journal of Biometry Studies

Comparison of traditional regression models and artificial neural network models for height-diameter modeling in uneven-aged fir stands

Mustafa Kağan ÖZKAL[1,2,*], Davut ATAR[1,3], Mehmet AYDIN[1], Fatih TUNÇ[1]

[1]General Directorate of Forestry, Forest Management and Planning Department, Ankara/TURKEY
[2]Isparta University of Applied Sciences, Faculty of Forestry, Isparta/TURKEY
[3]Çankırı Karatekin University, Faculty of Forestry, Çankırı/TURKEY

[*]Corresponding author: mustafakaganozkal@ogm.gov.tr

Received: 15/06/2021, Accepted: 26/06/2021

Abstract

Forests have been constantly growing with their dynamic structure. In order for this dynamic structure must be managed based on sustainable perspective. Diameter, height, age, stand structure, etc. parameters are used in the inventory stage of the planning. The easiest to measure among these parameters is the diameter. Therefore, the developed models are usually aimed at reaching other forest parameters from the diameter. In this article, 9 different height-diameter models were fitted using regression models, and feed-forward backpropagation artificial neural network model methods for uneven-aged fir (*Abies nordmanniana* subsp. *equi-trojani*) stands in Kökez, Bolu region of Turkey. The models compared based on adjusted R^2, bias, absolute bias, and root mean square error (RMSE). It was observed that the best result was obtained from the artificial neural network model, and the worst result was obtained from the power model.

Keywords: Linear models, Nonlinear models, Artificial Neural Network models (ANN), Height-diameter models, Feed-forward backpropagation ANN

Please cite this article as follows:
Özkal, M. K., Atar, D., Aydın, M., & Tunç, F. (2021). Comparison of Traditional Regression Models and Artificial Neural Network Models for Height-Diameter Modeling in Uneven-aged Fir Stands. *Journal of Biometry Studies, 1*(1), 1-7. https://doi.org/10.29329/JofBS.2021.348.01

1. Introduction

Fir species including Kazdağı fir (*Abies nordmanniana* subsp. *equi-trojani*), East Blacksea fir (*Abies nordmanniana* subsp. *nordmanniana*), and Taurus fir (*Abies cilicica*) is in 7th place in the general ranking with its distribution area of 511.703 ha in the forests of Turkey. It covers 5.23% of the general forest area (GDF, 2020).

Some researchers are primarily aimed at effective planning which should be based on the continuity of benefiting from forests. Forests are one of the most important natural resources. They spread over very large areas and have a long development process compared to other natural resources. For these reasons, it is very important to operate forests in a planned manner. In planned forests, there is a need for models that serve the planner by revealing the developments related to stands (Eraslan & Şad, 1993; Mısır & Yavuz, 2013; Seki, 2020).

In recent years, Turkey has adopted planning system approaches based on ecosystem-based functional planning principles. In order to, making such plans, growth and yield models for each tree species are needed. Although Turkey is a very rich country in terms of tree species diversity and stand structure, it lacks basic substrates for the growth and development characteristics of tree species in different stand establishments. One of the most important elements of growth and yield models is the height-diameter (*h-d*) models. Height-diameter equations are used to estimate tree height with the help of the dbh. This type of information becomes even more important because the diversity of tree species is rich and the differences of habitats are intense. This information is necessary to make more accurate yield estimations for decision mechanisms in ecosystem-based functional planning studies (Özçelik & Çapar, 2014).

Tree dbh is the most easily measured parameter of the tree in forestry. Thus, other variables of trees that are difficult to measure are estimated using this variable. Many simulation programs are used to determine the growth of the stand and height estimation. The growth status of trees at a point or period can be revealed with the equations of the relationship between dbh and height. In addition, height-diameter (*h-d*) equations guide determining tree height and volume in a region or stand (Kalıpsız, 1984; Carus, 1998; Eler, 2003; Temesgen et al., 2007; Çatal, 2009; Çatal, 2012).

There is always a positive curvilinear relationship between diameter and height. Linear and nonlinear regression models are used to reveal this relationship. Linear models are used in studies where high precision prediction is not desired. On the other hand, nonlinear models, which are more flexible than linear models, are easier to apply to data and are used more frequently (Carus & Çatal, 2017).

It is possible to calculate dbh data from height using new technologies such as LiDAR. Diameter estimation models can be fitted from the height by developing linear regression, randomForest, and randomForest imputation models (Ozkal, 2017).

With the development of technology, artificial intelligence technologies are used in many areas, as well as in the field of forestry. One of the sub-branches of these artificial intelligence technologies is an artificial neural network (ANN). Especially in the forestry works, stem taper models (Sakici & Ozdemir, 2018; Schikowski et al., 2018; Socha et al., 2020), tree bole volume models (Diamantopoulou, 2006; Özçelik et al., 2010) and diameter height models (Özçelik et al., 2013; Thanh et al., 2019) are developed with artificial neural networks.

Viera et al. (2018) used ANN to estimate diameter and height on eucalyptus trees in Brazil. They used many activation functions of ANN. As a result of the models they built, they had very high R^2 and very low RMSE.

One of the most efficient machine learning techniques is artificial neural network which involves fewer statistical assumptions predicting independent variables (Mehrotra et al. 2000; Atar, 2016).

Ercanlı and Bolat (2017) stated that they developed a model about diameter distribution using artificial neural networks for uneven-aged pines of Kunduz, Samsun region. NARX network, the feed-forward backprop, Cascade Correlation, Layer Recurrent, and Elman backprop were used for multiple layer network structures. The data were obtained from 637 sample plots. The best result was observed from feed-forward backprop ANN model with 0.784 R^2, 27.28 RMSE and 34360.8 AIC.

Diamantopoulou (2012) measured the stump diameter, diameter at breast height, sawn-timber height and, merchantable height of 471 fir trees in Greece. The several models were created from Back Propagation Artificial Neural Network (BPANN), Cascade Correlation Artificial Neural Network (CCANN), and Generalized Regression Neural Network (GRNN) models. The result shows that the best two models were from BPANN and CCANN.

In this study, height-diameter models were fitted using regression models, and feed-forward backpropagation artificial neural network model methods for uneven-aged fir (*Abies nordmanniana* subsp. *equi-trojani*) stands in Kökez, Bolu region of Turkey.

2. Material and Methods

2.1. Study area

The study area is within the boundaries of Bolu province, Turkey. The study area is called the Kökez Sub-district Forestry Directorate which is affiliated to the Aladağ District Forestry Directorate of the Bolu Regional Directorate of Forestry in the administrative structure of the General Directorate of Forestry in Turkey. The study area is located in the Western Black Sea Region as a geographical region. According to the equator; it is located between 40° 37' 05" - 40° 42' 42" north latitudes. According to Greenwich; it is located between 31° 29' 26" - 31° 38' 16" east longitudes (Figure 1).

Its lowest points are the route of the Büyüksu Stream, located in the northern part, at an altitude of 700-710 m. The highest point of the region is Ayıkaya Hill, which is located in the south of the region (north of Hıdırseyhler Plateau). Its altitude is around 1620 m.

The vegetation types of the region consisting of trees, shrubs, and herbaceous plants are listed below within the framework of those that can be determined according to the observations made in the field. The main tree types of the area are fir (*Abies nordmanniana* subsp. *equi-trojani*), black pine (*Pinus nigra*), Scots pine (*Pinus sylvestris*), juniper (*Juniperus oxycedrus*), beech (*Fagus orientalis*), hornbeam (*Carpinus betulus*), alder (*Alnus glutinosa*), maple (*Acer* sp.), planetree (*Platanus* sp.), hazelnut (*Corylus avellana*), poplar (*Populus* sp.), sapless oak (*Quercus petraea*), Cyprus oak (*Quercus infectoria*), willow (*Salix* sp.), and rowan (*Sorbus* sp.).

2.2. Data

High forests are operated in two ways as even-aged and uneven-aged in Turkey. The basic elements used in regulating utilization are age, age class, area, volume, and increment in the even-aged forests. They are dbh, diameter class, number of trees, basal area, volume, and increment in the uneven-aged forests (Keleş & Bulut, 2014). The data were collected from the sample areas of the forest management plan which were measured at intervals of 150 m x 300 m forests of uneven-aged. We chose to study the height-diameter models of fir trees because the fir is the dominant coniferous species for Kökez region which was managed under the uneven-aged.

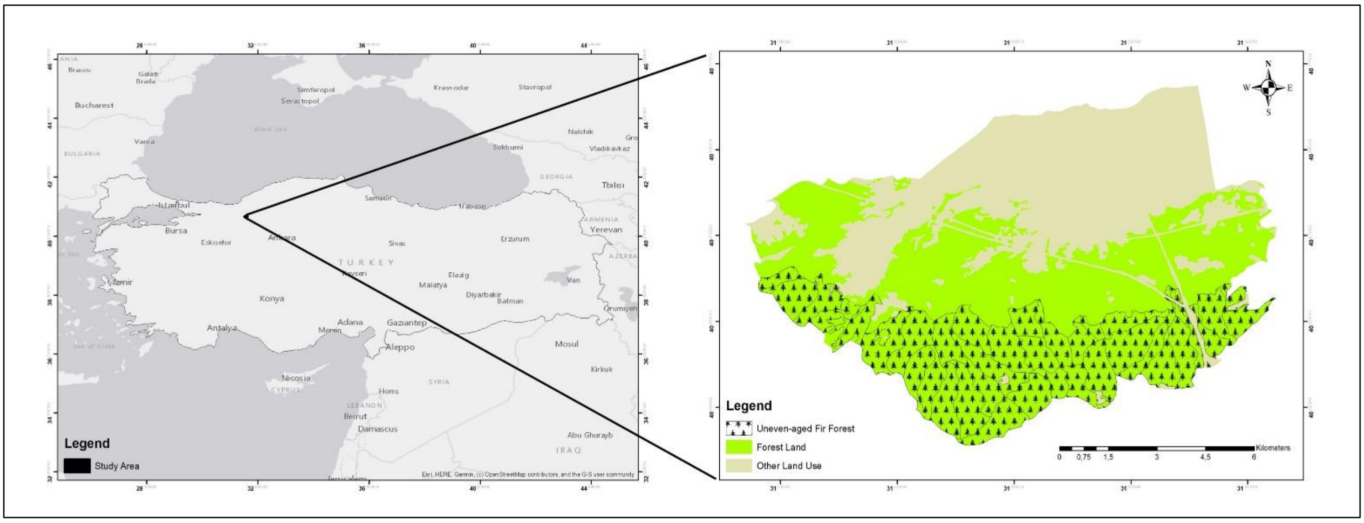

Figure 1. The location of the study area

Diameter at breast height (cm) and height (m) measurements of 862 trees were made in 309 sample plots area of 600 m^2. There was a requirement to measure the height at least two trees in each sample plot. Summary statistics including minimum, maximum, mean and standard deviation for h and dbh values of the sample trees are given in Table 1. Total data of the 862 sample trees were randomly divided into two groups: model development (80%, 633) and model evaluation (20%, 229) in R software (R Core Team, 2021) (Figure 2).

Table 1. Summary statistics of the sample trees

Data	Variable	n	Min	Max	Mean	Std. Dev.
Model development	dbh (cm)	633	8.1	97.8	34.7	14.8
	h (m)		4.1	38.1	20.2	6.9
Model evaluation	dbh (cm)	229	8.1	90.0	36.2	15.2
	h (m)		4.2	38.6	20.4	7.0

2.3. Candidate functions

The relationship between diameter and height has been tried to be explained by testing the models created with linear and non-linear equations before. (Huxley, 1932; Curtis, 1967; Korf, 1939; Meyer, 1940; Wykoff et al., 1982; Schumacher, 1939). In this study, we compared the models created using the artificial neural network and found the relationship between diameter and height to be appropriate at a sufficient confidence level were used. All models studied for this paper can be found in Table 2.

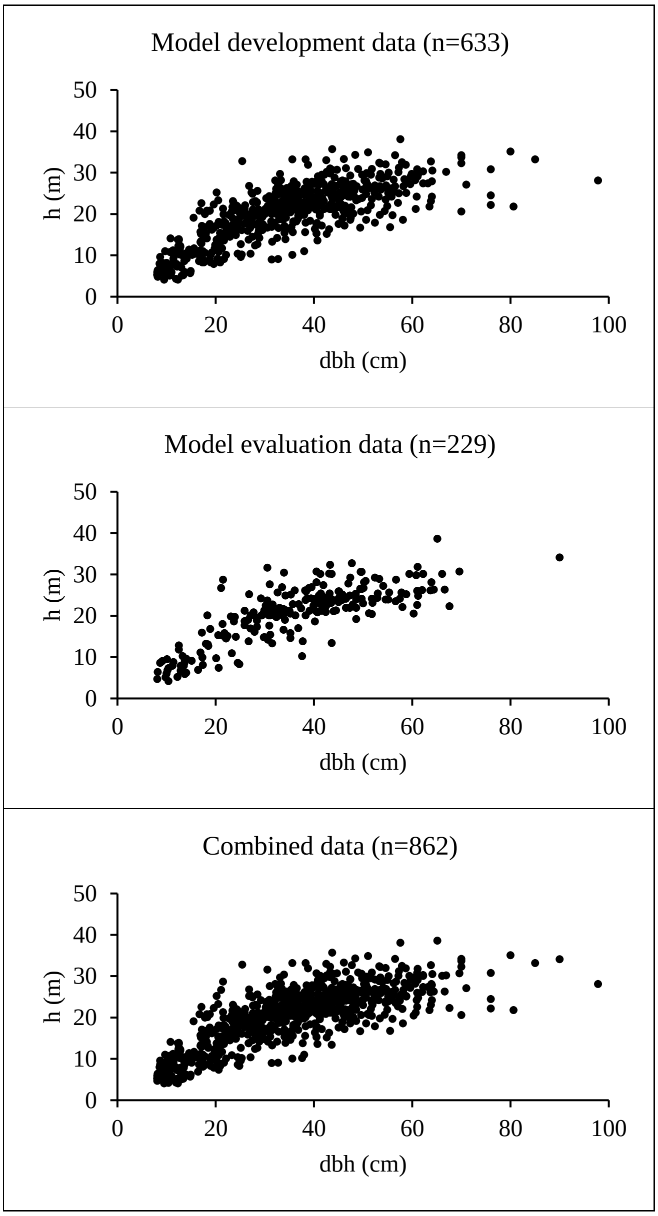

Figure 2. Plots of tree height against dbh

Table 2. Candidate linear and non-linear *h-d* models

No	Model	Function
M1	Linear	$h = b_0 + b_1 dbh$
M2	Power	$h = b_0 dbh^{b_1}$
M3	Logarithmic	$h = b_0 + b_1 \ln(dbh)$
M4	Curtis	$h = 1.30 + b_0 \left(\frac{dbh}{1 + dbh}\right)^{b_1}$
M5	Korf	$h = 1.30 + b_0 e^{(-b_1 dbh^{-b_2})}$
M6	Meyer	$h = 1.30 + b_0 \left(1 - e^{(-b_1 dbh)}\right)$
M7	Wykoff	$h = 1.30 + e^{(b_0 - b_1 (dbh+1)^{-1})}$
M8	Schumacher	$h = 1.30 + b_0 e^{(-b_1 dbh^{-1})}$
M9	ANN	

Note: b_0, b_1, and b_2: model parameters, *h*: tree height, *dbh*: diameter at breast height

Feed-forward backpropagation network type was selected for creating a model using an artificial neural network on MATLAB (MATLAB, R2021a). Network training function of trained data was used for updating bias and weight values based on the Levenberg-Marquadt optimization. The performance function of the model was selected as Mean Squared Error. Model 9 as ANN model was created after these steps (Figure 3).

2.4. Model evaluation and selection

In this study, the selection of the best model was executed using statistical evaluations. The statistics calculated for each model are given as follows:

The adjusted coefficient of determination:

$$R_{adj}{}^2 = 1 - \frac{(n-1)\sum\left(h_i - \hat{h}_i\right)^2}{(n-p-1)\sum\left(h_i - \bar{h}_i\right)^2} \tag{1}$$

Bias:

$$B = \frac{\sum\left(\hat{h}_i - h_i\right)}{n} \tag{2}$$

Absolute bias:

$$AB = \frac{\sum\left|\hat{h}_i - h_i\right|}{n} \tag{3}$$

Root mean square error:

$$RMSE = \sqrt{\frac{\sum\left(h_i - \hat{h}_i\right)^2}{n-p}} \tag{4}$$

Akaike Information Criterion:

$$AIC = n.\ln(RMSE) + 2p \tag{5}$$

where; h_i, $\hat{h}_i$, $\bar{h}_i$ are observed, estimated, and mean height values, *n* is the number of data used for fitting and *p* is the number of model parameters.

The best *h-d* model was determined through relative ranking system proposed by Poudel and Cao (2013). After the calculation of model statistics as above involved combining all the scores for each model and reranking them to determine the overall score. The model that scores 1 on the overall ranking is the best model, while the model that scores the highest on the overall ranking is the worst. The evaluation data set was used to critically assess the best model's accuracy, and t-tests were applied to examine whether there exists a significant difference between observed and predicted heights. Additionally, in order to examine the problem of unequal error variance, a graphical analysis was conducted of the residuals.

3. Results and Discussion

Measuring tree height is a difficult process in forestry. Even if the height of all trees is measured, their accuracy is controversial due to some application errors. Therefore, equations have been developed for tree height estimation using easily measured dbh in forestry. *h-d* models are one of the most fundamental components in the development of growth and yield models in forestry studies. With the help of these models, many tree and stand parameters can be easily estimated. In this study, 9 different *h-d* models were tested in order to develop *h-d* models for natural fir stands of Bolu-Kökez Region. For this purpose, 309 sample areas were taken and detailed *h-d* measurements were made on 862 trees. The following results were obtained in the comparisons made using the data obtained from these sample areas.

Adjusted R^2, Bias, Absolute Bias, Root Mean Square Error and, Akaike Information Criterion were calculated for 9 different models consisting of linear and nonlinear regression models, and Artificial Neural Networks. All models' information is given in Table 3.

When all these results are compared according to the ranking system developed by Poudel and Cao (2013), we obtained the best result from M9, which we obtained from Artificial Neural Networks. M7 developed by Wykoff et al. (1982) and M4 developed by Curtis (1967) were followed for the best models. The worst three models are followed by M2 called power, M1 called linear, and M6 developed by Meyer (1940) (Table 4).

Although most results of the models are consistent with the literature, unsuccessful models in this study were among the successful models in other studies. These results may be due to differences in the biology of the tree species, local conditions, and operational purposes.

Diamantopoulou (2012) fitted BPANN model from merchantable height of 471 fir trees in Greece that R value was 0.79 and RMSE was 2.99. Li and Jiang (2010) tested BPANN model about tree height model from plantation larch trees. They estimated R value as 0.83 and mean square error as 0.13. We calculated R values as 0.85 and RMSE as 3.80 for our data.

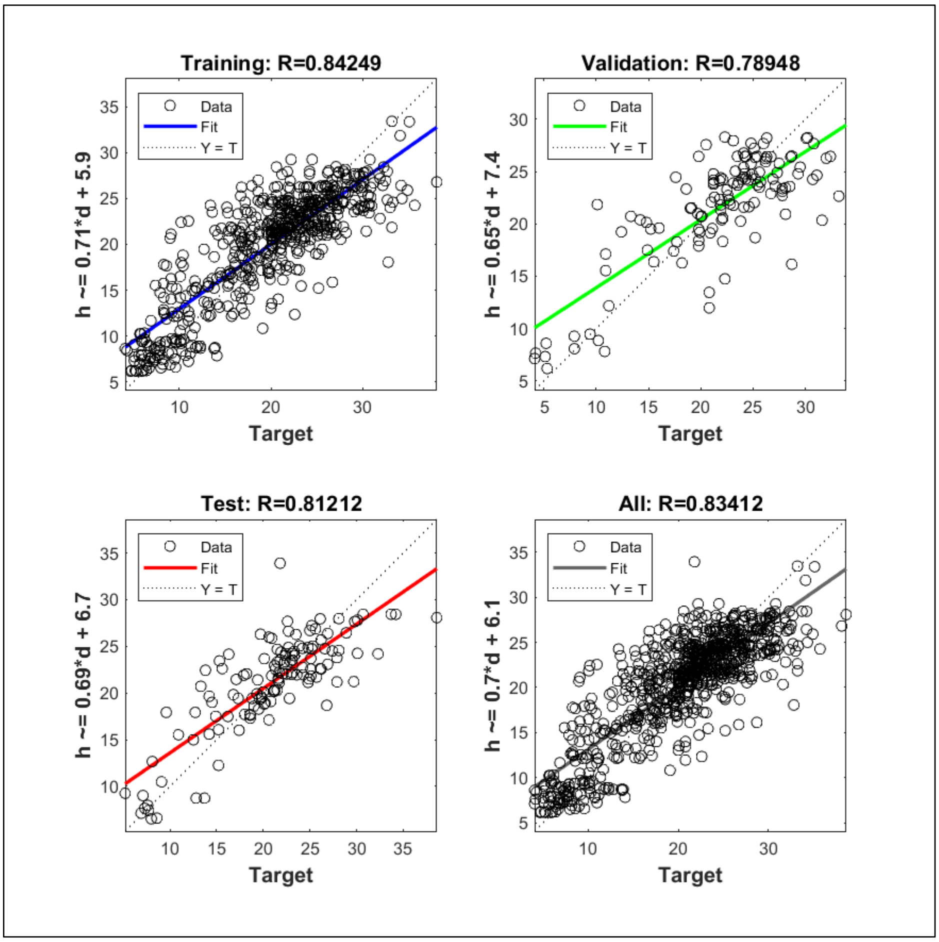

Figure 3. Plots of M9 ANN Model

When the error distribution of M9 is examined, it can be observed that the error distribution is relatively homogeneous for different diameter values and this distribution has heterogeneous variance. However, some larger predicted height errors were observed in 10-20 cm dbh interval. It has been observed that the number of sample trees in this interval is less than the sample trees in the other intervals, and the distribution of tree heights with a dbh of 10-20 cm interval is very wide (Figure 4).

Table 3. Parameter estimates and goodness-of-fit statistics for *h-d* models

Model	R_{adj}^2	Bias	Absolute bias	RMSE	AIC	Parameters	Estimate
M1	0.609	0.0	3.3	4.30	927.29	b_0 b_1	7.606 0.363
M2	0.621	-0.4	3.2	4.23	917.34	b_0 b_1	1.524 0.729
M3	0.682	0.0	2.9	3.88	861.73	b_0 b_1	-18.913 11.376
M4	0.686	0.0	2.9	3.86	858.64	b_0 b_1	35.721 19.625
M5	0.685	0.0	2.9	3.86	860.95	b_0 b_1 b_2	38.543 14.355 0.878
M6	0.677	0.2	3.0	3.91	866.79	b_0 b_1	33.081 0.027
M7	0.686	0.0	2.9	3.86	858.35	b_0 b_1	3.589 20.422
M8	0.685	0.0	2.9	3.86	859.06	b_0 b_1	35.272 18.855
M9	0.691	0.0	2.9	3.80	849.79		

Table 4. Ranking for *h-d* models

Model	Ranking for					Total Rank	Overall Ranking
	R_{adj}^2	Bias	Absolute bias	RMSE	AIC		
M1	9.00	1.00	9.00	9.00	9.00	37.00	8.38
M2	7.83	9.00	7.00	7.88	7.97	39.68	9.00
M3	1.88	1.00	1.00	2.28	2.23	8.39	1.78
M4	1.49	1.00	1.00	1.96	1.91	7.36	1.54
M5	1.59	1.00	1.00	1.96	2.15	7.70	1.62
M6	2.37	5.00	3.00	2.76	2.75	15.88	3.51
M7	1.49	1.00	1.00	1.96	1.88	7.33	1.54
M8	1.59	1.00	1.00	1.96	1.96	7.50	1.58
M9	1.00	1.00	1.00	1.00	1.00	5.00	1.00

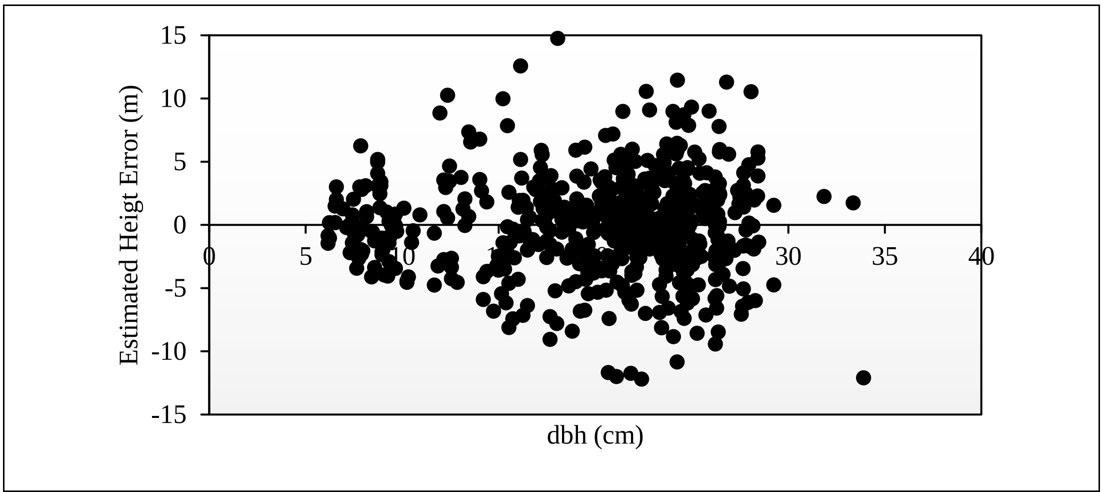

Figure 4. Plots of height estimation residuals vs. dbh

Acknowledgement

The data used in this study were taken from the Kökez Forest Management Plan, which was made under the control of the Forest Management and Planning Department of the General Directorate of Forestry. Thank you to those who contributed.

Conflict of interest

The authors declare that there is no conflict of interest.

References

Akaike, H. (1969). Fitting Autoregressive Models for Prediction. *Annals of the Institute of Statistical Mathematics*, *21*(1), 243-247. https://doi.org/10.1007/BF02532251

Atar, D. (2016). *Simulating the transition from even-aged longleaf plantations to old growth savannas with harvest management using a coupled multiple model framework* [Master's thesis, University of Florida].

Carus, S. (1998). *Aynıyaşlı doğu kayını (Fagus orientalis Lipsky.) ormanlarında artım ve büyüme* [PhD dissertation, Istanbul University]. (in Turkish)

Carus, S., & Çatal, Y. (2017). Ağlasun Yöresi Kızılçam (*Pinus brutia* Ten.) Ağaçlandırmaları İçin Bazı Çap-Boy Modellerinin Karşılaştırılması. *Turkish Journal of Forestry*, *18*(2), 94-101. (in Turkish) https://doi.org/10.18182/tjf.289330

Curtis, R. O. (1967). Height-Diameter and Height-Diameter-Age Equations for Second-Growth Douglas-Fir. *Forest Science*, *13*(4), 365–375. https://doi.org/10.1093/forestscience/13.4.365

Çatal, Y. (2009). *Batı Akdeniz Bölgesi kızılçam (Pinus brutia Ten.) meşcerelerinde artım ve büyüme* [PhD dissertation, Süleyman Demirel University]. (in Turkish)

Çatal, Y. (2012). Göller Yöresinde Yalancı Akasya, Anadolu Karaçamı ve Toros Sediri Ağaç Türleri İçin Çap-Boy Modeli. *SDÜ Orman Fakültesi Dergisi*, *13*(2), 92-96. (in Turkish)

Diamantopoulou, M. J. (2006). Tree-Bole Volume Estimation on Standing Pine Trees Using Cascade Correlation Artificial Neural Network Models.

Agricultural Engineering International: CIGR Journal, *8*, 1-6.

Diamantopoulou, M. J. (2012). Assessing A Reliable Modeling Approach of Features of Trees through Neural Network Models for Sustainable Forests. *Sustainable Computing: Informatics and Systems*, *2*(4), 190-197. http://dx.doi.org/10.1016/j.suscom.2012.10.002

Eler, Ü. (2003). *Dendrometri*. Süleyman Demirel Üniversitesi Yayınevi. (in Turkish)

Ercanli, İ., & Bolat, F. (2017). *Diameter distribution modelling based on artificial neural networks for Kunduz forests* [Oral presentation]. International Symposium on New Horizons in Forestry, Isparta, Turkey.

GDF (2020). *Orman varlığı*. General Directorate of Forestry Publications. (in Turkish)

Huxley, A. (1932). *Problems of relative growth*. The Dial Press.

Keleş, S., & Bulut, S. (2014). *Aynıyaşlı ve değişikyaşlı orman formlarının orman ekosistem fonksiyonları kapsamında karşılaştırılması* [Oral presentation]. II. Ulusal Akdeniz Orman ve Çevre Sempozyumu, Isparta, Turkey. (in Turkish)

Korf, V. (1939). Pfispevek k matematicke definici vzrustoveho zakona lesnich porostii. *Lesnicka Prace*, *18*, 339-356.

Li, Y. X., & Jiang, L. C. (2010). Application of ANN Algorithm in Tree Height Modeling. Applied Mechanics and Materials, 20, 756-761. https://doi.org/10.4028/www.scientific.net/AMM.20-23.756

MATLAB R2021a, The MathWorks, Inc., Natick, Massachusetts, United States.

Meyer, H. A. (1940). A Mathematical Expression for Height Curves. *Journal of Forestry*, *38*(5), 415-420.

Ozkal, M. K. (2017). *Models of forest inventory for Istanbul forest using airborne lidar and spaceborne imagery* [Master's thesis, Michigan Technological University].

Özçelik, R., Diamantopoulou, M. J., Brooks, J. R., & Wiant Jr., H. V. (2010). Estimating Tree Bole Volume Using Artificial Neural Network Models for Four Species in Turkey. *Journal of Environmental Management*, *91*(3), 742-753. https://doi.org/10.1016/j.jenvman.2009.10.002

Özçelik, R., Diamantopoulou, M. J., 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. https://doi.org/10.1016/j.foreco.2013.06.009

Özçelik, R., & Çapar, C. (2014). Antalya Yöresi Doğal Kızılçam Meşcereleri Için Genelleştirilmiş Çap-Boy Modellerinin Geliştirilmesi. *SDÜ Orman Fakültesi Dergisi*, *15*(1), 44-52. (in Turkish) https://doi.org/10.18182/tjf.01926

Poudel, K. P., & Cao, Q. V. (2013). Evaluation of Methods to Predict Weibull Parameters for Characterizing Diameter Distributions. *Forest Science*, *59*(2), 243-252. https://doi.org/10.5849/forsci.12-001

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Sakici, O. E., & Ozdemir, G. (2018). Stem Taper Estimations with Artificial Neural Networks for Mixed Oriental Beech and Kazdaği Fir Stands in Karabük Region, Turkey. *Cerne*, *24*(4), 439-451. https://doi.org/10.1590/01047760201824042572

Schikowski, A. B., Corte, A. P., Ruza, M. S., Sanquetta, C. R., & Montano, R. A. (2018). Modeling of Stem Form and Volume through Machine Learning. *Anais da Academia Brasileira de Ciências*, *90*(4), 3389-3401.

Schumacher, F. X. (1939). A new growth curve and its application to timber yield studies. *Journal of Forestry*, *37*(10), 819-820.

Seki, M. (2020). *Kastamonu Orman Bölge Müdürlüğü karaçam (Pinus nigra J.F. Arnold) meşcerelerine ilişkin ekolojik tabanlı büyüme modelleri* [PhD dissertation, Kastamonu University]. (in Turkish)

Socha, J., Netzel, P., & Cywicka, D. (2020). Stem Taper Approximation by Artificial Neural Network and A Regression Set Models. *Forests*, *11*(1), 79. https://doi.org/10.3390/f11010079

Temesgen, H., Hann, D. W., & Monleon, V. J. (2007). Regional Height-Diameter Equations for Major Tree Species of Southwest Oregon. *Western Journal of Applied Forestry*, *22*(3), 213-219. https://doi.org/10.1093/wjaf/22.3.213

Thanh, T. N., Tien, T. D., & Shen, H. L. (2019). Height-Diameter Relationship for *Pinus koraiensis* in Mengjiagang Forest Farm of Northeast China Using Nonlinear Regressions and Artificial Neural Network Models. *Journal of Forest Science*, *65*(4), 134-143. https://doi.org/10.17221/5/2019-JFS

Wykoff, W., Crookston, N., & Stage, A. (1982). *User's guide to the stand prognosis model*. Ogden.