ARTIFICIAL INTELLIGENCE APPLICATIONS FOR PREDICTING SOME STAND ATTRIBUTES USING LANDSAT 8 OLI SATELLITE DATA: A CASE STUDY FROM TURKEY

SAKICI, O. E. 1^* – GÜNLÜ, A. 2

¹Faculty of Forestry, Kastamonu University, Kastamonu, Turkey

²Faculty of Forestry, Çankırı Karatekin University, Çankırı, Turkey

**Corresponding author e-mail: oesakici@kastamonu.edu.tr; phone: +90-366-280-1740; fax: +90-366-215-2316*

(Received 13th Jun 2018; accepted 1st Aug 2018)

Abstract. Forest resources inventory is one of the essential parts of the sustainable forest management. Remote sensing applications have broad usage areas for this aim, since field measurements are costly, time consuming and laborious. Monitoring forest resources with various satellite images has found wide usage areas in forestry. In this study, the relationships between some stand attributes (mean diameter, basal area, stand volume and number of trees) and texture values obtained from Landsat 8 OLI satellite image were investigated for Crimean pine (Pinus nigra J.F. Arnold subsp. pallasiana (Lamb.) Holmboe) stands in Kastamonu region of Turkey. The multiple linear regression analysis and artificial neural networks (ANN) were utilized to fit stand parameters using texture values. To form the ANN architectures, various transfer functions in hidden and output layers and number of nodes ranged from 1 to 20 in hidden layer were used, and a total of 180 architectures were designed for each stand attribute. The results indicated that the regression models had low R^2 values (0.399 for mean diameter, 0.337 for basal area, 0.332 for stand volume, and 0.183 for number of trees), and the most of the ANN models were better than the regression models for predicting stand attributes. The model containing hyperbolic tangent transfer functions in both hidden and output layers for mean diameter ($R^2 = 0.593$), logistic transfer function in hidden layer and hyperbolic tangent function in output layer for basal area and stand volume $(R^2 = 0.632 \text{ and } 0.650, \text{ respectively})$, and hyperbolic tangent function in hidden layer and linear function in output layer for number of trees ($R^2 = 0.610$) were the best ANN models. This study concluded that the ANN models developed with Landsat 8 OLI were useful to predict stand parameters better than the regression models in Crimean pine stands located in Kastamonu, Turkey.

Keywords: forest inventory, satellite image, artificial neural networks, multiple linear regression

Introduction

Forest ecosystems provide so many different economic, ecological and social products and services such as wood and non-wood products, carbon sequestration, biodiversity conservation, wildlife contribution, water and soil protection, and recreation. However, these ecosystems face deforestation and desertification problems because of human activities, unusual climatic conditions due to global warming, insect attacks, erosion, wildfires, etc. In this negative situation, sustainable management of forests becomes more important. The inventory and monitoring of forest resources are the main parts of the sustainable forest management.

Traditional forest inventory based on ground measurements is very hard, timeconsuming and costly in a great forest area (Lu et al., 2004). In addition to ground measurements, remote sensing data are also widely used for forest management planning (Holmgren and Thuresson, 1998). Particularly for predicting stand-level circumstances across great forest areas, remote sensing data have been utilized efficiently to offer valuable information regarding forest constructions for forest managers (Cohen et al., 1995). Since the early 21st century, Landsat satellite data have had broad usage in many forestry studies such as land cover or land use change (Günlü et al., 2008; Huang et al., 2009; Hu et al., 2016), stand parameters predictions (Kahriman et al., 2014; Günlü and Başkent, 2015; Günlü and Kadıoğulları, 2018) and aboveground biomass estimations (Zheng et al., 2004; Lu et al., 2012; Günlü et al., 2014a).

Forest attributes such as mean diameter, stand basal area, stand volume and numbers of trees are important for forest management planning activities. In studies on predicting stand parameters using Landsat satellite images, the models were developed by using reflectance and vegetation indices values obtained from Landsat satellite data (Hall et al., 2006; Mohammadi et al., 2010; Günlü and Kadıoğulları, 2018). Besides, there are several studies to predict the stand parameters (especially for predicting aboveground biomass) by using texture values generated from Landsat satellite data (Kelsey and Neff, 2014; Safari and Sohrabi, 2016). In general, regression analysis was used to predict stand parameters with remote sensing data (Hyde et al., 2006; Gama et al., 2010). Recently, some studies have been performed to estimate the stand parameters using artificial neural networks (ANN) techniques (Ercanli et al., 2016; Reis et al., 2018).

The aims of this research are (i) to generate the ANN models estimating relationships between the stand parameters (mean diameter, stand basal area, stand volume and numbers of trees) obtained from ground measurements and texture values generated from Landsat 8 OLI satellite image, (ii) to assess the utilization of the ANN models for attaining the estimation of stand parameters by matching the regression analysis outcomes in pure Crimean pine stands of Kastamonu Regional Directorate of Forestry.

Materials and methods

Study area

This study was carried out on pure Crimean pine stands within the boundaries of Kastamonu Regional Directorate of Forestry located in the Black Sea Region of Turkey (*Fig. 1*). This directorate is the first among 28 regional directorates of Turkey in terms of growing stock (201.4 million m^3) and annual volume increment (4.3 million m^3), and these amounts equal to about 13% of the whole country. Forests cover a total of 1.26 million ha, which is about 66% of the total area of the region (General Directorate of Forestry, 2015). Crimean pine is also the most common tree species with 0.38 million ha distribution area in the region (General Directorate of Forestry, 2006).

Mean annual temperature and precipitation of the study area are 9.8 °C and 480 mm, respectively. The slope varies from 0% to 80%, and the elevation ranges between 604-1579 m above sea level, with an average of 1149 m. All sampled areas consist of naturally regenerated pure Crimean pine stands.

Field measurements

Field measurements for this study were conducted in 184 circular sample plots during summer seasons of 2015 and 2016. Sample plot sizes were determined considering stand crown closures, which is a key parameter to decide the sample plot sizes for forest inventory activities in Turkey. According to the crown closure classes of the stands (i.e., 11-40%, 41-70% and >71%), the sizes of circular sample plots were set

as 800, 600 or 400 m² with the radii of 15.96, 13.82 or 11.28 m, respectively. For some sample plots including excessive number of trees, the radii of the plots were reduced to 7.98 m (i.e., 200 m² in size). Sample plots were randomly selected to represent different stand characteristics, which directly affect the growth rate, such as stand densities, diameter classes and site qualities.



Figure 1. Landsat 8 OLI satellite image of study area

Measurements in the sample plots were initiated by recording UTM coordinates using a Global Positioning Systems (GPS) receiver placed at the center of every sample plot. In each sample plot, the diameter at breast height (*dbh*) over-bark was measured to the nearest 0.1 cm using calipers for each trees greater than 7.9 cm *dbh*, and the number of trees measured were counted. In total, 5757 trees were measured for *dbh*, and the number of trees measured in sample plots were ranged from 7 to 92 trees. Basal areas of each tree in sample plots were calculated. Volumes of the trees were predicted using the single-entry volume equation developed by Sakici et al. (2018) for Crimean pine stands located in the study area. After measuring and calculating the sample trees' dendrometric values, total basal area and total volume of each sampling plot were calculated by summing basal areas and volumes of sample trees in sample plot, respectively. The total basal area and total volume of sample plots were ranged between $0.341-4.371 \text{ m}^2$ and $0.848-45.203 \text{ m}^3$, respectively. Then, stand parameters such as mean diameter (d_q), stand basal area (G), stand volume (V) and number of trees per hectare (N) were calculated using Equations 1, 2, 3 and 4 for each sample plots.

$$d_q = \sqrt{\frac{\sum dbh^2}{n}}$$
(Eq.1)
$$G = \frac{10000}{A}g$$
(Eq.2)

APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH 16(4):5269-5285. http://www.aloki.hu • ISSN 1589 1623 (Print) • ISSN 1785 0037 (Online) DOI: http://dx.doi.org/10.15666/aeer/1604_52695285 © 2018, ALÖKI Kft., Budapest, Hungary

$$v = \frac{10000}{A}v \tag{Eq.3}$$

$$N = \frac{10000}{A} n \tag{Eq.4}$$

where, d_q is the quadratic mean diameter (cm), dbh is the tree diameter measured at 1.30 m from the ground, G is the stand basal area (m² ha⁻¹), g is the total basal area of sample plot (m²), V is the stand volume (m³ ha⁻¹), v is the total volume of sample plot (m³), N is the number of trees per hectare, n is the number of trees in the sample plot, and A is the sample plot area (m²).

The 184 sample plots were randomly divided into two groups to generate the model development and validation data sets. The data from 138 sample plots (75% of total data) were used to develop the models. As independent data set, the data from remaining 46 sample plots (25% of total data) were reserved for validation process of the developed models. The summary statistics for the sample plots were given in *Table 1* for modeling and validation data groups, separately.

	Mean diameter (cm)	Basal area (m ² ha ⁻¹)	Stand volume (m ³ ha ⁻¹)	Number of trees (ha ⁻¹)				
Modeling data $(n = 138)$								
Minimum	10.4	5.685	14.128	88				
Maximum	51.7	75.873	882.648	1800				
Mean	28.3	31.160	273.305	570				
Standard deviation	10.3	16.172	177.385	345.6				
Validation data $(n = 46)$								
Minimum	12.9	7.224	19.976	88				
Maximum	53.7	66.113	639.660	1850				
Mean	29.7	36.288	323.589	634				
Standard deviation	10.8	14.202	150.482	380.6				

Table 1. Summary statistics of sample plots

Remote sensing data and processing

The Landsat 8 Operational Land Imager (OLI) satellite image used in this research was free downloaded from United States Geological Survey Global Visualization Viewer (URL-1, 2014). The Landsat 8 OLI satellite image has already included atmospheric and geometric corrections, and radiometric calibration. In this study, five bands (Band 2, 3, 4, 5 and 7) of Landsat 8 OLI satellite image with a spatial resolution of 30 meters were used. Eight different texture metrics (contrast, correlation, dissimilarity, entropy, homogeneity, mean, second moment and variance) for each band were produced using four different window sizes (3 x 3, 5 x 5, 7 x 7 and 9 x 9). To produce texture values, ENVI software was used. After that, depending on the size of the sample areas, buffer zones (with radius of 7.98, 11.28, 13.82 or 15.96 m) were composed around the sample plots' centers in accordance with UTM coordinates recorded by GPS receiver. The texture images produced for each band were overlaid using GIS with the sample plots. The texture values of each sample plot were calculated

by two different methods. If the sample plot centers were at or near the center of a pixel, the texture values of the sample plots were calculated as the value of a single pixel for 200, 400 and 600 m² sample plots, while calculated by taking the average of some pixel values in the buffer zone for 800 m² sample plots (*Fig. 2a*). However, in the second method, if the sample plot centers were far from the pixel centers, the texture values of each sample plot were computed by taking the average of pixel values in the buffer zone for all sample plot sizes (*Fig. 2b*). In this way, the texture values for five bands, four window sizes and eight different texture metrics were calculated. Thus, a total of 160 texture values were obtained for each sample plot.



Figure 2. Texture value calculation of sample plots according to buffer zones on satellite image

Regression models

The multiple linear regression analysis was used to develop the equations modeling the interactions between stand parameters (mean diameter, stand basal area, stand volume and number of trees) and texture values obtained from Landsat 8 OLI satellite image. To obtain multiple linear regression models based on stepwise variable selection method, the ordinary least squares technique was used. The dependent variables in these models were quadratic mean diameter, stand basal area, stand volume and number of trees per hectare, and independent variables were 160 texture values of modeling data group containing 138 sample plots. The dependent variables were observed values from field works, while the independent ones were produced values from satellite image. The stepwise regression procedure in IBM SPSS 23 software was used to select the statistically significant (p < 0.05) texture values as predictor variables to estimate stand parameters. The relationships between stand parameters and texture values were assumed linear as given formula below (*Eq. 5*):

Stand Parameter =
$$b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n + e$$
 (Eq.5)

where *Stand Parameter* is quadratic mean diameter, stand basal area, stand volume or number of trees per hectare, X_i (i = 1 to n) are texture values, b_i represent the model coefficients, and e is the error term.

Artificial neural networks

Artificial intelligence techniques were also employed to model stand parameters beside multiple linear regression analysis to determine prediction success of these techniques and to compare with regression modeling. For this purpose, artificial neural network (ANN) models were developed for each stand parameter examined in this study. For defining the architecture of neural networks, there are several criteria such as number of layers, learning algorithms, transfer function forms, number of nodes in hidden layer and determination of data sizes for training, verification and test processes. The ANN models developed in this study consist of three layers: input, hidden and output layers. The feed-forward back-propagation network structure was selected because of its success popularity in forestry literature (e.g., Özçelik et al., 2014; Diamantopoulou et al., 2015). The learning algorithm used in ANN models was the Levenberg-Marquardt algorithm. Three transfer function forms (linear, logistic and hyperbolic tangent) were examined in hidden and output layers, separately (*Eqs. 6, 7* and δ).

$$f(s) = purelin(s) = s$$
 (Linear function) (Eq.6)

$$f(s) = logsig(s) = \frac{1}{1 + e^{(-s)}}$$
 (Logistic function) (Eq.7)

$$f(s) = tanh(s) = \frac{1 - e^{(-2s)}}{1 + e^{(-2s)}}$$
 (Hyperbolic tangent function) (Eq.8)

where $s = \sum w_i x_i$, w_i are weights and x_i are the input variables.

To determine the most predictive alternatives, the number of nodes in hidden layers ranged from 1 to 20 adding one by one in training process of ANN models. Thus, a total of 180 ANN model architectures were created for each stand parameter (*Fig. 3*). The ANN models were built using the neural network toolbox in R2015a version of MATLAB software. The modeling data obtained from 138 sample plots were used for ANNs' fitting. Input variables of ANN models were texture values of satellite image determined as the best predictor for each stand parameter in multiple linear regression analyses, and output (target) variable was observed mean diameter, stand basal area, stand volume or number of trees according to the stand attributes fitted.



Figure 3. ANN models architecture for stand parameters

APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH 16(4):5269-5285. http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online) DOI: http://dx.doi.org/10.15666/aeer/1604_52695285 © 2018, ALÖKI Kft., Budapest, Hungary

Model comparisons and validation tests

The regression and ANN models were evaluated based on four goodness-of-fit statistics; the coefficient of determination (R^2) , root mean square error (*RMSE*), bias and mean absolute error (*MAE*). Corresponding mathematical forms of statistical criteria utilized were defined as *Equations 9*, 10, 11 and 12.

$$R^{2} = 1 - \frac{\sum (y_{i} - \hat{y}_{i})^{2}}{\sum (y_{i} - \bar{y})^{2}}$$
(Eq.9)

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 1}}$$
(Eq.10)

$$Bias = \frac{\sum(y_i - \hat{y}_i)}{n}$$
(Eq.11)

$$MAE = \frac{\sum |y_i - \hat{y_i}|}{n}$$
(Eq.12)

In these equations; y_i and \hat{y}_i are observed and estimated values of corresponding stand attribute, and *n* is sample size.

When comparing alternative models, it is desirable that the R^2 values are high while the others (*RMSE*, *Bias* and *MAE*) are low. For ranking of models, taking into account of all goodness-of-fit statistics together is better than the ranking of each criteria separately. The relative ranking method proposed by Poudel and Cao (2013) was used for model comparisons. In this method, the relative rank of model *i* for a statistical criterion is defined using *Equation 13*.

$$R_i = 1 + \frac{(m-1)(S_i - S_{min})}{(S_{max} - S_{min})}$$
(Eq.13)

where R_i is the relative rank of model *i* (*i* = 1, 2, ..., m), S_i is the goodness-of-fit statistic of model *i*, S_{min} and S_{max} are the minimum and maximum values of S_i , respectively.

For each stand parameter, relative rankings of ANN models were first implemented according to number of nodes in hidden layer for transfer function pairs of hidden and output layers, separately, for each statistical criterion. So, four rankings with 20 ANN models were formed for nine transfer function pairs (i.e., linear, logistic and hyperbolic tangent functions were used in both hidden and output layers) for every stand parameter fitting. The model with the highest R^2 was ranked as 1 and the lowest R^2 was ranked as 20 for coefficient of determination, while the model with the lowest value was ranked as 1 and the highest value was ranked as 20 for *RMSE*, *Bias* and *MAE*. Then, four relative ranks of each model according to statistical criteria were summed. The second relative ranking was generated using total relative ranks of ANN models based on number of nodes ranged from 1 to 20 for each transfer function pairs. Thus, the most successful ANN models were specified for every transfer function pairs of each stand parameter fitting.

To compare estimation success of regression and ANN models, the relative ranks of 10 models (a regression model and the best ANN models for nine transfer function pairs) were rebuilt for each stand parameter. So, the most successful models were identified for predicting stand parameters evaluated in this study. The validities of successful models were tested by comparing the validation data observed from 46 sample plots and models' results by Student's t-test (also called paired samples t-test) procedure in IBM SPSS Statistics 23 software. To exhibit the prediction abilities of the models, residual graphs based on observed vs. predicted stand parameters were also prepared.

Results

In this study, multiple regression and artificial intelligence techniques were conducted for predicting some stand parameters such as mean diameter, basal area, stand volume and number of trees using texture values obtained from Landsat 8 OLI satellite image, and regression and ANN models were developed with these techniques. To determine the best predictor texture values for predicting stand attributes and to develop linear regression models for each stand parameter, multiple linear regression technique was used. In this analyses, all stand parameters were tried to fit using 160 texture values as candidate independent variables. These variables were produced by combining five bands of Landsat 8 OLI satellite image, four window sizes and eight texture metrics. The results of regression analyses are given in *Table 2*. As seen in this table, six, seven, six and three variables included in regression models of mean diameter, basal area, stand volume and number of trees, respectively. All coefficients of these variables were statistically significant as well as constants of models except for number of trees (p < 0.05). It might be said that all stand parameters were fitted poorly with regression models due to the low R^2 values. The maximum coefficient of determination value was obtained for mean diameter ($R^2 = 0.399$), while the minimum was for number of trees ($R^2 = 0.183$). The R^2 values of basal area and stand volume models were similar, which were 0.337 and 0.332, respectively.

In regression models developed for stand parameters, texture value combinations containing Band 2 were not statistically significant for number of trees as well as Band 4 for mean diameter and number of trees and Band 7 for all stand parameters except mean diameter, while some combinations with Band 3 and Band 5 were significant for all. The combinations including 7 x 7 window sizes were found insignificant for number of trees, and 9 x 9 ones were for the parameters except basal area. Some combinations with 3 x 3 and 5 x 5 window sizes were significant for all stand parameters. Texture values comprising *contrast* for mean diameter, *dissimilarity* for all, *entropy* and *homogeneity* for the parameters except mean diameter, *mean* for number of trees, *second moment* for all except number of trees, and *variance* for basal area and number of trees were not statistically significant. Some combinations with *correlation* were significant for all stand parameters.

ANN models were also fitted for stand parameters using input variables found statistically significant in regression models. Totally 180 ANN architectures including various transfer functions in hidden and output layers and number of nodes from 1 to 20 in hidden layer were designed for every stand parameter. Thus, a total of 720 ANN procedures were implemented in this study. For all stand parameters evaluated, all ANN models were successfully completed the training process except architectures

comprising logistic transfer function in output layer. The ANN models with logistic transfer function in output layer could not give any result. It means that 120 ANN models could be ranked instead of 180 models for each stand parameter.

Stand	Independent	Variab	le combination	(Texture value)	Coefficients	p ²	CEEa
parameter	variable	Band	Band Window size Texture metric		(\boldsymbol{b}_i)	ĸ	SEE
	Constant				49.531*		
Mean diameter $(d - cm)$	\mathbf{X}_1	2	3x3	Correlation	-0.026*		
	\mathbf{X}_2	3	7x7	Entropy	-0.149*		
	X_3	5	5 5x5 Homogeneity -0.086*		0.399	8.158	
(u_q, cm)	X_4	5 5x5 Mean 0.072		0.072*			
	X_5	5 5x5 Variance -0.133*					
	X_6	7	3x3	Variance	0.073*		l
	Constant				31.709***		
	\mathbf{X}_1	2	9x9	Mean	-0.774**		
	X_2	3	9x9	Mean 0.556** Correlation 0.066** Contrast -0.195*** Contrast 0.583***		0.337	15.348
Basal area	X_3	4	7x7				
$(G, m^2 ha^{-1})$	X_4	5	3x3				
	X_5	5	5x5				
	X_6	5	9x9	Contrast -0.196*			
	X_7	X ₇ 5 9x9 Correlatio		Correlation	-0.078**		
	Constant				225.797***		
	\mathbf{X}_1	2	3x3	Mean -6.949** Mean 4.149* Correlation 0.640** Contrast -1.514**		0.332	172.487
C 1 1	\mathbf{X}_2	3	3x3				
Stand volume $(V \text{ m}^3 \text{ ha}^{-1})$	X_3	4	7x7				
(, 111 11a)	X_4	5	3x3				
	X_5	5	5x5	Contrast	5.536***		
	X_6	₆ 5 5x5 Variance		-2.619**			
Number of trees (N, ha^{-1})	Constant				-21.209 ^{ns}		
	\mathbf{X}_1			Contrast	20.264***	0 102	425.452
	X_2			Correlation	1.693**	0.183	
	X ₃	5	3x3	Second moment	1.402**		

Table 2. Multiple linear regression results for stand parameters

^aStandard error of estimate

*p < 0.05, **p < 0.01, ***p < 0.001, ^{ns}non-significant (p > 0.05)

Four statistical criteria (R^2 , *RMSE*, *Bias* and *MAE*) were used to rank the ANN models. These rankings were applied within six transfer function pairs, i.e. linear, logistic and hyperbolic tangent functions in hidden layer combined with linear and hyperbolic tangent functions in output layer, separately. Hence, a total of 24 ranking lists were produced due to the presence of four stand attributes and six transfer function pairs. Finally, the best ANN architectures within each ranking list were determined to compare with each other and regression models.

The regression models and all of the best ANN models for each stand parameter were re-ranked to decide the most successful models for predicting stand parameters.

The final ranking lists are given in *Table 3*. In this table, it is seen that the regression models and the ANN models including linear transfer function in hidden layer were unsuccessful than the other ANN models. The ANN₄ models (i.e. the logistic function in hidden layer and hyperbolic tangent function in output layer) were the most successful for predicting basal area (with 12 nodes in hidden layer) and stand volume (with 9 nodes in hidden layer) having high R^2 values as 0.632 and 0.650, respectively. The ANN₆ model (i.e. the hyperbolic tangent functions in both hidden and output layers, and 10 nodes in hidden layer) was the best for mean diameter prediction $(R^2 = 0.593)$, while ANN₅ (i.e. the hyperbolic tangent function in hidden layer, the linear function in output layer, and 9 nodes in hidden layer) was the best to predict number of trees ($R^2 = 0.610$). The other goodness-of-fit statistics had similar trends to the coefficient of determinations among the models compared. In the other words, the ANN models having higher R^2 values had lower *RMSE*, *Bias* and *MAE* values. As it can be understood from these conclusions, the powers of stand parameter predictions were considerably increased with artificial intelligence applications. These increases of prediction powers are approximately 50% for mean diameter, nearly 100% for basal area and stand volume, and more than 200% for number of trees.

If the ranking lists in *Table 3* are examined in detail, the regression models were weaker than the ANN models except the architectures containing linear transfer functions in both hidden and output layers (ANN₁). For number of trees prediction, the regression model was also stronger than the ANN model comprising the linear transfer function in hidden layer and the hyperbolic tangent transfer function in output layer (ANN₂). According to the transfer function in hidden layer, the ANN models with hyperbolic tangent function were more successful for mean diameter and number of trees predictions, while the models with logistic function for basal area and stand volume. Besides, according to the transfer function in output layer, the ANN model with linear function were more successful for number of trees prediction, while the models with hyperbolic tangent function for the other stand parameters. There was not any general trend for the ANN models with respect to the number of nodes in hidden layer. However, the node numbers were around 10 in the best models for all parameters.

The validities of the developed regression models and the best ANN models were tested with Student's t-test using independent data set obtained from 46 sample plots. For all the models tested, there were no significant differences between observed and predicted values (p > 0.05) and consequently it was decided that the models were statistically usable for predicting the aforementioned stand parameters (*Table 4*). Thus, because of their statistical successes, the ANN₆ model for mean diameter, the ANN₄ models for basal area and stand volume, and the ANN₅ model for number of trees estimations can be used.

The residual distributions of predicted stand parameters obtained by the best ANN models and regression models using all data from 184 sample plots are shown in *Figure 4*. When the residual distributions are examined, it is seen that the residuals were distributed randomly in all regression and ANN models. The regression and ANN models had similar residuals and the mean residuals were close to zero for each stand parameter. Mean residual values for mean diameter, basal area, stand volume and number of trees estimates were -0.4 cm, -0.37 m² ha⁻¹, -3.96 m³ ha⁻¹ and 18.6 ha⁻¹ for regression models (*Fig. 4a-d*), and -0.1 cm, 0.64 m² ha⁻¹, -10.04 m³ ha⁻¹ and 5.5 ha⁻¹ for the best ANN models (*Fig. 4e-h*), respectively.

		ANN architecture								
Stand parameter	Model	Transfer function in hidden layer	Transfer function in output layer	Number of nodes in hidden layer	R^2 (R_i)	RMSE (R _i)	Bias (R _i)	MAE (R _i)	Total <i>R</i> i	General rank
Mean diameter (d_q, cm)	Regression				0.399 (6.94)	7.950 (6.93)	-0.13 (3.86)	6.49 (7.00)	24.72	6.04
	ANN_1	Linear	Linear	6	0.397 (7.00)	7.967 (7.00)	-0.24 (7.00)	6.49 (7.00)	28.00	7.00
	ANN_2	Linear	Hyp. Tangent	16	0.398 (6.97)	7.956 (6.95)	-0.03 (1.00)	6.48 (6.96)	21.89	5.22
	ANN ₃	Logistic	Linear	15	0.550 (2.32)	6.879 (2.41)	-0.03 (1.00)	5.37 (2.80)	8.52	1.32
	ANN_4	Logistic	Hyp. Tangent	8	0.497 (3.94)	7.273 (4.07)	0.09 (2.71)	5.86 (4.64)	15.36	3.31
	ANN ₅	Hyp. Tangent	Linear	9	0.538 (2.68)	6.971 (2.79)	0.09 (2.71)	5.10 (1.79)	9.98	1.74
	ANN ₆	Hyp. Tangent	Hyp. Tangent	10	0.593 (1.00)	6.546 (1.00)	0.15 (4.43)	4.89 (1.00)	7.43	1.00
	Regression				0.337 (6.80)	14.897 (6.83)	-0.0873 (1.00)	12.0117 (7.00)	21.63	6.92
	ANN_1	Linear	Linear	5	0.327 (7.00)	15.011 (7.00)	-0.0977 (1.07)	11.8592 (6.76)	21.83	7.00
	ANN ₂	Linear	Hyp. Tangent	11	0.359 (6.37)	14.655 (6.45)	-0.1442 (1.36)	11.8403 (6.73)	20.92	6.63
Basal area $(G, m^2 ha^{-1})$	ANN ₃	Logistic	Linear	16	0.544 (2.73)	12.354 (2.92)	-0.2436 (2.00)	9.3266 (2.79)	10.44	2.34
	ANN ₄	Logistic	Hyp. Tangent	12	0.632 (1.00)	11.100 (1.00)	0.5824 (4.15)	8.1835 (1.00)	7.15	1.00
	ANN ₅	Hyp. Tangent	Linear	8	0.465 (4.29)	13.387 (4.51)	1.0289 (7.00)	10.2516 (4.24)	20.04	6.27
	ANN_6	Hyp. Tangent	Hyp. Tangent	18	0.491 (3.77)	13.057 (4.00)	0.6217 (4.41)	9.8674 (3.64)	15.82	4.54
	Regression				0.332 (6.94)	168.055 (6.95)	0.0467 (1.00)	130.3944 (6.93)	21.82	6.23
	ANN ₁	Linear	Linear	11	0.329 (7.00)	168.471 (7.00)	-2.9128 (2.84)	130.8383 (7.00)	23.84	7.00
	ANN ₂	Linear	Hyp. Tangent	4	0.392 (5.82)	160.306 (5.95)	-3.2678 (3.07)	125.7314 (6.16)	21.01	5.92
Stand volume $(V, m^3 ha^{-1})$	ANN ₃	Logistic	Linear	8	0.540 (3.06)	139.422 (3.28)	-0.1988 (1.10)	101.8162 (2.24)	9.68	1.62
	ANN ₄	Logistic	Hyp. Tangent	9	0.650 (1.00)	121.599 (1.00)	-6.3624 (5.05)	94.2181 (1.00)	8.05	1.00
	ANN ₅	Hyp. Tangent	Linear	8	0.585 (2.21)	132.460 (2.39)	-5.9340 (4.78)	101.5374 (2.20)	11.58	2.34
	ANN_6	Hyp. Tangent	Hyp. Tangent	5	0.513 (3.56)	143.435 (3.80)	-9.4027 (7.00)	103.9747 (2.60)	16.95	4.38
	Regression				0.183 (6.91)	419.241 (6.92)	0.03 (1.00)	289.77 (7.00)	21.82	6.62
Number of trees (<i>N</i> , ha-1)	ANN ₁	Linear	Linear	14	0.176 (7.00)	421.113 (7.00)	-9.25 (2.00)	284.57 (6.61)	22.61	6.88
	ANN ₂	Linear	Hyp. Tangent	7	0.216 (6.46)	410.800 (6.54)	-22.41 (3.44)	283.41 (6.52)	22.97	7.00
	ANN ₃	Logistic	Linear	10	0.599 (1.30)	293.818 (1.37)	-12.20 (2.33)	209.70 (1.00)	5.99	1.30
	ANN ₄	Logistic	Hyp. Tangent	19	0.621 (1.00)	285.546 (1.00)	-55.11 (7.00)	216.10 (1.48)	10.48	2.81
	ANN ₅	Hyp. Tangent	Linear	9	0.610 (1.15)	289.685 (1.18)	6.92 (1.75)	209.71 (1.00)	5.08	1.00
	ANN ₆	Hyp. Tangent	Hyp. Tangent	17	0.551 (1.94)	310.820 (2.12)	-8.17 (1.89)	217.52 (1.59)	7.54	1.82

Table 3. Relative ranks of models based on goodness-of-fit statistics^{1,2}

¹Values in parenthesis indicate the relative ranks of the models based on statistical criteria

²The best models are marked in bold for each stand parameters



Figure 4. Residual distributions of the regression models (a-d) and selected ANN models (e-h)

When the observed stand parameters and predicted values of these parameters with two modeling approaches were compared considering *Table 4*, the mean diameter predictions of the regression and ANN_6 models were similar to observed values with mean differences of 1.4 and 1.5 cm, respectively. The basal area predictions of the ANN₄ model were very close to the observed values with mean difference of 0.860 m², while the regression model's predictions were quite different with of 5.221 m² per hectare. Since mean differences of the regression and ANN₄ models were 65.586 and

44.559 m^3 per hectare, the ANN₄ model was more prosperous for the stand volume estimations. The comparison for the number of trees was resulted different from the other stand parameters, since the mean difference of the regression model was positively 45.4 trees while of the ANN₅ model was negatively 63.5 trees per hectare. However, the ANN₅ model was superior than the regression model for the number of trees prediction when the mean residuals of the models took into consideration.

Stand	Data	Madal	Maan	Pair	red differe	4		
parameter	Data	Widdei	Mean	Mean	SD ^a	SEE ^b	l	p
Mean diameter $(d \text{ cm})$	Observed		29.7					
	Predicted	Regression	28.3	1.4	10.5	1.55	0.894	0.376
(<i>uq</i> , em)		ANN_6	28.1	1.5	10.8	1.60	0.948	0.348
Basal area $(G, m^2 ha^{-1})$	Observed		36.288					
	Predicted	Regression	31.067	5.221	22.949	3.384	1.543	0.130
		ANN_4	35.427	0.860	23.279	3.432	0.251	0.803
Stand volume $(V, m^3 ha^{-1})$	Observed		323.589					
	Predicted	Regression	258.003	65.586	238.683	35.192	1.864	0.069
		ANN_4	279.030	44.559	312.397	46.060	0.967	0.339
Number of trees (N, ha ⁻¹)	Observed		633.7					
	Predicted	Regression	588.3	45.4	469.7	69.26	0.655	0.516
		ANN ₅	697.2	-63.5	673.3	99.27	-0.640	0.525

Table 4. Student's t-test results for the regression and ANN models selected

^aStandard deviation

^bStandard error of estimate

Discussion and conclusions

In this study, the relationships between some stand attributes (mean diameter, basal area, stand volume and number of trees) and texture values obtained from Landsat 8 OLI satellite image were predicted for Crimean pine stands in Kastamonu region of Turkey. The multiple linear regression analysis and ANN technique were performed to fit stand attributes using texture values. The stand attributes estimation models developed by ANN models performed better than multiple linear regression models except ANN_1 for basal area, stand volume and number of trees and except ANN_1 and ANN_2 for mean diameter.

The band reflectance and vegetation indices were used as independent variables in some published studies performed to estimate the stand parameters using the Landsat satellite images (Mohammadi et al., 2010; Kahriman et al., 2014; Günlü and Kadıoğulları, 2018). However, in the literature no studies have been found to predict mean diameter, basal area, stand volume and number of trees using texture values obtained from Landsat satellite images. On the other hand, there are some studies on estimating stand parameters such as aboveground biomass and carbon stock using textures values obtained from Landsat satellite data images (Kelsey and Neff, 2014; Zhao et al., 2016; Safari and Sohrabi, 2016 In addition, there are many studies on estimating stand parameters such as stand volume, basal area, number of trees, aboveground biomass and carbon stock using texture values obtained from the other satellite data images (Castillo-Santiago et al., 2010; Özdemir and Karnieli, 2011;

Eckert, 2012; Pu and Cheng, 2015; Wallner et al., 2015; Xie et al., 2017). In general, regression analysis is used in the studies to estimate stand parameters using remote sensing data images (Castillo-Santiago et al., 2010; Mohammadi et al., 2011; Günlü et al., 2014b; Kahriman et al., 2014).

The reflectance values, vegetation indices and texture values were utilized as independent variables of a multiple linear regression model in some studies on the subject. For instance, the models using vegetation indices obtained from Landsat ETM+ data had the R^2 values as 0.43 for stand volume and 0.73 for tree density (Mohammadi et al., 2010), while the texture variables of IKONOS-2 explained 35% of the variation in basal area of even-aged stands in temperate forests (Kayitakire et al., 2006). In other studies, the reflectance and vegetation indices values obtained from Landsat TM satellite data explained 61% and 70% of tree density, respectively, in mixed forest areas, and generated from pan-sharpened IKONOS satellite data explained 41% and 43% of stand volume, and 55% and 59% of basal area, respectively (Kahriman et al., 2014; Günlü et al., 2014b). In the multiple linear regression models with digital number values of Landsat TM, the R^2 values for mean diameter, basal area, stand volume and number of trees were founded as 0.25, 0.32, 0.37 and 0.44, respectively (Günlü and Kadıoğulları, 2018). In the same study, the R^2 values of the multiple linear regression analyses obtained from vegetation indices values for mean diameter, basal area, stand volume and number of trees were founded as 0.17, 0.34, 0.36 and 0.28, respectively.

The results of our study indicated that the R^2 values of the multiple linear regression analyses generated with texture values from Landsat 8 OLI satellite image for mean diameter, basal area, stand volume and number of trees were founded as 0.40, 0.34, 0.33 and 0.18, respectively. When the results obtained from our study are compared with some other studies mentioned above (Mohammadi et al., 2010: Kahriman et al., 2014; Günlü and Kadıoğulları, 2018), the regression model generated from texture values is not suitable for estimating mean diameter, basal area, stand volume and number of trees due to the low R^2 values.

In recent years, the ANN method used in some published studies on predicting leaf area index (Shoemaker and Cropper, 2010), stand volume and basal area (Shataee, 2013; Santi et al., 2015), stand carbon stock (Ercanli et al., 2016), and aboveground biomass (Ni et al., 2017; Yan et al., 2018). We also used the ANN method to predict the stand parameters in this study. The comparisons between multiple linear regression and ANN models for estimating stand parameters show that the abilities of the ANN models are higher. In other words, the R^2 values of the best ANN models increased between 48% and 239% for the stand parameters compared to the regression models. Similar results were published in some rare studies. For instance, Ercanli et al. (2016) modelled aboveground stand carbon stock using multiple linear regression analysis and ANN with vegetation indices obtained from Landsat TM satellite data and founded that the R^2 values of 0.43 and 0.65, respectively. In another study, a comparative analysis was conducted for different satellite data and modeling techniques (e.g. ANN, random forest, linear regression, k-nearest neighbor, support vector regression) for aboveground biomass prediction in different forest types (Gao et al., 2018). The results obtained from this study showed that the ANN models were more accurate for mixed, broadleaf and conifer forest types (RMSE values ranged 30.0-36.0 Mg/ha, 24.0-26.4 Mg/ha and 28.2-30.9, respectively) than linear regression model (*RMSE* values ranged 32.6-37.0 Mg/ha, 24.4-31.0 Mg/ha and 28.2-32.4, respectively). It is seen that there are few studies to predict the stand parameters using both regression analysis and ANN method. Thus, our study is one of the rare studies that use both the multiple linear regression and ANN models in estimating of stand parameters using remote sensing data, especially texture values. Therefore, investigating the estimation success of the ANN models using satellite data for stand attributes is essential for further studies.

The ANN models can be suggested as more suitable than regression models for predicting stand attributes such as mean diameter, basal area, stand volume and number of trees using remote sensing data, since the estimation power of the ANN models are higher than the regression models. The texture values might be used successfully as predictor variables in these models. The utilization of different model techniques such as random forest, *k*-nearest neighbor, support vector regression and mixed effect modeling and various satellite data such as optical, active and combined can improve the model achievement criteria in various forest ecosystems.

Acknowledgements. The field data of this study were obtained from a research project supported by The Scientific and Technological Research Council of Turkey (TUBITAK), Project no: TOVAG-214O217.

REFERENCES

- [1] Castillo-Santiago, M. A., Ricker, M., de Jong, B. H. J. (2010): Estimation of tropical forest structure from SPOT-5 satellite images. International Journal of Remote Sensing 31(10): 2767-2782.
- [2] Cohen, W. B., Spies, T. A., Fiorella, M. (1995): Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, USA. – International Journal of Remote Sensing 16(4): 721-746.
- [3] Eckert, S. (2012): Improved forest biomass and carbon estimations using texture measures from WorldView-2 satellite data. Remote Sensing 4(4): 810-829.
- [4] Ercanli, İ., Günlü, A., Şenyurt, M., Bolat, F., Kahriman, A. (2016): Artificial neural network for predicting stand carbon stock from remote sensing data for even-aged scots pine (*Pinus sylvestris* L.) stands in the Taşköprü-Çiftlik forests. – Proceedings of the 1st International Symposium of Forest Engineering and Technologies (FETEC 2016), 02-04 June, Bursa, Turkey.
- [5] Diamantopoulou, M. J., 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. – Biosystems Engineering 133: 33-45.
- [6] Gama, F. F., Santos, J. R., Mura, J. C. (2010): Eucalyptus biomass and volume estimation using interferometric and polarimetric SAR data. Remote Sensing 2(4): 939-956.
- [7] Gao, Y., Lu, D., Li, G., Wang, G., Chen, Q., Liu, L., Li, D. (2018): Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sensing 10(4): 627.
- [8] General Directorate of Forestry (2006): Orman Varlığımız. General Directorate of Forestry Publications, Ankara.
- [9] General Directorate of Forestry (2015): Türkiye Orman Varlığı 2015. General Directorate of Forestry Publications, Ankara.
- [10] Günlü, A., Başkent, E. Z. (2015): Estimating crown closure of forest stands using Landsat TM data: A case study from Turkey. – Environmental Engineering and Management Journal 14(1): 183-193.
- [11] Günlü, A., Kadıoğulları, A. İ. (2018): Modeling forest stand attributes using Landsat ETM+ and QuickBird satellite images in western Turkey. Bosque 39(1): 49-59.

- [12] Günlü, A., Sivrikaya, F., Baskent, E. Z., Keles, S., Çakir, G., Kadiogullari, A. İ. (2008): Estimation of stand type parameters and land cover using Landsat-7 ETM image: A case study from Turkey. – Sensors 8(4): 2509-2525.
- [13] Günlü, A., Ercanlı, İ., Başkent, E. Z., Çakır, G. (2014a): Estimating aboveground biomass using Landsat TM imagery: A case study of Anatolian Crimean pine forests in Turkey. – Annals of Forest Research 57(2): 289-298.
- [14] Günlü, A., Ercanlı, İ., Sönmez, T., Başkent, E. Z. (2014b) Prediction of some stand parameters using pan-sharpened IKONOS satellite image. European Journal of Remote Sensing 47(1): 329-342.
- [15] Hall, R. J., Skakun, R. S., Arsenault, E. J., Case, B. S. (2006): Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest Ecology and Management 225(1-3): 378-390.
- [16] Holmgren, P., Thuresson, T. (1998): Satellite remote sensing for forestry planning-A review. Scandinavian Journal of Forest Research 13(1-4): 90-110.
- [17] Hu, T., Yang, J., Li, X., Gong, P. (2016): Mapping urban land use by using Landsat images and open social data. Remote Sensing 8(2): 151.
- [18] Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., Zhu, Z. (2009): Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States. – Remote Sensing of Environment 113(7): 1430–1442.
- [19] Hyde, P., Dubayah, R., Walker, W., Blair, B., Hofton, M., Hunsaker, C. (2006): Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. – Remote Sensing of Environment 102(1-2): 63-73.
- [20] Kahriman, A., Günlü, A., Karahalil, U. (2014): Estimation of crown closure and tree density using Landsat TM satellite images in mixed forest stands. – Journal of the Indian Society Remote Sensing 42(3): 559-567.
- [21] Kayitakire, F., Hamel, C., Defourny, P. (2006): Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. – Remote Sensing of Environment 102(3-4): 390-401.
- [22] Kelsey, K. C., Neff, J. C. (2014): Estimates of aboveground biomass from texture analysis of Landsat imagery. Remote Sensing 6(7): 6407-6422.
- [23] Lu, D., Mausel, P., Brondizio, E., Moran, E. (2004): Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. – Forest Ecology and Management 198(1-3): 149-167.
- [24] Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., Laurin, G. V., Saah, D. (2012): Aboveground forest biomass estimation with Landsat and LIDAR date and uncertainty analysis of the estimates. – International Journal of Forestry Research Article ID: 436537.
- [25] Mohammadi, J., Shataee, S., Yaghmaee, F., Mahiny, A. S. (2010): Modelling forest stand volume and tree density using Landsat ETM data. – International Journal of Remote Sensing 31(11): 2959-2975.
- [26] Mohammadi, J., Shataee, S., Babanezhad, M. (2011): Estimation of forest stand volume, tree density and biodiversity using Landsat ETM + Data, comparison of linear and regression tree analyses. – Procedia Environmental Sciences 7: 299-304.
- [27] Ni, X., Cao, C., Zhou, Y., Ding, L., Choi, S., Shi, Y., Park, T., Fu, X., Hu, H., Wang, X. (2017): Estimation of forest biomass patterns across Northeast China based on allometric scale relationship. – Forests 8(8): 288.
- [28] Özçelik, R., Diamantopoulou, M. J., Brooks, J. R. (2014): The use of tree crown variables in over-bark diameter and volume prediction models. iForest-Biogeosciences and Forestry 7(3): 132-139.
- [29] Ozdemir, I., Karnieli, A. (2011): Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. International Journal of Applied Earth Observation and Geoinformation 13(5): 701-710.

- [30] Poudel, K. P., Cao, Q. V. (2013): Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. Forest Science 59(2): 243-252.
- [31] Pu, R., Cheng, J. (2015): Mapping forest leaf area index using reflectance and textural information derived from WorldView-2 imagery in a mixed natural forest area in Florida, US. – International Journal of Applied Earth Observation and Geoinformation 42: 11-23.
- [32] Reis, A. A., Acerbi, F. W., Mello, J. M., Carvalho, L. M. T., Gomide, L. R. (2018): Relationship between spectral data and dendrometric variables in Eucalyptus sp. stands. – Floresta e Ambiente 25: e20150170.
- [33] Safari, A., Sohrabi, H. (2016): Ability of Landsat-8 OLI derived texture metrics in estimating aboveground carbon stocks of Coppice oak forests. – Proceedings of the XXIII ISPRS Congress, 12-19 July, Prague, Czech Republic.
- [34] Sakici, O. E., Sağlam, F., Seki, M. (2018): Single- and double-entry volume equations for Crimean pine stands in Kastamonu Regional Directorate of Forestry. – Turkish Journal of Forestry 19(1): 20-29.
- [35] Santi, E., Paloscia, S., Pettinato, S., Chirici, G., Mura, M., Maselli, F. (2015): Application of Neural Networks for the retrieval of forest woody volume from SAR multifrequency data at L and C bands. European Journal of Remote Sensing 48(1): 673-687.
- [36] Shataee, S. (2013): Forest attributes estimation using aerial laser scanner and TM data. Forest Systems 22(3): 484-496.
- [37] Shoemaker, D. A., Cropper, W. P. (2010): Application of remote sensing, an artificial neural network leaf area model, and a process-based simulation model to estimate carbon storage in Florida slash pine plantations. Journal of Forestry Research 21(2): 171-176.
- [38] URL-1 (2014): USGS Global Visualization Viewer. http://glovis.usgs.gov (accessed on 30 September 2014).
- [39] Wallner, A., Elatawneh, A., Schneider, T., Knoke, T. (2015): Estimation of forest structural information using RapidEye satellite data. – Forestry: An International Journal of Forest Research 88(1): 96-107.
- [40] Xie, S., Wang, W., Liu, Q., Meng, J., Zhao, T., Huang, G. (2017): Estimation of forest stand parameters using SPOT-5 satellite images and topographic information. – Preprints 2017100017.
- [41] Yang, S., Feng, Q., Liang, T., Liu, B., Zhang, W., Xie, H. (2018): Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. Remote Sensing of Environment 204: 448-455.
- [42] Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., Yu, S. (2016): Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. – Remote Sensing 8(6): 469.
- [43] Zheng, D., Rademacher, J., Chen, J., Crow, T., Bressee, M., Le Moine, J., Ryu, S. R. (2004): Estimating aboveground biomass using Landsat ETM+ data across a managed landscape in northern Wisconsin, USA. – Remote Sensing of Environment 93(3): 402-411.