



Assessment of the interactions among net primary productivity, leaf area index and stand parameters in pure Anatolian black pine stands: A case study from Türkiye

✉ Sinan BULUT*, ✉ Alkan GÜNLÜ and ✉ Sedat KELES

Department of Forest Management and Planning, Faculty of Forestry, Çankırı Karatekin University, 18200 Çankırı, Türkiye.

*Correspondence should be addressed to Sinan Bulut: sbulut@karatekin.edu.tr

Abstract

Aim of study: To examine the relationships between net primary productivity (NPP) and leaf area index (LAI) and modeling these parameters with stand parameters such as stand median diameter (dg), dominant height (htop), number of trees (N), stand basal area (BA) and stand volume (V).

Area of study: Pure Anatolian black pine (*Pinus nigra* J.F. Arnold) stands in semi-arid climatic conditions in the Black Sea backward region of Türkiye.

Materials and methods: In this study, the Carnegie-Ames-Stanford Approach model was used to calculate NPP; LAI, dg, h_{top}, N, BA and V were calculated in 180 sample plots. The relations between NPP and LAI with stand parameters were modeled using multiple regression analysis, support vector machines (SVM) and deep learning (DL) techniques. Relationships between NPP and LAI were investigated according to stand developmental stages and crown closure classes.

Main results: The highest level of relations was obtained in the stands containing the a-b developmental stages ($r=0.84$). The most successful model in modeling NPP with stand parameters was obtained by DL method (model $R^2=0.64$, test $R^2=0.51$). Although DL method had higher success in modeling LAI with stand parameters, SVM method was found to be more successful in terms of model-test fit, and modeling success in independent data set.

Research highlights: Grouping parameters affecting NPP and LAI increased the level of correlation between them. In modeling NPP and LAI in relation to stand parameters, machine learning algorithms performed better than linear approach. The overfitting problem can be eliminated substantially by including arguments such as early stopping, network reduction and regularization in the network structure.

Additional key words: machine learning algorithms; support vector machines; deep learning.

Abbreviations used: ANN (artificial neural networks); BA (basal area); CASA (Carnegie-Ames-Stanford Approach); dg (median diameter); DL (deep learning); FE (Forest Enterprise); FPU (Forest Planning Unit); htop (dominant height); LAI (leaf area index); MRA (multiple linear regression); N (number of trees); NPP (net primary productivity); RDF (Regional Directorate of Forestry); SVM (support vector machines); V (stand volume).

Citation: Bulut, S; Günlü, A; Keles, S (2023). Assessment of the interactions among net primary productivity, leaf area index and stand parameters in pure Anatolian black pine stands: A case study from Türkiye. Forest Systems, Volume 32, Issue 1, e003. <https://doi.org/10.5424/fs/2023321-19615>

Received: 02 Jun 2022. **Accepted:** 14 Feb 2023.

Copyright © 2023 CSIC. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Funding agencies/institutions	Project / Grant
Scientific Research Project Unit of Çankırı Karatekin University	OF061218D07

Competing interests: The authors have declared that no competing interests exist.

Introduction

In terrestrial ecosystems, forests have an important role in the global carbon cycle with the functions they undertake. The effect of forest ecosystems on this process is evaluated by the amount of carbon they store. The net amount of carbon added by forests is determined by net primary productivity (NPP). NPP is the difference between the total amount of carbon produced by plants and the amount they spend through autotrophic respiration. NPP researched on global and regional scales is valuable data for the interpretation of carbon emissions and climate change. NPP is also a very important parameter used in ecological-based models (Yan et al., 2021; Quinto-Mosquera et al., 2021; Wang & Yue, 2022). Interactions among variables in terrestrial ecosystems are affected by factors such as temperature and precipitation. NPP is a very successful indicator in evaluating the effects of climate change on vegetation in terrestrial ecosystems which is a complex system (Gower et al., 1999; Pan et al., 2015; Zhang et al., 2019).

The Carnegie-Ames-Stanford Approach (CASA) model, which is a process-based model, is frequently used for the estimation of NPP. This model has four main computational variables. These are: absorbed photosynthetically active radiation, light utilization efficiency, water, and temperature (Potter et al., 1993). The CASA model is remote sensing-based and flexible model. By enabling the use of remote sensing and geographic information systems technologies, databases can be created especially for large-scale areas, and calibration processes can be performed using data related to local areas. In addition, the explanatory variables used can be updated for different regions and different time periods, and calculations can be made in different variations. It allows calculation of NPP for global, national, and local scale areas in monthly, seasonal, and annual periods (Gülbeyaz, 2018).

The physiological processes controlling NPP are directly related to the leaf area index (LAI) and are frequently used in NPP models (Davidson, 2002). LAI is the amount of leaf area per unit area that is actively photosynthesizing. Two methods, direct and indirect, are used in the determination of LAI. The direct method is applied in the form of debris collection and determining the amount of leaves per unit area captured by leaf traps. This method is labor intensive and time consuming. The indirect method is the estimation of LAI using optical sensors (hemispherical photography). The indirect method is both easy to apply for large areas and advantageous in terms of time and cost (Chen et al., 1997; Jonckheere et al., 2004; Chen et al., 2022).

LAI, which is an important indicator of growth and productivity of forest ecosystems, is also an effective indicator of stand crown closure. Any change in the LAI as a result of natural disasters such as storms, snow and wind-induced overturns, drought and management interventions directly

affect growth and productivity. The fact that a forest stand has a high LAI value indicates that the site productivity is good. Therefore, it is widely used in modeling stand parameters and shows high correlations with stand characteristics such as increment, growth, and volume. It is also a variable that affects photosynthesis, interception, evapotranspiration, carbon, nutrient, and energy balance. With long-term monitoring of LAI, changes in forest ecosystem productivity and climate dynamics can be better interpreted (Bonan, 1993; Sprintsin et al., 2011; Yan et al., 2012; Günlü et al., 2017; Bahrami et al., 2022).

NPP and LAI are widely used to assess ecosystem structures and dynamics. Luo et al. (2004) investigated climatic and soil controls on NPP and LAI in the Tibetan Plateau. They reported that the climatic conditions highly limited the NPP, and increased precipitation generally increased LAI. Chen et al. (2020) explored the spatiotemporal NPP and LAI of natural and plantation forests in China. They quantified relative contributions and assessed the interannual variabilities-trends of NPP and LAI. Pan et al. (2021) investigated the relationships between NPP and LAI on a global scale. They calculated the NPP/LAI ratio of global vegetation to determine the net primary production per unit leaf area in interpretation of interactions. In addition to these studies, there are also studies in which NPP is modeled with LAI or stand parameters and the relationships between them are investigated. Gholz (1982), Scurlock et al. (2001) and Bulut et al. (2019) modeled NPP using LAI and stand parameters for different study areas and tree species.

In this study, we focused on pure Anatolian black pine stands, which shows an important spreading area in the region. According to the Turkish National Forest Inventory, the Anatolian black pine stands cover an area of 4,199,623 hectares and 18.31% of the total forest areas in Türkiye (GDF, 2021). As a result of the ecological resilience of the Anatolian black pine, it is one of the most widely used tree species for reforestation worldwide (Cenni et al., 1998). For instance, it is widely used in afforestation especially in semi-arid climatic regions in Türkiye. Also, it is the most common and dominant tree species in the study area. Thus, we selected pure Anatolian black pine stands for modelling the relationships between NPP-LAI, NPP-stand parameters, and LAI-stand parameters. The main objectives of the study were to: i) predict the NPP with the CASA model, ii) calculate the stand parameters and LAI values from ground measurements in pure Anatolian black pine stands, and iii) explore and model the relationships between NPP-LAI, NPP-stand parameters, and LAI-stand parameters, using multiple regression analysis (MRA), support vector machines (SVM) and deep learning (DL) methods. We hypothesized that there would be a non-linear relationship between NPP and LAI, and the SVM and DL modeling techniques will increase the model success levels compared to MRA modeling technique.

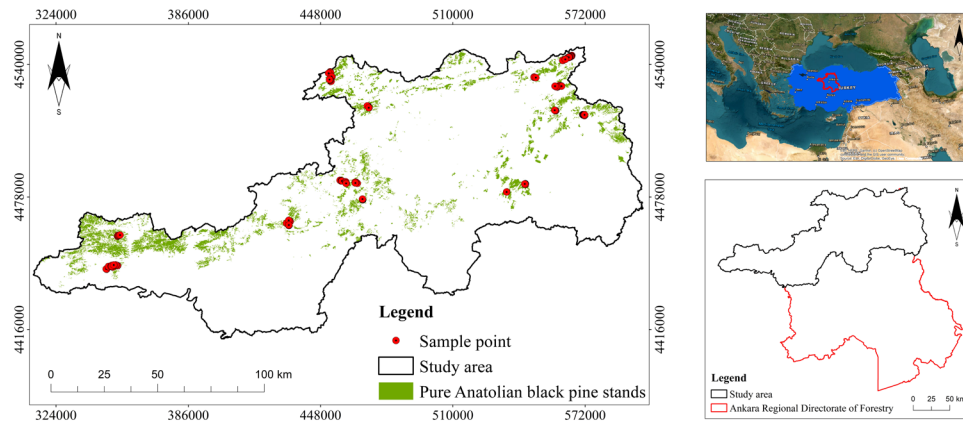


Figure 1. Location of the study area.

Material and methods

Study area

This study was carried out in pure Anatolian black pine stands, which are distributed within the borders of Ankara Regional Directorate of Forestry (RDF) (Fig. 1). Ankara RDF consists of 11 Forest Enterprises (FE) and 44 Forest Planning Units (FPU). The study area is located between 313096-598185 east longitudes and 4414251-4544558 north latitudes (WGS 1984 UTM Zone 36N). The total area of Ankara RDF is 4,563,247.07 ha.

The mean elevation is 1071 m, and the mean slope is 17%. The annual mean temperature is 11.6 °C, and the mean annual total precipitation amount, 399.75 mm. The month with the highest mean monthly precipitation is May (70.85 mm), and the lowest month is July (20.07 mm). Anatolian black pine stands constitute the majority in the study area among the other tree species. Anatolian black pine is also frequently used in plantation areas in the region, since it is well adapted to the semi-arid climatic conditions. The stands under the dominance of Anatolian black pine in the study area are 213,868.56 ha, and pure Anatolian black pine stands cover an area of 136,985.96 ha. Other tree species in the study area are *Pinus sylvestris*, *Pinus brutia*, *Abies nordmanniana*, *Fagus orientalis*, *Carpinus betulus*, *Quercus cerris*, *Quercus infectoria*, *Quercus frainetto*, *Quercus robur*, *Quercus petraea*, *Cedrus libani*, *Juniperus* sp. and *Populus tremula*.

Field measurement

The field work was carried out in FEs where pure Anatolian black pine stands are distributed in Ankara RDF. The spatial distribution of sample plot numbers was arranged by proportioning the areas of pure Anatolian black pine covered in the relevant FPUs. Accordingly, a total of 180 sample plots were taken. Taking into account the stand crown closure in the field studies, sample plots were taken

in the form of a circle, 400 m² in 3 (71-100%), 600 m² in 2 (41-70%) and 800 m² in 1 (11-40%) closed stands. Diameter values from the breast height level were measured bilaterally in trees with a diameter of less than 8 cm in young stands, and 8 cm and thicker in other stands, and the average values were taken. To determine the stand mean age, age was measured on 4-5 trees close to the stand mean diameter. Tree heights were measured on the dominant trees according to the method of 100 trees/ha to determine the site productivity. As a result of the inventory study, stand median diameter (dg), dominant height (h_{top}), number of trees (N), stand basal area (BA) and stand volume (V) values were calculated for each sample plot (Table 1). In the calculation of stand volume, the diameter and height-dependent double-entry volume equation (Eq. 1) developed for pure and natural Anatolian black pine stands distributed in Ankara RDF was used (Ercanlı, 2020):

$$v = 0.000097 \times d^{1.740878} \times h^{1.005521} \quad (1)$$

Calculation of net primary productivity

The NPP can be determined directly by making local measurements or can be predicted with statistical models or process-based models in forest ecosystems. In this study, CASA (Potter et al., 1993), which is a process-based model (Eqs. 2-9), was used for the estimation of NPP:

$$NPP = \varepsilon_n \times APAR \quad (2)$$

$$\varepsilon_n = \varepsilon_{max} \times f(T) \times f(W) \quad (3)$$

$$f(T) = \frac{(T-T_{min}) \times (T-T_{max})}{(T-T_{min}) \times (T-T_{max}) - (T-T_{opt})^2} \quad (4)$$

$$f(W) = \frac{1+LSWI}{1+LSWI_{max}} \quad (5)$$

$$LSWI = \frac{P_{nir} - P_{swir}}{P_{nir} + P_{swir}} \quad (6)$$

Table 1. Stand parameters, equations, and abbreviations.

Stand parameter	Equation	Abbreviation
dg	$\bar{d}_g = \sqrt{\frac{\bar{d}_1^2 + \bar{d}_2^2 + \dots + \bar{d}_n^2}{n}}$	dg= stand median diameter (cm) d ₁ ...d _n = measured tree diameter
h _{top}	$h_{top} = \frac{h_1 + h_2 + \dots + h_n}{n}$	n=number of measured trees h _{top} = stand dominant height (m) h ₁ ...h _n = measured tree height
N	$N = \frac{10000}{SSA} \times n$	N= number of trees/ha SSA= size of sample plot
BA	$BA = \frac{10000}{SSA} \times \left(\frac{\pi}{4} \times d^2\right)$	BA= stand basal area (m ² /ha) d= measured tree diameter
V	$V = \frac{10000}{SSA} \times \sum_{i=1}^n v_i$	V= stand volume (m ³ /ha) v _i = volume of tree

$$APAR = FPAR \times PAR \quad (7)$$

$$FPAR = \frac{(NDVI - NDVI_{min}) \times 0.95}{(NDVI_{max} - NDVI_{min})} + 0.05 \quad (8)$$

$$PAR = R_s \times 0.50 \quad (9)$$

where ϵ_n is light use efficiency; PAR is the photosynthetically active radiation; APAR is absorbed PAR; ϵ_{max} represents the maximum light use efficiency, and the ϵ_{max} values of young, middle-aged, and old stands used were 0.72, 0.57 and 0.52, respectively (Li & Zhou, 2015); $f(T)$ is the temperature function; T is the atmospheric temperature; T_{min} , T_{max} and T_{opt} are minimum, maximum and optimal temperatures for photosynthetic activities, respectively; $f(W)$ is the precipitation function; LSWI is the land surface water index; LSWI_{max} is the maximum of LSWI values; the p_{nir} and p_{swir} variables represent the surface reflectance of the NIR (near-infrared) and SWIR (short-wavelength infrared) bands in Sentinel-2 satellite image, respectively; FPAR is the fraction of PAR absorbed by the vegetation cover; NDVI is the normalized difference vegetation index; NDVI_{max} and NDVI_{min} are the maximum and minimum of NDVI, respectively; and R_s denotes total solar radiation (MJ m⁻² day⁻¹). At the end of this process, maps with a spatial resolution of 10 m were produced for the NPP, which was calculated on a monthly (gC m⁻² month⁻¹) and annual (gC m⁻² year⁻¹) scale for the study area.

LAI values

A camera (Canon EOS 50 D) with a fisheye (Sigma 8 mm Fisheye lens) was used to determine the LAI value in each sample plot. Hemispherical photographs covering an area of 180° were taken. These photographs were

taken from 5 different points to represent the sample plot, from a height of approximately 1.5 m, and towards the top of the stand. However, care was taken not to take photos when direct sunlight comes in. Then, the digitized photographs were analyzed using HemiView, H. version: 2.1 SR4 software program. By taking the average of the values obtained as a result of the analysis, the LAI values for each sample plot were calculated.

Modeling NPP and LAI with stand parameters

NPP was calculated using the CASA model, stand parameters such as LAI, dg, h_{top}, N, BA, and V were obtained from the field inventory studies, and datasets were made ready for the modeling process. To test the created models, the data sets were divided into two groups: 70% were reserved for the model (126 items) and 30% for the test (54 items). MRA, SVM and DL modeling techniques were applied to the separated data sets. The modeling process using these techniques was carried out in three stages. In the first stage, normality tests of the NPP, LAI, dg, h_{top}, N, BA, and V data sets were performed using Kolmogorov-Smirnov normality test ($p > 0.05$). In the second stage, the methods were applied using the model data set. In the third stage, the test data were run in the models created with the model data set.

Modeling techniques

The MRA method was performed to determine the linear relationships among NPP, LAI and stand parameters. IBM SPSS 23 software was used for the implementation of the multiple linear regression (MLR) method with stepwise procedure.

The SVM method is based on structural risk minimization approach and statistical learning theory. The purpose of this method is to obtain the optimal hyperplane that can separate the existing classes from each other. There are four different kernel functions such as linear, radial, polynomial, and sigmoid that can be used to create these limits (Zhao et al., 2019). The SVM method was applied in “e1071” package within the “R” software (R Development Core Team, 2014; Meyer et al., 2015).

The DL method is one of the machine learning algorithms and has a more complex structure than artificial neural networks (ANN). While ANN are implemented with one hidden layer, the DL method is implemented using more than one hidden layer. The flowchart of the DL method consists of the input layer, hidden layers, and output layer. Inputs are first transferred to the input layer. Each node provides an output value through its activation function. The outputs of the input layer are used as inputs in the next hidden layer. The neuron numbers in the input layer are the number of inputs in the processed data, and the neuron numbers in the output layer are equal to the number of outputs associated with inputs. The DL method was applied in “H2O” package within the “R” software (R Development Core Team, 2014; Aiello et al., 2015). Early stopping, network reduction and regularization arguments were used in “H2O” package. Early stopping argument was used so that the learning speed does not slow down. Network reduction argument was used to avoid noise learning and to reduce noise by eliminating irrelevant data. The regularization argument ensures that only the useful variable is added the model. On the other hand, it reduces the weights of the variables with less effect.

Model evaluation criteria

Different criteria were used to compare the observation data through the inventory study and the estimation results using modeling techniques. The performance measures used were correlation coefficient (r), coefficient of determination (R^2), and root-mean squared error (RMSE).

$$r = \frac{\sum XY - (\sum X)(\sum Y)/n}{\sqrt{(\sum X^2 - (\sum X)^2/n)(\sum Y^2 - (\sum Y)^2/n)}} \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_g - y_t)^2}{\sum_{i=1}^n (y_g - y_{og})^2} \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (h_i)^2} \quad (12)$$

where h_i , error; X , primary variable; Y , secondary variable; y_g , observation; y_t , estimation; y_{og} , mean observation; and n , number of samples. For high model success, r and R^2 values are expected to be high; RMSE value is expected to be small.

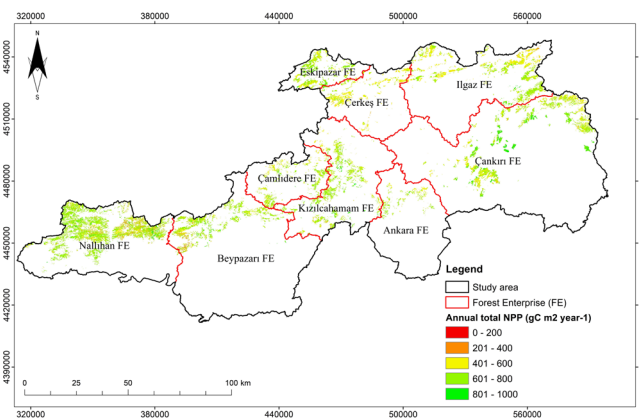


Figure 2. Annual total net primary productivity (NPP) map of the pure Anatolian black pine stands.

Results

Net primary productivity

NPP was calculated monthly and annually for 2019 using the CASA model (Fig. 2). The minimum annual NPP amount of pure Anatolian black pine stands in the sample plots was 284 gC m^{-2} and the maximum amount was 960 gC m^{-2} . The annual average NPP amount of the stands was calculated as 602 gC m^{-2} .

Spatial distribution of annual total NPP amounts of Anatolian black pine stands was mapped (Fig. 2). Average NPP amounts of FEs were $664 \text{ gC m}^{-2} \text{ year}^{-1}$ in Nallıhan, $651 \text{ gC m}^{-2} \text{ year}^{-1}$ in Beypazarı, $623 \text{ gC m}^{-2} \text{ year}^{-1}$ in Çamlıdere, $646 \text{ gC m}^{-2} \text{ year}^{-1}$ in Kızılcahamam, $635 \text{ gC m}^{-2} \text{ year}^{-1}$ in Eskişehir, $562 \text{ gC m}^{-2} \text{ year}^{-1}$ in Çerkeş, $594 \text{ gC m}^{-2} \text{ year}^{-1}$ in Ilgaz, $621 \text{ gC m}^{-2} \text{ year}^{-1}$ in Çankırı and $630 \text{ gC m}^{-2} \text{ year}^{-1}$ in Ankara. The highest average NPP in the Regional Directorate was found in Nallıhan FE ($664 \text{ gC m}^{-2} \text{ year}^{-1}$). It should be noted that these data only belong to pure Anatolian black pine stands. It would be inappropriate to make a more efficient/inefficient interpretation among FEs based on this information alone. There are many factors that can cause differences in NPP, and stand characteristics, topographic and edaphic factors in FEs are the most important ones. In addition, the timely and correct implementation of the silvicultural interventions on the stands can greatly contribute the productivity in the stands.

Leaf area index

Correlation levels of dg , h_{top} , N , BA , V and LAI for sample plots were -0.30 , 0.11 , 0.58 , 0.35 and 0.36 , respectively. LAI was grouped according to developmental stages and crown closure classes, and the relationships between NPP and these classes were examined. When classified according to their developmental stages, the highest average

Table 2. Relationships between net primary productivity (NPP) and leaf area index (LAI) according to developmental stages and crown closure classes.

	Class	N	Mean LAI (m ² m ⁻²)	Mean NPP (gC m ⁻²)	NPP model ^[a]	R ²	r
		180	1.254	601.610	Eq. 13	0.22	0.47
Developmental stage							
a	0-7.9 cm	2	2.729	884.995	Eq. 14	0.70	0.84
b	8-19.9 cm	21	1.276	685.000			
c	20-35.9 cm	88	1.202	617.822	Eq. 15	0.52	0.72
d	>36 cm	69	1.012	547.341	Eq. 16	0.20	0.45
Crown closure							
1	11-40%	26	0.654	569.744	Eq. 17	0.38	0.62
2	41-70%	68	0.974	572.195	Eq. 18	0.21	0.45
3	71-100%	86	1.494	660.714	Eq. 19	0.37	0.61

^[a] NPP model equations:

$$\text{NPP} = 543.949 + (36.119 \times \text{LAI}^2) \quad (13)$$

$$\text{NPP} = 1093.488 - \left(649.391 \times \frac{1}{\text{LAI}}\right) + \left(183.788 \times \frac{1}{\text{LAI}^2}\right) \quad (14)$$

$$\text{NPP} = 894.244 - \left(322.221 \times \frac{1}{\text{LAI}}\right) - (350.944 \times \ln \text{LAI}) + (55.021 \times \text{LAI}^2) \quad (15)$$

$$\text{NPP} = 456.589 + (91.165 \times \text{LAI}) \quad (16)$$

$$\text{NPP} = 521.380 + (128.738 \times \text{LAI}^2) \quad (17)$$

$$\text{NPP} = 388.628 + \left(171.250 \times \frac{1}{\text{LAI}^2}\right) + (569.962 \times \ln \text{LAI}) \quad (18)$$

$$\text{NPP} = 639.162 - \left(64.486 \times \frac{1}{\text{LAI}^2}\right) + (24.652 \times \text{LAI}^2) \quad (19)$$

LAI value (2.729 m² m⁻²) was found in the young stands (*a* developmental stages). The crown closure classes are determined according to the coverage ratio of the top roof to the ground surface, the increase in the amount of overlap is due to the increase in the amount of needles per unit area. Therefore, the highest average LAI value was determined in the 3rd (1.494 m² m⁻²) and the lowest average value was determined in the 1st (0.654 m² m⁻²) crown closure class.

Relationships among NPP, LAI, developmental stages, and crown closure classes

A parabolic relationship was determined between NPP and LAI (Fig. 3). The maximum value of NPP determined for pure Anatolian black pine stands was 960.038 gC m⁻² year⁻¹. NPP reached its maximum value at the breaking point where the LAI was 2.456 m² m⁻². A decreasing trend in NPP was determined when the LAI value exceeded 2.456 m² m⁻². While the relationship between NPP and LAI was positive until the breaking point ($r = 0.42$), it was found to be negative after the breaking point ($r = 0.64$).

While examining the relationships between NPP and LAI, they were evaluated separately according to developmental stages and crown closure classes (Table 2). The LAI used as an independent variable was included in the analysis by applying transformations. For this purpose, the transformations of LAI², 1/LAI, 1/LAI², lnLAI, logLAI, $\sqrt{\text{LAI}}$ and $1/\sqrt{\text{LAI}}$ were applied to the data sets of LAI. Since there were two stands belonging to the “a” developmental stage, they were combined with the “b” developmental stage in the analysis process. The highest mean LAI and NPP were found in “a” (young stands) developmental stage and 3-closed stands. The lowest NPP and LAI were at “d” developmental stage and 1-closed stands.

The model with the highest relationship between NPP and LAI was produced from stands belonging to “a” and “b” developmental stages (Fig. 4). In the scatterplots, the slope of the red line was expected to be close to 45 degrees, and the black points were expected to spread close to red line. When the scatterplots of NPP estimations vs CASA model NPP values were examined, it was seen that the best spread was obtained for “a-b” developmental stages ($r = 0.84$). Comparing the whole data set and the re-

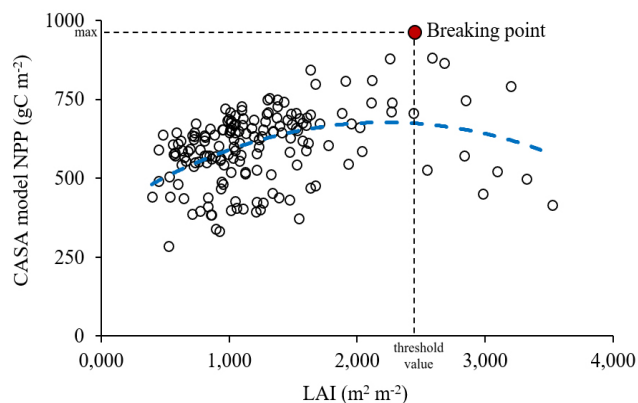


Figure 3. Relationship between net primary productivity (NPP) and leaf area index (LAI) including graphic illustrating the maximum NPP value on LAI threshold value.

relationships that were divided into different groups, it was observed that certain classes gave higher correlations than the total data set. It was also observed that the relations between LAI and NPP were stronger in the groups devoted for “a-b” and “c” developmental stages and “1” and “3” crown closure stands.

Modeling of NPP and LAI with stand parameters

The terrestrial data sets obtained from the inventory study consist of the calculated stand parameters (dg, h_{top} , N, BA, and V). NPP and LAI were modeled using stand parameters. In addition, the NPP-LAI model was created to determine the level of modeling of NPP using only LAI. For modelling, 70% of the data sets were devoted for the

model and 30% for testing. In the modeling phase, MRA, SVM and DL techniques were used.

As the first step, the MRA method was applied. Thus, the independent variables that were significant in the model to be created for the dependent variable were determined. Stand parameters included in the model created for NPP are h_{top} , N, BA, and V (Fig. 5, Eq. 20). The coefficient of determination of the resulting model was found to be 0.50, and when the test data were run in the same model, this value decreased to 0.30 (Fig. 5). A slight decrease in test success was expected, because the model was tested with a separate data set that was never encountered during the modeling phase. The test error of the NPP model was found to be 93.757 $gC\ m^{-2}\ year^{-1}$. In the model created for LAI, dg, N and BA were entered as significant parameters (Fig. 5, Eq. 21). Although the model success was not high, there was no difference in success between the model and the test (model $R^2 = 0.48$, test $R^2 = 0.47$). This shows that the model can be applied to different data according to the level of success and that it has a stable structure. Although there was no difference between the model and test in the NPP model created with LAI (Fig. 5, Eq. 22), the model success was observed at a low level ($R^2 = 0.20$). The error value of the NPP-LAI model created was 99.807 $gC\ m^{-2}\ year^{-1}$.

Significant independent variables were determined with the MRA models created. Relationships between dependent and independent variables were modeled using SVM and DL methods. In modeling NPP with stand parameters, the performance of the SVM method was found to be higher than the MRA method (Fig. 6). When the similarities between the predictions and observations of the models were compared, the similarity of the MRA method was 0.71, while this value was 0.79 for the SVM method. As a result of the test, SVM was more successful ($r = 0.66$). A similar situation was observed in terms of

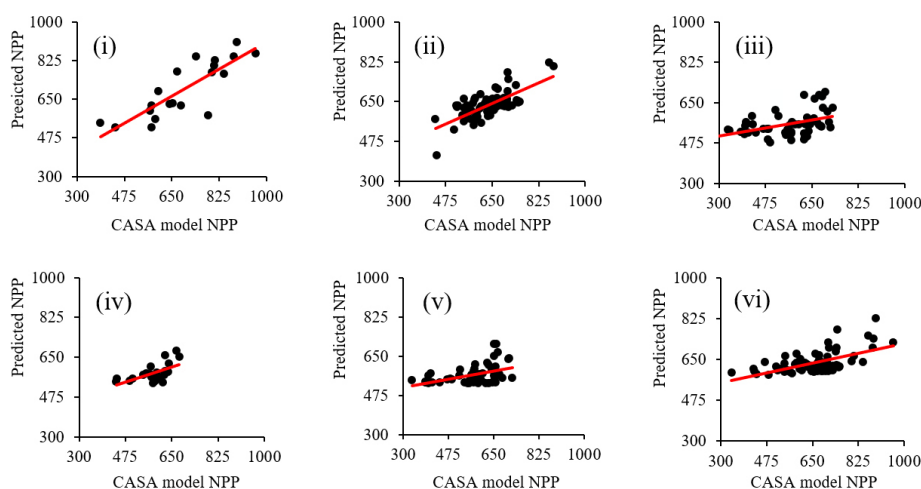


Figure 4. Scatterplots of NPP ($g\ C\ m^{-2}\ year^{-1}$) based on i) a-b developmental stages, ii) c developmental stage, iii) d developmental stage, iv) 1 crown closure, v) 2 crown closure, and vi) 3 crown closure.

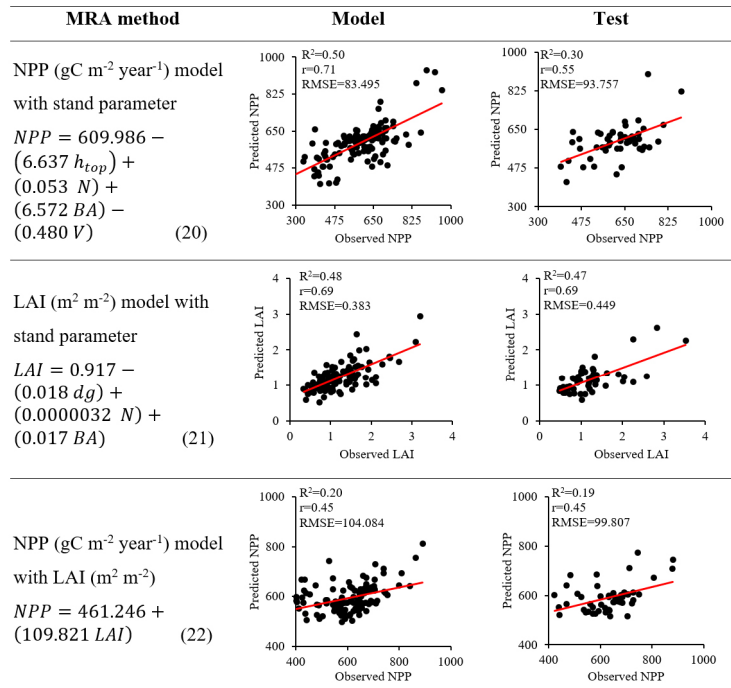


Figure 5. Observed and predicted values obtained from NPP-stand parameters, LAI-stand parameters and NPP-LAI models using the MRA method. Fitted model equations (20, 21, 22) are shown.

success in modeling LAI. While the MRA method modeled the LAI with an error of $0.383 \text{ m}^2 \text{ m}^{-2}$ for the model data set, the error of the SVM method was 0.322 m^2

m^2 . In addition, no difference was observed between the model and test success of the SVM method (Model $R^2 = 0.64$, Test $R^2 = 0.64$). In the SVM model between NPP

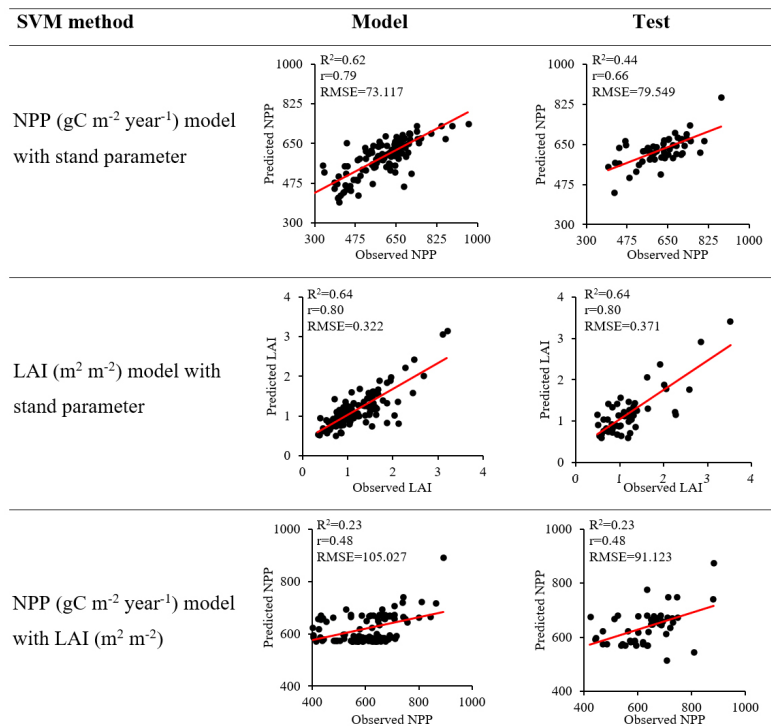


Figure 6. Observed and predicted values obtained from NPP-stand parameters, LAI-stand parameters and NPP-LAI models using the SVM method.

and LAI, a notable increase in success was not obtained compared to the MRA method.

With the independent variables taken from the MRA models, the DL method was applied for different neuron and layer numbers (Fig. 7). The layers were tried as 3, 7 and 10, and the number of neurons starting from 10 up to 100 in ten intervals. Thus, 30 model network trainings were conducted for each model. Among the models developed, the network with the closest model-test successes and the highest test success was selected and marked with black circles on the success graphs in Fig. 7. Findings of the selected best fit model and other methods were compared (Fig. 8). In the estimations made, it was expected that there will not be a high difference between the model and the test. However, due to the overfitting problem frequently encountered in the DL method, the model showed high success, but when the same model was tested with an independent data set, the success went down to very low levels. Model estimations trained with stand parameters of NPP achieved high success ($r = 0.95$). However, test success was low and not equivalent to model success ($r = 0.54$). As seen in the success graphs prepared, the difference between model and test success levels for NPP-stand parameters was quite large (model $R^2 = 0.90$, test $R^2 = 0.29$). For LAI-stand parameters, this difference was lower (model $R^2 = 0.72$, test $R^2 = 0.46$), while for NPP-LAI success results were similar (model $R^2 = 0.24$, test $R^2 = 0.24$).

In the applied DL method, overfitting problem was observed especially in the models of NPP and LAI created with stand parameters, and sequential predictions were observed in the modeling of NPP with LAI. To eliminate the problems that occurred, the method was rerun by adding early stopping, network reduction and regularization

arguments. By using these arguments, we aimed to learn the model data set of the network, to achieve the desired level of modeling success in independent data groups and to create a model network with high representation ability. The findings obtained by adding the relevant arguments are presented in Fig. 9. The overfitting problems encountered in the DL method were improved with the arguments arranged for the model network. In modeling NPP with stand parameters, while the coefficients of determination for model-test were 0.90-0.29, these values increased to 0.64-0.51 with the regulation arguments (Fig. 9). As a result of these findings, a better performance was obtained than with the SVM method. The coefficients of determination for the model-test, which were 0.72-0.46 in the model network applied for the stand parameters with LAI, also reached the level of 0.68-0.53. For the NPP-stand parameters model, DL performed better than the SVM method. However, the SVM method was found to be more successful for the LAI-stand parameters model because its success in the test data set was higher (SVM test $R^2 = 0.64$, DL test $R^2 = 0.53$). But there was no difference in the performance criteria obtained for NPP and LAI.

Discussion

Relationships of NPP and LAI

The relationships between NPP and LAI are highly affected by environmental and climatic conditions (Pan et al., 2015; Yan et al., 2021). Subtropical forests have a high NPP but a low LAI (Luo et al., 2004). NPP/LAI ratio is higher in arid than humid regions due to high net produc-

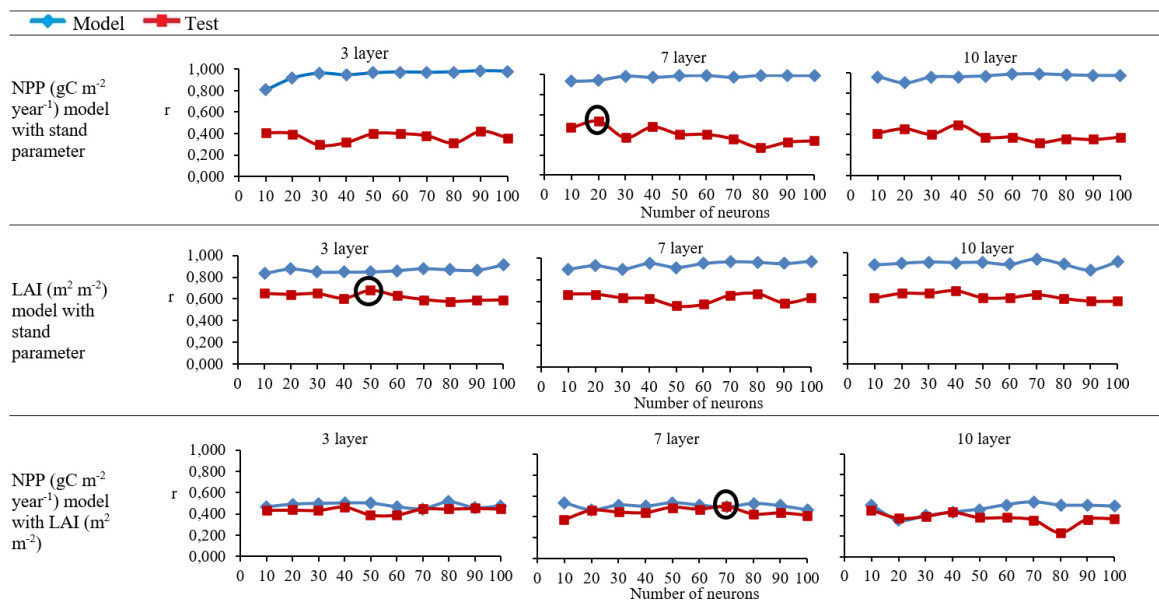


Figure 7. Correlation changes of NPP-stand parameters, LAI-stand parameters and NPP-LAI models and tests according to the changing number of neurons and layers using the DL method.

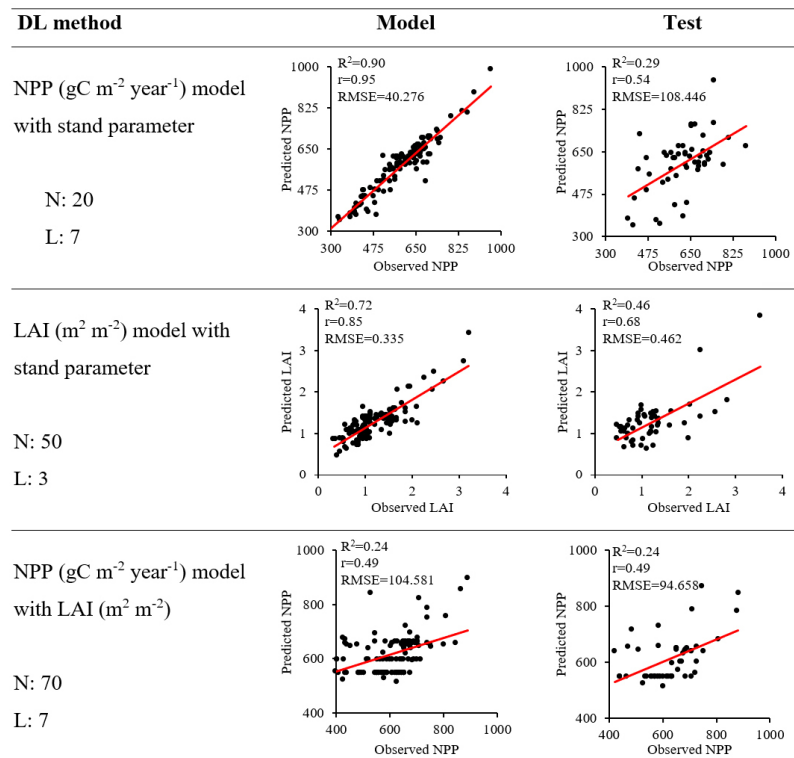


Figure 8. Observed and predicted values obtained from NPP-stand parameters, LAI-stand parameters and NPP-LAI models using the DL method. N: number of neurons, L: number of layers.

tion and small leaf surfaces (Pan et al., 2021). These conditions also affect the function of the relationship between NPP and LAI. There are studies in which the NPP-LAI relationship is positive in different study areas (Gholz, 1982; Scurlock et al., 2001; Luo et al., 2004; Kushida et al., 2007). Pan et al. (2021) investigated the relationship between NPP and LAI on a global scale and found that there is a parabolic relationship between NPP and LAI. They reported that parabolic relationships between NPP and LAI are rarely in regions with poor site index conditions. Especially, it has been reported that relation of NPP and LAI is parabolic in humid areas and positively related in dryland areas. In this study, sample plot data up to the breaking point made up most of the data set, and the relationship between NPP-LAI was positive in this section. NPP tended to decrease especially in sample plots where LAI was high after the breaking point. This caused the overall relationship between NPP and LAI to be parabolic. This study was carried out in a semi-arid region backward the Black Sea region. Therefore, it can be said that the parabolic relationship between NPP and LAI is due to the semi-arid region.

Evaluation of modeling findings on net primary productivity

NPP was modeled using stand parameters with MRA method. Significant independent variables in this model

were h_{top} , N, BA, and V. The coefficients of determination for the model-test were found to be 0.50-0.30. The coefficients of determination for the model-test for the SVM and DL methods were 0.62-0.44 and 0.64-0.51, respectively. In addition to stand parameters, NPP was also modeled with LAI. The coefficient of determination for the model-test was obtained as 0.24 using DL method. The change in the level of success was controlled by testing the regulation arguments in the DL technique. However, no change was observed in model-test successes as a result of the added arguments (model-test $R^2 = 0.24$). Bulut et al. (2019) modeled the NPP with stand volume, volume increment and stand carbon separately using the MRA method in their study. The coefficients of determination obtained by stand volume and volume increment were 0.78 and 0.75, respectively. The highest success was obtained with stand carbon ($R^2 = 0.85$). Since the independent variables used represent the accumulation in the stand, the correlation levels were obtained at a high rate. Since NPP gives the net produced carbon amount in the stands, it is expected that the highest level of relationship between them will be determined with stand carbon. In this study, stand parameters such as h_{top} , N, BA, and V were used to model NPP. Although a parameter characterizing the accumulation in the stand, such as V, was used in this study, stand parameters were able to predict only 50% of the explainable variance in NPP modeling. Differences between model performances may be due to the different characteristics of the datasets.

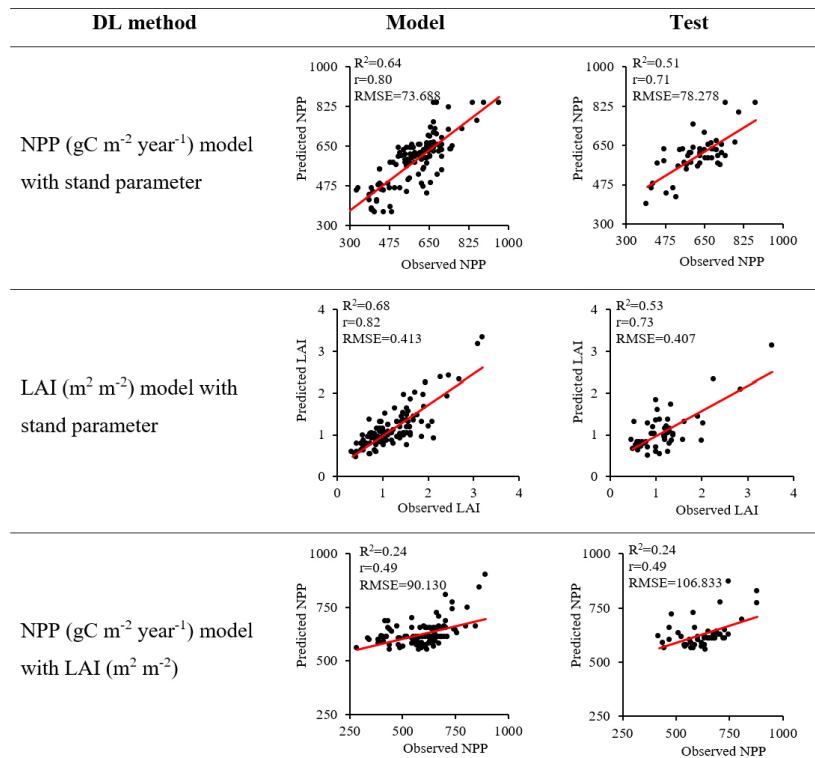


Figure 9. Observed and predicted values obtained from NPP-stand parameters, LAI-stand parameters and NPP-LAI models run by adding arguments to the DL method.

Gholz (1982) examined the relations of NPP with BA and LAI in *Tsuga heterophylla*-*Picea sitchensis*, *Tsuga mertensiana*, *Pinus ponderosa* and *Juniperus occidentalis* stands. The coefficient of determination of the MRA model between NPP and BA was found to be 0.95. This value was 0.96 between NPP and LAI. The findings show that there is a high level of linear relations between the parameters. There are quite high differences between our study and Gholz's (1982). In our study, MRA method and stand parameters (h_{top} , N, BA, and V) were used in the modeling of NPP and a determination coefficient of 0.50 was obtained. When LAI was used to model the NPP, a determination coefficient of 0.20 was found. The different results between Gholz (1982) and our study may be due to the number of samples used. As the number of data increases, the error in the model is expected to decrease (Sterenczak et al., 2018). Gholz (1982) examined the relationships with 8 sample data taken from 12 different main vegetation types, and the model success was high. However, the low number of data also reduces the representativeness of the model.

The performances of parametric MRA method and non-parametric SVM and DL methods in modeling NPP with LAI were found to be close to each other, because the correlation between NPP and LAI was not high ($r = 0.47$). Artificial intelligence techniques may be insufficient in terms of modeling success among data that do not have a significant relationship and trend (Aertsen et al., 2010). This may be due to the inclusion of all data from sample plots

that different stand structures in the modeling (Çil, 2014). To see this effect, the data were modeled separately according to stand developmental stages and crown closure classes. When classification was made according to stand developmental stages and crown closure, which have a great effect on NPP and LAI, the strength of the relations between the two increased and affected the success of the model. The highest correlation in the classes devoted to stand developmental stages was obtained for the stands at the "a-b" developmental stages ($R^2 = 0.70$, $r = 0.84$). In the crown closure classes, although the highest relationships were close to each other, 1 ($R^2 = 0.38$, $r = 0.62$) and 3 ($R^2 = 0.37$, $r = 0.61$) closed stands gave better results. As a result, it was observed that the success rate increases when the parameters that do not have a direct strong relationship with each other are grouped in terms of the factors that are decisive for them. It is important for such groupings that the number of sample plots in the inventory studies is similar in terms of the study-specific criteria. Thus, each class can have enough data for modelling. In particular, the fact that the number of data belonging to the classes is separable for model-testing also allows the application of machine learning algorithms.

Evaluation of modeling findings on LAI

In the modeling in which stand parameters were used as the independent variable for LAI, the significant pa-

rameters selected according to the MRA method were dg, N and BA. Among the applied techniques, the SVM method showed the highest test success. The error of the estimation made by the DL method is $0.407 \text{ m}^2 \text{ m}^{-2}$ after adding arguments, while the error of the SVM method was $0.371 \text{ m}^2 \text{ m}^{-2}$. Considering the model performances, the DL method (Model $R^2 = 0.68$) is more successful than the SVM method (Model $R^2 = 0.64$). Khosravi et al. (2012) made an estimation of LAI for oak forests using stand parameters. The coefficients of determination obtained were 0.36 with dg, 0.36 with BA, and 0.45 with mean stand height (h). Ercanlı et al. (2018) determined the LAI for 108 sample plots taken from pure Anatolian black pine stands by hemispherical photographic acquisition. Stand parameters were also measured in the sample plots and the LAI was modeled using MRA and ANN techniques. The coefficients of determination obtained for the MRA, and ANN methods are 0.5431 and 0.6392, respectively. With the ANN method, the estimation was improved, and the error amount was reduced from $0.3935 \text{ m}^2 \text{ m}^{-2}$ to $0.3497 \text{ m}^2 \text{ m}^{-2}$.

In our study, coefficient of determination of 0.48 was obtained for LAI model with the MRA method. Coefficients of determination of LAI model by using machine learning algorithms such as SVM and DL methods were 0.64 and 0.68, respectively. It has been observed that more accurate predictions can be made with nonlinear techniques. Here, if the aim is to analyze linear relationships and interpret simple relationships, MRA method may be appropriate. Linear models produced with easy-to-measure parameters that can be obtained from stands are considered useful in estimating LAI. In cases where model success is important in estimating LAI, non-parametric techniques such as SVM and DL can be used. Machine learning methods are thought to be useful in making predictions and mappings with higher success rates and usable levels. In addition to the applied techniques, the independent variables used also affect the model success and representation.

In our study, the relationships between LAI and stand characteristics were investigated. Considering that the stand age will increase as the diameter increases, there is a decrease in the LAI values inversely proportional to the stand age. In young stands, an increase in LAI may occur in the early years during the canopy formation process (Jagodzinski & Kalucka, 2008). However, a gradual decrease in LAI occurs with the increase of stand age (Clough et al., 2000; Kushida et al., 2007). However, it is insufficient to use only stand features in linear models to be produced, because besides stand characteristics, all edaphic conditions affect allometric relations with LAI. Therefore, the use of soil properties, which are edaphic factors, as well as stand parameters in linear models to be produced for LAI estimation, contributes to the success and representativeness of the model (Gower et al., 1999; Asner et al., 2003).

Evaluation of overall findings

Model-test successes were compared for the overfitting problem, which is frequently encountered in SVM and DL methods. According to the findings we obtained, the model-test success levels for the SVM method were found to be compatible with each other. NPP-stand parameters (Model $R^2 = 0.90$, Test $R^2 = 0.29$) and LAI-stand parameters (Model $R^2 = 0.72$, Test $R^2 = 0.46$) models were found to have an overfitting problem due to the difference between model and test success. Some of the reasons for this problem are the insufficient training data set, poor representation ability and high outlier data rate (Ying, 2019). To eliminate the problem, early stopping (Raskutti et al., 2014), network reduction (Bramer, 2002) and regularization (Warde-Farley et al., 2013) arguments were added to the model (Candel et al., 2016; MathWorks, 2019). As a result of this process, the effects of the overfitting problem were largely eliminated. Findings obtained were model $R^2 = 0.64$ and test $R^2 = 0.51$ for NPP-stand parameters, model $R^2 = 0.68$ and test $R^2 = 0.53$ for LAI-stand parameters.

In the DL method, the trend of the relationship between dependent and independent variables is determined during the training process of the network. When this trend which is established with the existing data reflects the reality, the model accuracy rate and the representation ability can be sufficient. Thus, when the trained model network is applied to independent data groups, satisfactory results can be obtained. In this case, the characteristics of the data sets to be used for modeling are of great importance. There is no obvious trend in the modeling of parameters that do not have any organic link between them (Aertsen et al., 2010). It is very difficult to generalize the model network created in this way. Thus, a consistent estimation cannot be made with the independent data run in the model network created. Although the success in the training process is high, the problem of overfitting, which is frequently encountered in machine learning algorithms, arises. As a result, it causes inconsistent estimates that are not representative and do not allow the application of independent data sets (Ying, 2019).

Overfitting problem was encountered in the models produced for NPP and LAI. NPP is a parameter that is estimated using many data groups (photosynthetic data, meteorological data, remote sensing data), while LAI is a parameter that can vary greatly according to stand structures. Since there is no direct relationship and connection between the independent variables, a clear trend could be formed in the trained network. Thus, the test data also fails in the generated model network. The application of these methods in parameters such as age and site index (Ercanlı et al., 2018), which have a clear relationship, can reduce the incidence of this problem and increase model-test successes to higher levels.

With the application and development of artificial intelligence techniques in different fields, modeling studies with high success rates are carried out. Although these

techniques have high success potential, they are not simple and user-friendly (Aertsen et al., 2010). In artificial intelligence studies, it is very difficult and time-consuming to find the necessary arguments and the most appropriate number of neurons and layers for the training process of the model network. In addition, the fact that it has uninterpretable weights originating from the hidden layer and nonlinear activation functions limits the interpretation of ecological processes (Peng & Wen, 1999). As a solution to this problem, there are techniques in the literature to determine the contribution of independent variables (Maier & Dandy, 2000). The use of simpler, transparent, and interpretable techniques can be useful and explanatory to reveal the relationships in the parameters related to the productivity and development of stands such as NPP and LAI.

Conclusions

In this study, the relationships between NPP-LAI, NPP-stand parameters and LAI-stand parameters using MRA, SVM and DL methods in pure Anatolian black pine stands were investigated. A non-linear parabolic relationship was found between NPP and LAI. Compared to the MRA method, the SVM and DL were found to be more successful in modeling NPP-LAI, NPP-stand parameters, and LAI-stand parameters. These models, especially developed with SVM and DL modeling techniques, may considerably contribute to the understanding of stand structure and dynamics. As a result, they will contribute to sustainable forest management planning, developing forest growth and yield models, and determine the most suitable silvicultural prescriptions to the forest stands.

Acknowledgements

The authors thank the Turkish General Directorate of Forestry for the data provision. Special thanks for their helping to Assoc. Prof. Onur Şatır in the statistical process and Dr. Ferhat Bolat in the field work.

Authors' contributions

Conceptualization: S. Bulut, A. Günlü, S. Keles

Data curation: S. Bulut

Formal analysis: S. Bulut, A. Günlü

Funding acquisition: S. Bulut, A. Günlü

Investigation: S. Bulut

Methodology: S. Bulut, A. Günlü

Project administration: S. Bulut, A. Günlü

Resources: S. Bulut, A. Günlü

Software: S. Bulut

Supervision: S. Bulut, A. Günlü, S. Keles

Validation: S. Bulut

Visualization: S. Bulut

Writing – original draft: S. Bulut, A. Günlü, S. Keles

Writing – review & editing: S. Bulut, A. Günlü, S. Keles

References

- Aertsen W, Kint V, Van Orshoven J, Özkan K, Muys B, 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol Model* 221(8): 1119-1130. <https://doi.org/10.1016/j.ecolmodel.2010.01.007>
- Aiello S, Kraljevic T, Maj P, 2015. Package 'h2o'. *dim*, 2, 12.
- Asner GP, Scurlock JMA, Hicke J, 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecol Biogeogr* 12(3): 191-205. <https://doi.org/10.1046/j.1466-822X.2003.00026.x>
- Bahrami B, Hildebrandt A, Thober S, Rebmann C, Fischer R, Samaniego L, et al., 2022. Developing a parsimonious canopy model (PCM v1. 0) to predict forest gross primary productivity and leaf area index. *Geoscientific Model Development Discussions*, 1-40. <https://doi.org/10.5194/gmd-2022-87>
- Bonan GB, 1993. Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. *Remote Sens Environ* 43(3): 303-314. [https://doi.org/10.1016/0034-4257\(93\)90072-6](https://doi.org/10.1016/0034-4257(93)90072-6)
- Bramer M, 2002. Using J -pruning to reduce overfitting in classification trees. *Knowledge-Based Systems* 15(5-6): 301-308. [https://doi.org/10.1016/S0950-7051\(01\)00163-0](https://doi.org/10.1016/S0950-7051(01)00163-0)
- Bulut S, Şatır O, Günlü A, 2019. Determining the interactions of black pine net primary productivity and forest stand parameters in northern Turkey. *Appl Ecol Environ Res* 17(2): 4459-4473. https://doi.org/10.15666/aer/1702_44594473
- Candel A, Parmar V, LeDell E, Arora A, 2016. Deep learning with H2O. H2O. ai Inc.
- Cenni E, Bussotti F, Galeotti L, 1998. The decline of a *Pinus nigra* Arn. reforestation stand on a limestone substrate: the role of nutritional factors examined by means of foliar diagnosis. *Ann Sci Forest* 55(5): 567-576. <https://doi.org/10.1051/forest:19980504>
- Chen JM, Rich PM, Gower ST, Norman JM, Plummer S, 1997. Leaf area index of boreal forests: Theory, techniques, and measurements. *J Geophys Res: Atmospheres* 102(D24): 29429-29443. <https://doi.org/10.1029/97JD01107>
- Chen Y, Chen L, Cheng Y, Ju W, Chen HY, Ruan H, 2020. Afforestation promotes the enhancement of forest LAI and NPP in China. *For Ecol Manage* 462: 117990. <https://doi.org/10.1016/j.foreco.2020.117990>
- Chen Y, Jiao S, Cheng Y, Wei H, Sun L, Sun Y, 2022. LAI-NOS: An automatic network observation system for leaf area index based on hemispherical photography. *Agr For Meteorol* 322: 108999. <https://doi.org/10.1016/j.agrformet.2022.108999>
- Clough B, Tan DT, Buu DC, 2000. Canopy leaf area index and litter fall in stands of the mangrove *Rhizophora*

- apiculata* of different age in the Mekong Delta, Vietnam. *Aquat Bot* 66(4): 311-320. [https://doi.org/10.1016/S0304-3770\(99\)00081-9](https://doi.org/10.1016/S0304-3770(99)00081-9)
- Çil B, 2014. Bazı meşcere parametrelerinin farklı uydu görüntüleri yardımıyla tahmin edilmesi: Kelkit ve İğdir Planlama Birimi örneği. Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 72 s., Trabzon.
- Davidson DP, 2002. Sensitivity of ecosystem net primary productivity models to remotely sensed leaf area index in a Montane Forest environment. PhD Thesis, Univ. Lethbridge, Lethbridge, Alberta, Canada.
- Ercanlı İ, Günlü A, Şenyurt M, Keleş S, 2018. Artificial neural network models predicting the leaf area index: a case study in pure even-aged Crimean pine forests from Turkey. *For Ecosyst* 5(1): 29. <https://doi.org/10.1186/s40663-018-0149-8>
- Ercanlı İ, 2020. Ankara Orman Bölge Müdürlüğü Anadolu karaçamı meşcereleri için tek ağaç gövde çapı ve gövde hacminin uyumlu gövde çapı denklemleri ve yapay sinir ağları ile tahmin edilmesi. Tübitak, 119O061 nolu 1002 projesi.
- GDF, 2021. State of Turkey's Forests 2020. Ministry Agr Forest, General Directorate of Forestry, Forest Management and Planning Dept., Ankara, Türkiye.
- Gholz HL, 1982. Environmental limits on aboveground net primary production, leaf area, and biomass in vegetation zones of the Pacific Northwest. *Ecology* 63(2): 469-481. <https://doi.org/10.2307/1938964>
- Gower ST, Kucharik CJ, Norman JM, 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. *Remote Sens Environ* 70(1): 29-51. [https://doi.org/10.1016/S0034-4257\(99\)00056-5](https://doi.org/10.1016/S0034-4257(99)00056-5)
- Gülbeyaz Ö, 2018. Estimating net primary productivity of forest ecosystems over Turkey using remote sensing approach. Degree of Doctor of Philosophy, METU, Natural and Applied Sciences.
- Günlü A, Keleş S, Ercanlı İ, Şenyurt M, 2017. Estimation of leaf area index using WorldView-2 and Aster satellite image: a case study from Turkey. *Environ Monit Assess* 189(11): 538. <https://doi.org/10.1007/s10661-017-6254-2>
- Jagodzinski AM, Kalucka I, 2008. Age-related changes in leaf area index of young Scots pine stands. *Dendrobiology* 59: 57-65.
- Jonckheere I, Fleck S, Nackaerts K, Muys B, Coppin P, Weiss M, Baret F, 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. *Agr For Meteorol* 121(1-2): 19-35. <https://doi.org/10.1016/j.agrformet.2003.08.027>
- Khosravi S, Namiranian M, Ghazanfari H, Shirvani A, 2012. Estimation of leaf area index and assessment of its allometric equations in oak forests: northern Zagros, Iran. *J For Sci* 58(3): 116-122. <https://doi.org/10.17221/18/2011-JFS>
- Kushida K, Isaev AP, Maximov TC, Takao G, Fukuda M, 2007. Remote sensing of upper canopy leaf area index and forest floor vegetation cover as indicators of net primary productivity in a Siberian larch forest. *J Geophys Res: Biogeosci* 112(G2). <https://doi.org/10.1029/2006JG000269>
- Li Z, Zhou T, 2015. Optimization of forest age-dependent light-use efficiency and its implications on climate-vegetation interactions in China. *ISPRS Archiv* 40: 449-454. <https://doi.org/10.5194/isprsarchives-XL-7-W3-449-2015>
- Luo T, Pan Y, Ouyang H, Shi P, Luo J, Yu Z, Lu Q, 2004. Leaf area index and net primary productivity along subtropical to alpine gradients in the Tibetan Plateau. *Global Ecol Biogeogr* 13(4): 345-358. <https://doi.org/10.1111/j.1466-822X.2004.00094.x>
- Maier HR, Dandy GC, 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environ Model Softw* 15(1): 101-124. [https://doi.org/10.1016/S1364-8152\(99\)00007-9](https://doi.org/10.1016/S1364-8152(99)00007-9)
- MathWorks Inc, 2019. Improve shallow neural network generalization and avoid overfitting. Software Manual.
- Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, 2015. Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2015. R package vers 1.6-7.
- Pan S, Tian H, Dangal SR, Ouyang Z, Lu C, Yang J, et al., 2015. Impacts of climate variability and extremes on global net primary production in the first decade of the 21st century. *J Geograph Sci* 25(9): 1027-1044. <https://doi.org/10.1007/s11442-015-1217-4>
- Pan N, Wang S, Wei F, Shen M, Fu B, 2021. Inconsistent changes in NPP and LAI determined from the parabolic LAI versus NPP relationship. *Ecol Ind* 131: 108134. <https://doi.org/10.1016/j.ecolind.2021.108134>
- Peng C, Wen X, 1999. Recent applications of artificial neural networks in forest resource management: an overview. *Transfer* 1(X2): W1. <https://doi.org/10.1109/9780470545355.ch0>
- Potter CS, Randerson JT, Field CB, Matson PA, Vitousek PM, Mooney HA, Klooster SA, 1993. Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochem Cycl* 7: 811-841. <https://doi.org/10.1029/93GB02725>
- Quinto-Mosquera H, Valois-Cuesta H, Moreno-Hurtado F, 2021. Effects of soil fertilization on the allocation of net primary productivity in tropical rainforests of Chocó, Colombia, Preprints, 2021070522. <https://doi.org/10.20944/preprints202107.0522.v1>
- R Development Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
- Raskutti G, Wainwright MJ, Yu B, 2014. Early stopping for non-parametric regression: An optimal data-dependent stopping rule. *J Machine Learn Res* 15: 335-366.
- Scurlock JMO, Asner GP, Gower ST, 2001. Worldwide historical estimates of leaf area index, 1932-2000. ORNL/TM-2001/268, 34. <https://doi.org/10.2172/814100>
- Sprintsin M, Cohen S, Maseyk K, Rotenberg E, Grünzweig J, Karnieli A, et al., 2011. Long term and seasonal courses

- of leaf area index in a semi-arid forest plantation. *Agr For Meteorol* 151(5): 565-574. <https://doi.org/10.1016/j.agrformet.2011.01.001>
- Sterenczak K, Lisanczuk M, Parkitna K, Mitelsztedt K, Mroczek P, Misnicki S, 2018. The influence of number and size of sample plots on modelling growing stock volume based on airborne laser scanning. *Drewno. Prace Naukowe. Doniesienia. Komunikaty*, 61(201).
- Wang Y, Yue T, 2022. Make forests better play their role in mitigating climate change. *Forests* 13(2): 249. <https://doi.org/10.3390/f13020249>
- Warde-Farley D, Goodfellow IJ, Courville A, Bengio Y, 2013. An empirical analysis of dropout in piecewise linear networks. Cornell University, arXiv:1312.6197 [stat. ML].
- Yan H, Wang SQ, Billesbach D, Oechel W, Zhang JH, Meyers T, et al., 2012. Global estimation of evapotranspiration using a leaf area index-based surface energy and water balance model. *Remote Sens Environ* 124: 581-595. <https://doi.org/10.1016/j.rse.2012.06.004>
- Yan W, He Y, Cai Y, Qu X, Cui X, 2021. Relationship between extreme climate indices and spatiotemporal changes of vegetation on Yunnan Plateau from 1982 to 2019. *Global Ecol Conserv* 31: e01813. <https://doi.org/10.1016/j.gecco.2021.e01813>
- Ying X, 2019. An overview of overfitting and its solutions. *Journal of Physics: Conference Series* 1168(2): 022022. <https://doi.org/10.1088/1742-6596/1168/2/022022>
- Zhang M, Lin H, Wang G, Sun H, Cai Y, 2019. Estimation of vegetation productivity using a Landsat 8 time series in a heavily urbanized area, Central China. *Remote Sens* 11(2): 133. <https://doi.org/10.3390/rs11020133>
- Zhao Q, Yu S, Zhao F, Tian L, Zhao Z, 2019. Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments. *For Ecol Manage* 434: 224-234. <https://doi.org/10.1016/j.foreco.2018.12.019>