NICHE BASED DISTRIBUTION MODELLING AND MAPPING OF BRUTIAN PINE IN THE GÖLHISAR DISTRICT

MODELIRANJE DISTRIBUCIJE TEMELJENO NA NIŠAMA I MAPIRANJE BRUCIJSKOG BORA U OKRUGU GÖLHISAR

Özdemir ŞENTÜRK[1]*

SUMMARY

The purpose of this work was to elucidate the fundamental characteristics and application areas of various modelling techniques that are widely employed in current ecological modelling research. Using five distinct distribution modelling techniques, possible species distribution modelling and mapping of the Brutian pine species in the Gölhisar district were conducted. The data was collected from Brutian pine species in 400 sampling plots in the area. The variables used in the models were elevation, slope, aspect, radiation index, heat index, topographic position index and bedrock types. Logistic regression, classification tree, random forest, generalized additive model and maximum entropy were used as the species distribution modelling methods. Receiver Operating Characteristics (ROC) curves were created and the performance of the species distribution models was evaluated with the Area Under the ROC curve (AUC). The statistical analyses revealed that the best models were generalized additive model, random forest, classification tree, maximum entropy and logistic regression, respectively. Elevation and bedrock types had the highest contribution to the Brutian pine distribution models. The outputs of the generalized additive model technique that had the highest AUC value were mapped. Some ecological and statistical differences were found between the models and their reasons were presented. Compared to the methods commonly used in species distribution modelling studies, generalized additive model technique has a specific smoothing function which ensures both fittings between the environmental changes and explanatory curves and more accurate ecological interpretation of the models obtained.

KEY WORDS: classification tree, ecological modelling, generalized additive model, logistic regression, MaxEnt, random forest

INTRODUCTION
UVOD

Human-related factors such as rapid population growth from past to present, industrialization, agricultural activities, mining, utilization of forests and forest products, have led to an increased pressure on ecosystems. This has resulted in changes, destruction, and degradation of forest ecosystems. Particularly, it is known that deforestation is at a very severe scale even in important forest ecosystems such as the Amazon forests in Brazil (Malhi et al., 2008), Colombia (Armenteras et al., 2006), and mangroves (Blanco et al., 2012; Jones et al., 2014) due to agricultural activities, smuggling, timber production etc. Similarly, the fact that rural people in Türkiye usually live in forests or areas adjacent to forests increases the pressure on forest ecosystems. This also applies to the ecosystems in the Mediterranean region, which have been exposed to severe degradation and damage for centuries. As a result of this, most forest ecosystems in

[1] Assoc. Prof. Dr. Özdemir Şentürk, Burdur Mehmet Akif Ersoy University, Department of Forestry, Gölhisar Vocational School, Gölhisar, Burdur, 15400, Türkiye
* Corresponding address: osenturk@mehmetakif.edu.tr

the Mediterranean region have become semi-natural ecosystems. Combined with climate change, these ongoing human-related degradations have resulted in decreased biological diversity, reduction in the carbon storage capability of ecosystems, losses in food cycle, erosion, soil and air pollution, habitat losses and fragmentation, reduction in the number of individuals or dislocation of flora and fauna (Hobbs, 1994; Matthews et al., 2000).

In conclusion, it is imperative to develop ecosystem-based management plans aimed at protecting and improving ecosystems to ensure their sustainability. Within this scope, the concept of ecological restoration in forest ecosystems comes to the fore. Jackson et al. (1995) defined ecological restoration as the repair of the human-related damage on the ecosystem diversity and dynamics. Information about the relationships of living organisms with their environment is of great importance to repair the damage, destruction, and degradation of an ecosystem. Therefore, one of the most important steps of ecological restoration is to identify the environments that species can adapt to. For that reason, the ecological features of the species that compose an ecosystem should be identified and mapped. To this end, geographical information systems and modelling methods are the most important tools used for obtaining this ecological information.

Species distribution models are divided into two groups: mechanistic methods and correlative approaches. Mechanistic methods are based on the ecophysiological features of the species to determine their geographical distribution, while correlative approaches model the binary data of the dependent variable or presence-only data according to the environmental factors. If the dependent variable consists of binary data, group discrimination techniques (logistic regression, discriminant analysis, classification tree, generalized additive model, generalized linear model, random forest, artificial neural network, boosted regression trees, multiple linear regression etc.) are used, but if it consists of presence-only data, profile techniques (maximum entropy, genetic algorithm for rule-set prediction, DOMAIN, ENFA etc.) are used (Beerling et al., 1995; Caithness, 1995; Robertson et al., 2003; Brotons et al., 2004; Mateo et al., 2010; Özkan et al., 2015; Mert et al., 2016; Shabani et al., 2016). Correlative methods are preferred in the species distribution models, and target species may include tree or plant species, wild animals, insects, fungi, birds, or reptiles (Ortega-Huerta and Peterson, 2008; Kramer-Schadt et al., 2013; Sarmiento-Ramírez et al., 2014; Filz and Schmitt, 2015; Rameshprabu and Swamy, 2015; Mert and Kıraç, 2019). In recent years, species distribution models have been used for various purposes such as restoration, planning, protection and sustainability of ecosystems, monitoring, wildlife management, identification of different habitats, planning of endemic or rare species, climate change crisis, assessment

of biological diversity and similar (Franklin et al., 2013; Hamann and Aitken, 2013; Fajardo et al., 2014; Ramirez-Villegas et al., 2014; Mert et al., 2016).

There are several methods and approaches regarding species distribution models (Ferrier et al., 2002; Phillips et al., 2006; Austin, 2007). Methods that have different predictions in ecosystem studies especially play an important role for the explanation or interpretation of relations. As a matter of fact, the relationships in natural ecosystems may not always be linear. They are usually imbalanced, incomplete and complicated; therefore, non-parametric methods should be used to determine such relationships (De'ath and Fabricius, 2000; Cutler et al., 2007; Özkan, 2012a).

Due to different topographical features and local climate types even at short distances, forest ecosystems in the Mediterranean region of Türkiye are complicated. This is why this region is rich in plant diversity and contains many main forests tree species. Gölhisar district is in the mountainous and karstic area of the Mediterranean region. Thus, it has a high species diversity. Gölhisar is an old settlement area that was established around 200 BC. People in this locality were engaged in forging, agriculture, horse and animal raising, wild animal hunting and timber production under the Roman rule (Özüdoğru, 2018). The forest ecosystems have been severely damaged due to the utilization of forests in this area for centuries. Main forest tree species play a crucial role for the continuity of utilization in forest ecosystems. In this context, Brutian pine (*Pinus brutia* Ten., also known as Turkish red pine) is one of the species with the widest distribution in Türkiye and is among the species that are preferred from a functional perspective. This s because Brutian pine in Türkiye plays an important role in ecosystem-based functional planning for the reduction of effects of climate change, protection of biological diversity, afforestation, ensuring the sustainability, etc. Due to its significance as a key forest tree in Türkiye, this species was utilized more extensively. Although Brutian pine is distributed in a relatively small area in the Gölhisar district, it has an important place in utilization. Brutian pine stands that have a production function in the area are not only utilized for economic purposes but are also used for non-wood products such as resin, root wood, and firewood. On the other hand, there is a Brutian pine gene conservation forest around Gölhisar and Gölova. Gene conservation forests, seed stands, and seed orchards are needed for successful afforestation. In this regard, Brutian pine trees that are distributed up to an elevation of around 1600 m in the locality have an important role in genetic improvement studies (Carus and Çatal, 2012). The utilization of Brutian pine stands for ecological and economic purposes in the district necessitates the protection and ensures the sustainability of Brutian pine ecosystems. Therefore, Brutian pine is one of the crucial species for the ecology-based functional

planning of the locality, while one of the priorities should be to identify the locations appropriate for its distribution for effective protection and sustainability.

In the study conducted in the Gölhisar district to achieve the abovementioned objective, logistic regression, classification tree technique, random forest, generalized additive model and maximum entropy were used for modelling. The results were obtained from each model, the differences between the models were compared and discussed statistically and ecologically, and then the most appropriate distribution model was chosen in order to develop the potential distribution map of the species.

MATERIALS AND METHODS
MATERIJALI I METODE

Site description – *Opis područja*

The study site covers Gölhisar, Çavdır, Tefenni and Altınyayla districts in Burdur province, Çameli district in Denizli province, and some parts of Fethiye district in Muğla province. It is located between the east longitudes 35° 44' 54" – 35° 19' 27" and north latitudes 36° 44' 03" – 37° 21' 33". The elevation of the site ranges from 540 m to 2300 m. The study site is surrounded by Boncuk mountains in the southwest, Akdağ in the west and Eren Mountains in the south. The site has an area of 252680 hectares with 136743 hectares of forest and 115937 hectares of nonforest area. Crimean juniper, black pine, oak, Brutian pine and Taurus cedar are tree species with the widest distribution. Mediterranean climate is dominant within and around the study site which is located within the boundaries of the Lakes Region. However, since the site is located in the transition zone between the Mediterranean climate and continental climate, features of the transition climate have also been observed. According to the Köppen climate classification, Mediterranean climate (Csa) with mild winters and very hot and dry summers is dominant in the study site (Rubel et al., 2017). According to the Thornthwaite precipitation effectiveness index and climate classification, semiarid less humid (C1) climate is observed in the site (Yılmaz and Çiçek, 2016). The geological structure of the Gölhisar locality consists of three layers of Lycian, Middle-Upper Miocene and Gölhisar formations in the age range from Mesozoic to the present. These formations contain limestone, sandstone, coarse-grained conglomerate, marl and ophiolitic bedrock types (Elitez et al., 2018; Varol, 2011). Considering the site where the study was conducted, it has poor soil structure since most of the area is located on a sloped and rugged terrain.

Environmental variables –*Varijable okoliša*

In order to create models and maps through statistical analyses, environmental variables should be created. As-pect, slope, and elevation were created in the raster format using the digital elevation model (DEM) of the Gölhisar district. To determine the landform of the study area, "Topographic Tools" extension was used, and topographic position index map was created (Jennes, 2006). Each bedrock type in the parent materials obtained from the DG of Mineral Research and Exploration was drawn as a polygon considering the boundaries of the study area, and bedrock types were recorded in the attribute table. At the last stage, those bedrock types (limestone, sandstone, chert, serpentine, alluvium, quartzite, basalt) in the form of vectors were exported in raster format using the conversion options. The cellular values of slope and aspect variables created for the study area were used to calculate heat index and radiation index (Equation 1 and 2).

$$HI = \cos(radyan((\theta) - \theta\ maks)) \times (\tan(radyan(\Delta)))) \quad (1)$$
$$RI = (1 - \cos((\pi / 180) \times (\theta - 30))) / 2 \quad (2)$$

HI = Heat Index

RI = Radiation Index

In these equations, θ represents aspect and Δ represents slope, while θ maks represents 202.5°. The results of equation 1 ranged from -1 to 1 (Parker, 1988; Olsson et al., 2009) and of equation 2 they ranged from 0 to 1 (Roberts and Cooper 1989).

At the final stage, heat index and radiation index variables were created in raster format with "Raster Calculator" tool in ArcMap 10.2 software based on the cellular values of the indices calculated for the entire study area.

Statistical evaluation

At this stage, binary data of Brutian pine were collected from 400 sampling plots each with a size of 20×20 m. Presence-absence data were used for the group discrimination techniques among the correlative approaches, while presence-only data were used for the profile technique. Receiver Operating Characteristic (ROC) curves were created and the accuracy of all models obtained was checked with the Area Under the ROC curve (AUC).

Pearson's chi square (for categorical variables: limestone, sandstone, chert, serpentine, alluvium, quartzite, basalt) and Mann-Whitney U-tests (for continuous variables: elevation, aspect, slope, topographic position index, heat index, radiation index) were performed to detect whether there was a relationship between Brutian pine and the independent variables (Pearson, 1900; Mann and Whitney, 1947).

Logistic Regression (LR) analysis is a parametric method that predicts the relation between the dependent variable and the independent variable according to the possibility rules. Furthermore, this method is referred to as the logarithmical conversion of the linear regression analysis. Here

the dependent variable is categorical and it is applied in three different ways (binary logistic, nominal logistic, ordinal logistic) (Hosmer and Lemeshow, 2000; Şenel and Alatlı, 2014). During the analysis, the dependent variable may be continuous or categorical (Peng et al., 2002; Aksaraylı and Saygın, 2011). The prediction values of the model range from 0 to 1.

Classification and Regression Tree Technique (CART) is called classification tree technique (CTT) if the dependent variable contains binary data, and regression tree technique (RTT) if it contains continuous data. As the presence-absence data of Brutian pine was used in this study, CTT was used. Classification tree is a non-parametric rule-based method and it divides environmental variables into homogenous sub-groups as the dependent variable. These sub-groups form a hierarchical tree model and are connected to one another with leaves and nodes (Breiman et al., 1984; De'ath and Fabricius 2000; Chu et al., 2009; Özkan, 2012b). The independent variable values at the last nodes of the resulting tree model are written with rules, and in this way, the estimated values of the model are obtained (Özkan, 2012b; Özkan, 2013; Mert et al., 2016).

Random Forest (RF), which is a more sophisticated version of CTT, is a non-parametric algorithm-based method. RF produces several classification trees to get the most accurate result and makes assessment by combining the trees it produces. After creating several classification trees, RF combines the predictions from all trees to finalize the process (Cutler et al., 2007; Evans and Cushman, 2009; Grossmann et al., 2010). RF can use categorical, sequenced, and continuous data in the same analysis, and it does not need data conversion (Evans et al., 2011). To create the prediction model, it needs the number of the classification trees and that of the prediction variables on each node. However, the number of the classification trees and the prediction values on each node should be optimized to minimize the generalization error during the analysis (Rodriguez-Galiano et al., 2012). As RF eliminates the overfitting that occurs in CTT method and applies cross validation during the analysis, it is used in many disciplines and has become more popular (Breiman, 2001; Cutler et al., 2007; Grossmann et al., 2010; Evans et al., 2011). In this study, "randomForest"

package in R studio software was used for distribution modelling of Brutian pine (Breiman, 2001).

Generalized Additive Model (GAM) is the non-parametric version of the generalized linear modelling technique, and it allows the determination of non-linear distributions with its smoothing function. GAM has a prediction capability in determining the complicated relations between dependent variables and environmental variables. Different assessment functions are used depending on the characteristics of the dependent variable (Guisan et al., 2002; Moisen et al., 2006). In this study, Generalized Regression Analysis and Spatial Prediction (GRASP) in S – Plus 6.1 package software was used to create the species distribution model and map using the GAM technique (Lehmann et al., 2002; Özkan, 2012a; Gülsoy and Özkan, 2013). ANOVA (F test) stepwise model in GRASP was used during the species distribution modelling and mapping process.

Maximum Entropy (MaxEnt) that is one of the methods which uses only presence data in species distribution modelling to create model(s) for the relationships between the dependent variable and environmental variables based on the possibility data in areas with similar features (Phillips et al., 2006; Baldwin, 2009). For the prediction of the model(s) obtained, maximum entropy theory is used. This method, which makes predictions based on incomplete information and determines the possibility distribution, provides the opportunity to assess the categorical and continuous data at the same time (Phillips et al., 2006). On the other hand, MaxEnt can create models using any of the logistic, raw, and logarithmical options. In this study, logistic option was chosen because its possibility values ranged from 0 to 1 and it provided near-real prediction values (Phillips and Dudik, 2008; Baldwin, 2009). While predicting the presence possibility, 0 gets the lowest prediction value while 1 gets the highest value (Phillips et al., 2004; Kumar and Stohlgren, 2009).

RESULTS
REZULTATI

Mann-Whitney U-test results show that elevation, slope, topographic position index, and heat index are significant

Table 1. Mann-Whitney U-test results for continuous variables
Tablica 1. Rezultati Mann-Whitney U-testa kontinuiranih varijabli

	U	W	Z	p
Elevation / *Nadmorska visina*	1628	13104	−15.319	**0.000**
Aspect / *Orijentacija*	18616	49741	−0.164	0.870
Slope / *Nagib*	15640	27116	−2.819	**0.005**
Radiation index / *Indeks zračenja*	17107	48232.5	−1.509	0.131
Heat index / *Toplinski indeks*	15830	46955.5	−2.649	**0.008**
Topographic position index / *Indeks topografskog položaja*	16160	27636.5	−2.354	**0.019**

Table 2. Pearson's chi-square (χ^2) test results for continuous variables
Tablica 2. Rezulati Pearsonova hi-kvadrat (χ^2) testa kontinuiranih varijabli

	a	b	c	d	χ^2	p
Limestone / *Vapnenac*	161	88	127	24	17.634	0.000
Sandstone / *Pješčenjak*	175	74	40	111	72.510	0.000
Serpentine / *Serpentinit*	213	36	140	11	4.664	0.031
Alluvium / *Aluvij*	241	8	150	1	2.780	0.095
Chert / *Rožnjak*	210	39	149	2	21.006	0.000
Basalt / *Bazalt*	249	0	149	2	3.315	0.069
Quartzite / *Kvarcit*	245	4	151	0	2.450	0.118

for the occurrences of Brutian pine (Table 1). According to the result of Pearson's chi-squared test, limestone, sandstone, chert, and serpentine were detected playing an important role in the site properties factors of Brutian pine (Table 2). Environmental variables which are considerably related to the independent variables according to the results of Mann-Whitney U-test and Pearson's chi-squared tests were involved in the modelling processes.

In the model obtained from LR analysis, sandstone, which was a bedrock type, and elevation were found to be the important variables. In the analysis, backward LR method was used and the adequacy of the fit between the variables that composed the models was checked with Hosmer-Lemeshow test. The significance of the model according to Hosmer-Lemeshow test was found to be 0.146 (p>0.05). This value shows that the model is valid. Elevation provided the highest contribution to the model. Both variables in the model had a negative correlation. In other words, as elevation increased, the species distribution decreased and preferred the areas without sandstone. The validity of the model was checked with the AUC, while the values were found to be 0.882 for the training dataset and 0.881 for the test dataset. The probability value of the model obtained from the LR analysis was calculated using the Equation 3 below.

$$p = \frac{exp^{20.182+\left(Elevation*(-0.015)\right)-\left(Sandstone*(-0.880)\right)}}{1+exp^{20.182+\left(Elevation*(-0.015)\right)-\left(Sandstone*(-0.880)\right)}} \quad (3)$$

Sandstone as one of the bedrock types and elevation were the variables that constructed the tree model obtained from CTT. The tree model comprised 6 groups and 4 terminal nodes. Elevation provided the highest contribution to the model. According to the tree model, areas below 1345 m and areas with sandstone between 1345–1400 m represented the potential distribution areas of the Brutian pine (Figure 1). The AUC values of the training dataset and the test dataset of the model were found to be 0.927 and 0.913, respectively.
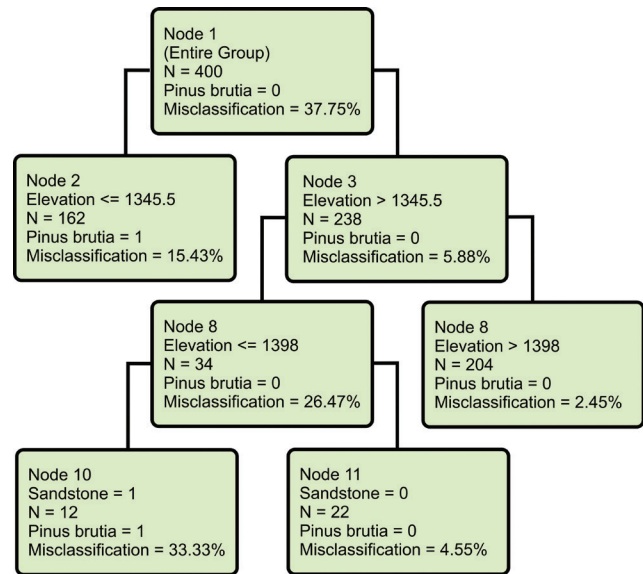

Figure 1. Tree model of the Brutian pine
Slika 1. Model stabla brucijskog bora

The number of the most ideal decision trees in the model created with the RF method was selected as 500 as reported by Breiman (2001). The variables constructing the model included elevation, heat index, topographic position index and bedrock types. Regarding the contribution of the variables in accordance with the Gini value, elevation (109.17), heat index (17.26), topographic position index (15.39) and bedrock types (sandstone (12.43), chert (2.12), limestone (1.93), serpentine (1.23)) had the highest contribution (Figure 2). Because a low Gini value was obtained from the variable of the slope, it wasn't used in the RF modelling process. According to the constructed random forest model, Brutian pine showed potential distribution up to 1600 m altitude, hot or hotter aspects, plains, open slopes, and convex lands and sandstone areas. On the other hand, the AUC value of the model was found to be 0.938.

Elevation, bedrock types and heat index were the variables used in the GAM, and among them elevation had the highest contribution. According to the model, elevations of 950–1500 m were the most suitable potential distribu-
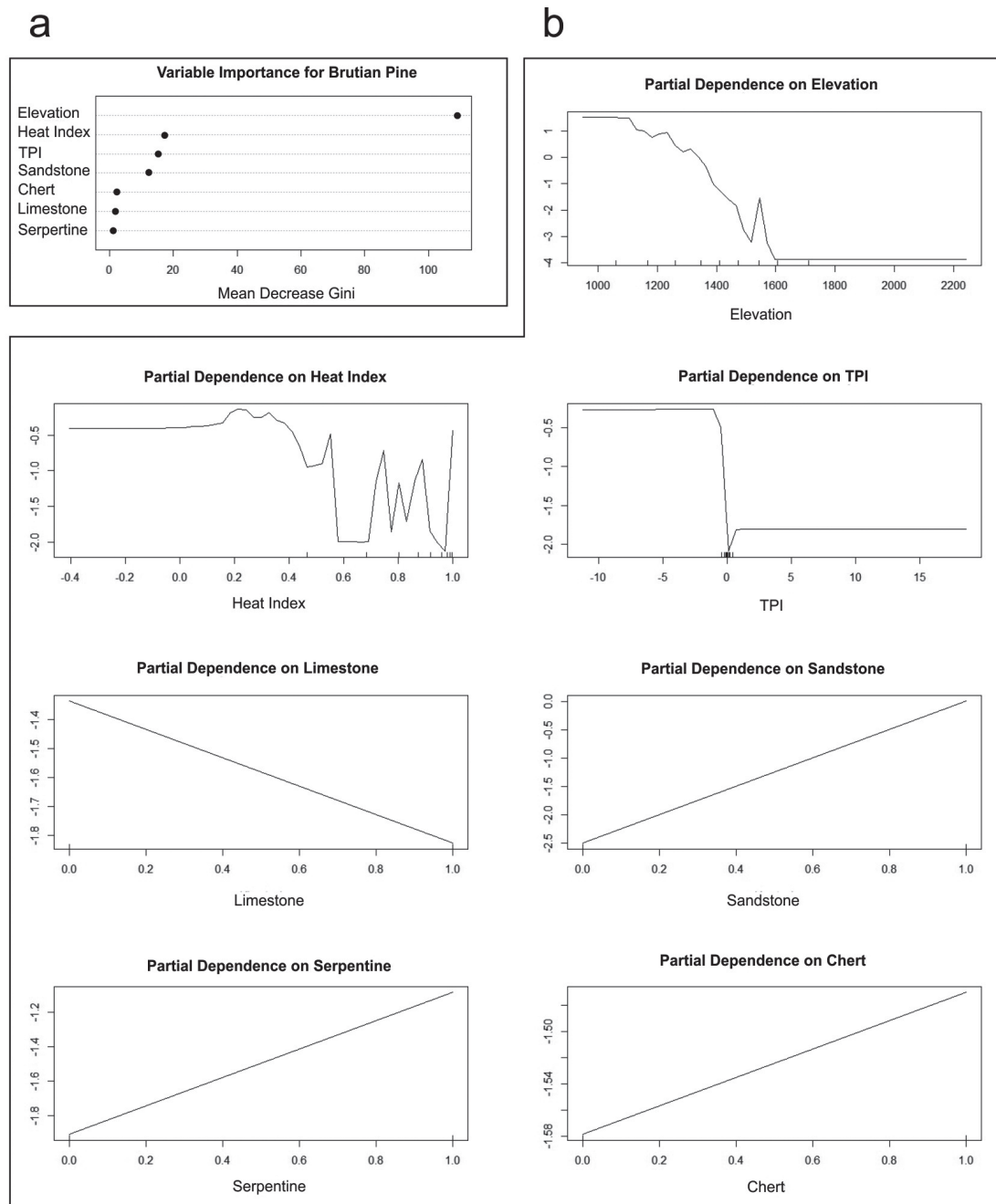
Figure 2. Random forest variable importance (a) and model results of response curves (b) of environmental variables for Brutian pine
Slika 2. Značajnost varijable prema modelu slučajne šume (a) i rezultati modela krivulja odgovora (b) varijabli okoliša za brucijski bor
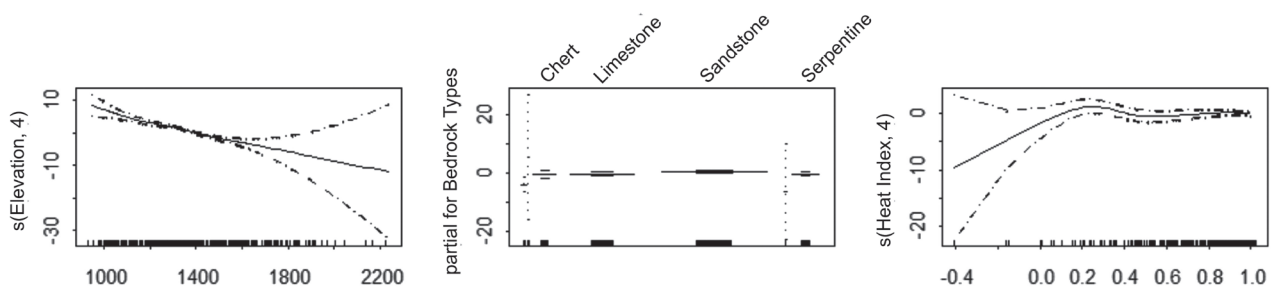

Figure 3. Response curves of environmental variables in the GAM analysis
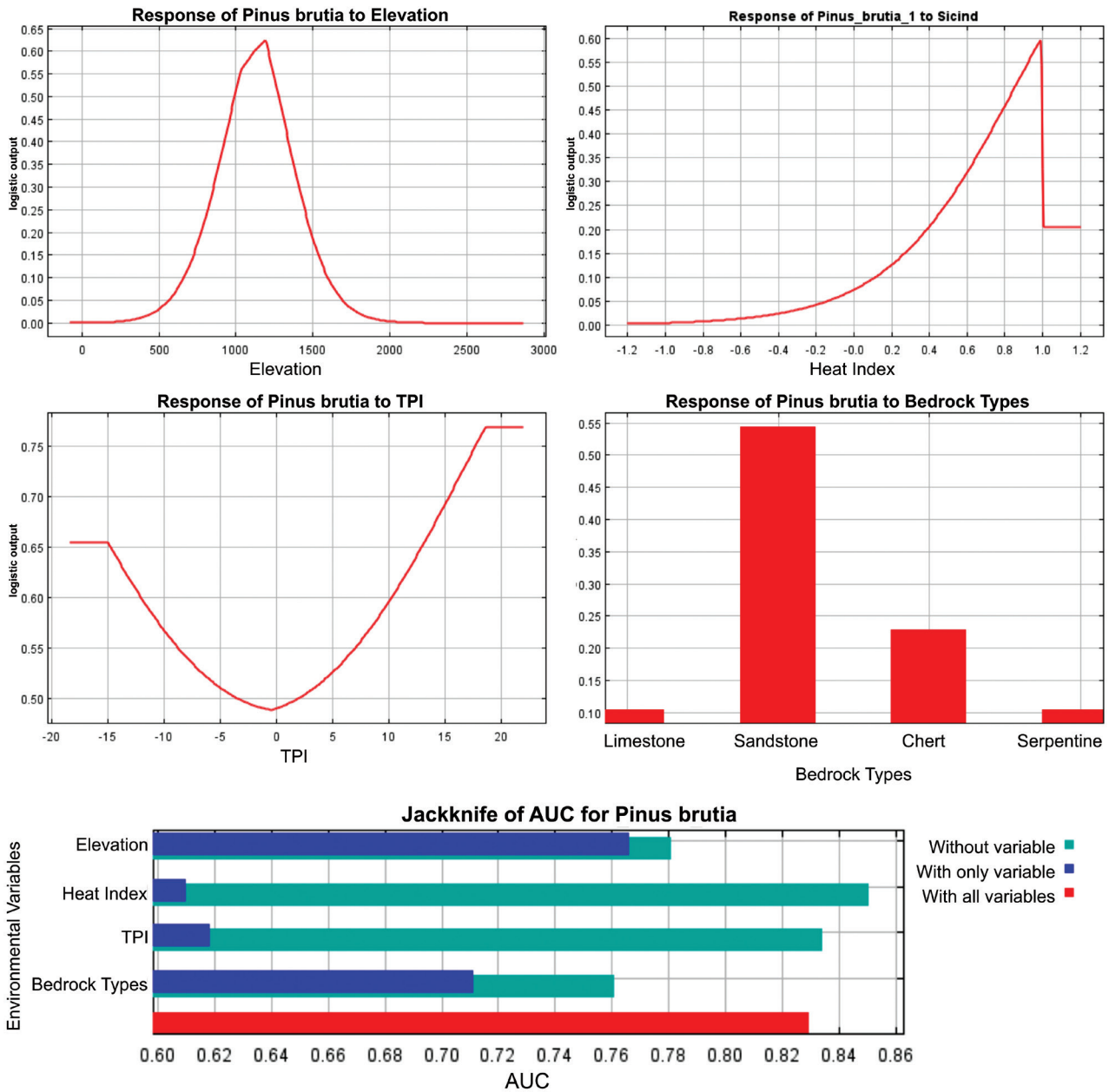Slika 3. Krivulje odgovora varijabli okoliša u GAM analizi

Figure 4. Response curves and results of jackknife estimation of relative importance of environmental variables for Brutian pine
Slika 4. Krivulje odgovora i rezultati *jackknife* procjene relativnog značaja varijabli okoliša za brucijski bor

tion areas for Brutian pine, while the distribution of the species decreased gradually starting from the elevation of 1600 m. Heat index, which was another variable in the model, showed that hot and hotter aspects created appropriate areas for the species, while as the index value got closer to 1, its distribution gradually increased. Finally, areas with sandstone constituted the most suitable potential distribution areas for Brutian pine (Figure 3). The formula of the model was = s (Elevation, 4) + Bedrock Types + s (Heat Index, 4). The validation value of the model was 0.964, while the cross validation result was 0.937.

According to the MaxEnt results related to *P. brutia*, elevation, heat index, topographic position index, and bedrock types were the variables used for constructing the model.

The variable of the slope wasn't important contribution to model. So it wasn't used to consider. Elevation and bedrock types had the highest impact in the model. The distribution of Brutian pine increased up to the elevation of around 1100 m, while its distribution decreased as the elevation increased. The distribution of Brutian pine in the area was found to be related to sandstone, serpentine, limestone, and chert among bedrock types. Hot and hotter aspects were found to be the most suitable potential distribution areas. Finally, the species was found to have a potential distribution area on plains, open slopes, and convex lands such as hills and summits according to the topographic position index (Figure 4). The AUC values of the model were 0.854 for the training dataset and 0.803 for the test dataset.
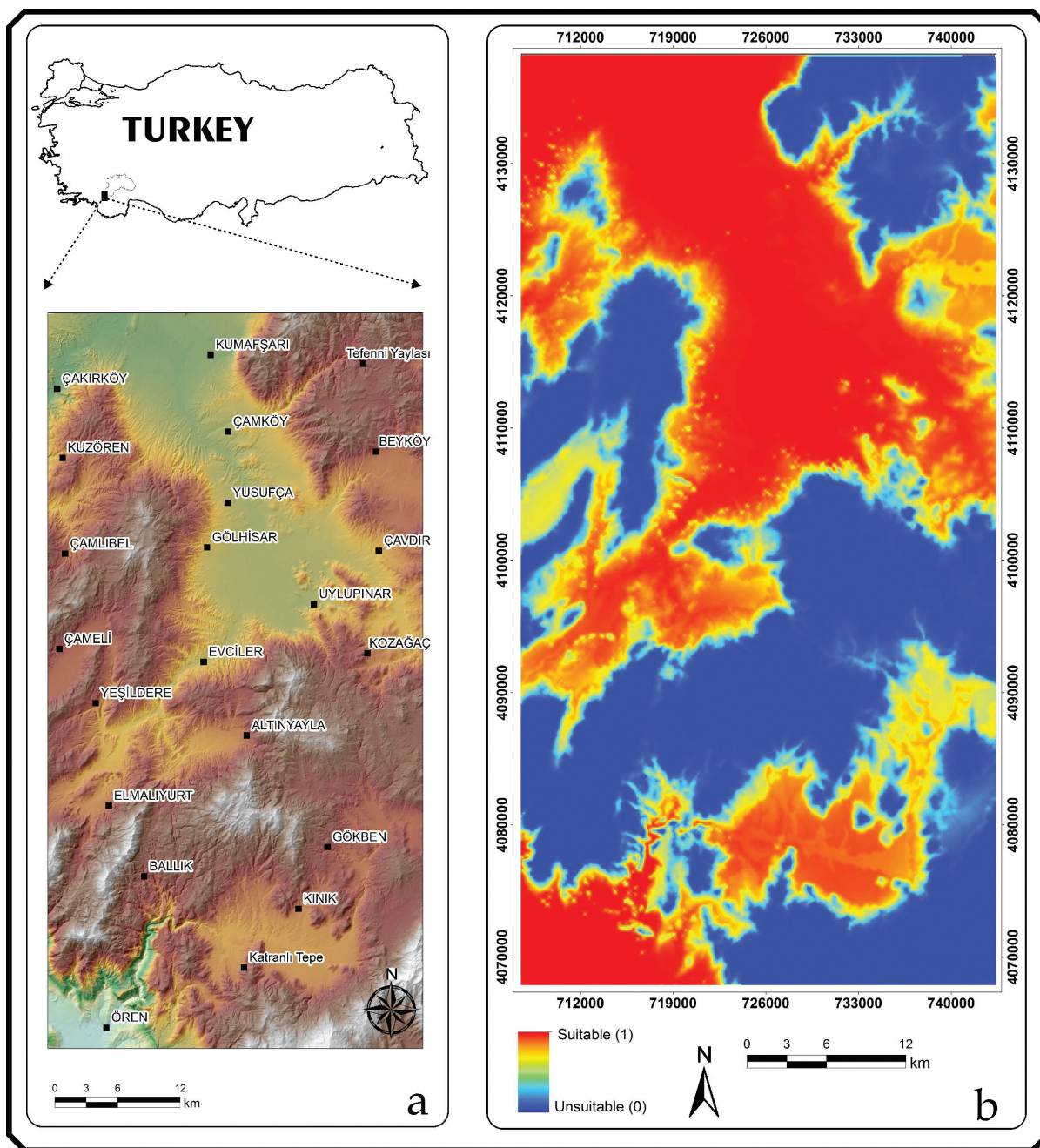
Figure 5. a) Location map of the study area; **b)** Species distribution model maps of GAM
Slika 5. a) Karta lokacije područja istraživanja; b) Karte distribucije vrste prema modelu GAM

The model with the highest AUC value according to the correlative approaches was obtained from the GAM technique. Potential distribution maps were created based on the distribution model obtained from the GAM technique and visualized using ArcMap 10.2 software (Figure 5).

DISCUSSION
RASPRAVA

In this study, which was conducted using five different modelling techniques, potential distribution models and maps were created for Brutian pine distribution in the Gölhisar

district. The differences between the models were compared and the most appropriate distribution model's map was visualized. The models obtained with each method were constructed by elevation, bedrock types, heat index, topographic position index, and slope. Furthermore, elevation and bedrock types had the highest contribution to the models. The accuracy of the models was checked with AUC. The highest performance was found in GAM, RF, CTT, MaxEnt and LR methods, respectively.

Brutian pine, which is mainly distributed in the Mediterranean and Aegean regions in Türkiye, is locally distributed

in the inlands of the Central and Western Black Sea region (Arbez, 1974; Kandemir et al., 2010; Şentürk et al., 2019). Brutian pine establishes pure stands up to the elevation of 800 m in the Mediterranean and Aegean regions where it has its primary distribution, whereas it also establishes mixed stands with other main forest tree species up to the elevation of 1200 m. It is distributed up to 1600 m in the Gölhisar district (Kılıç and Güner, 2000), while the models created revealed that the most suitable distribution areas were located up to the elevation of 1400–1500 m. In most parts of the Taurus Mountains which are karstic, Brutian pine develops well. Brutian pine grows on such bedrock types as primarily limestone and marl, conglomerate, serpentine, basalt sandstone and schist (Atalay et al., 2008; Atalay and Efe, 2015). It mainly prefers areas with sandstone and limestone in that locality, while it avoids areas with alluvium. Elevation and bedrock types constructed all the models in the study. In the study, Brutian pine was found to prefer sunny aspects, which is because it is a light-demanding tree species. At sufficient light concentrations, it develops better. On the other hand, open slopes and midslope lands, and convex lands such as hills and summits were the appropriate distribution areas for Brutian pine by topographic position index. Slope areas plays an important role in the growth of Brutian pine stands (Atalay et al., 2008). Some studies found that the productivity of Brutian pine stands increased as the slope increased depending on elevation (Özkan and Kuzugüdenli, 2010). In a study conducted in the same district, midslope was found to be one of the important factors for the distribution of Brutian pine (Şentürk, 2012).

The model from the LR analysis was constructed by elevation and sandstone and AUC values of these variables were found to be 0.882 and 0.881. LR had the lowest AUC value compared to the other methods used in the study. The reason is that LR analysis is a non-linear method and thus it can hardly classify the variables (Holden et al., 2011). As known, relationships in forest ecosystems are usually complex; therefore, non-parametric methods provide more accurate results to determine these relationships (Özkan, 2012b; Özkan, 2013).

CTT and RF are rule-based non-parametric methods and construct the models by assessing the best options in modelling (Holden et al., 2011). There are differences between CTT and RF. RF creates decision trees without a need for pruning, and validates the model's performance. While creating each decision tree without requiring the training and test sets, it can determine the relationships and distance between the variables composing the model and eliminate the overfitting; which makes it superior to the classification tree technique (Breiman, 2001; Grossman et al., 2010; Özdemir, 2018). However, the tree model and the process steps are not visible entirely (black box), which is a disad-

vantage of the RF technique (Breiman, 2001; Cutler et al., 2007; Evans et al., 2011).

In this study, MaxEnt technique had a lower performance compared to the other non-parametric methods. Brutian pine, which was selected as the target species, is distributed widely in the Gölhisar locality due to its high ecological tolerance range. The most important difference of MaxEnt from other techniques is that it creates species distribution models using presence-only data. This method produces absence data artificially using the cell values in the background and creates the distribution model by calculating the possibilities based on these cells (Phillips et al., 2006). In particular, it provides better results when creating habitat suitability models in wildlife studies, when determining the distribution areas of endemic or rare, endangered species, and identifying the present and future impacts of climate change on the distribution of species (Ray et al., 2011; Babar et al., 2012; Clements et al., 2012; Kumar, 2012; Marcer et al., 2013; Garcia et al., 2013; Ma et al., 2014; Kıraç and Mert, 2019; Süel, 2019; Acarer, 2024). Since it is difficult to validate the absence data of the target species in habitat suitability modelling studies (Mateo et al., 2010), presence-only data is suitable for such studies. On the other hand, the use of presence-absence data is more suitable in developing the distribution models of plant species. However, if the plant species is distributed in a narrow area, rare or endemic, MaxEnt can be effective. Moreover, MaxEnt provides more effective results in case the number of sampling plots is low, which is why it is preferred in some studies (Phillips et al. 2004; Elith et al., 2006; Hernandez et al., 2006; Pearson et al., 2007; Wisz et al., 2008; Kumar and Stohlgren, 2009).

In this study, the most successful model was obtained by the GAM technique. Compared to the methods commonly used in species distribution modelling studies, GAM technique has a specific smoothing function which ensures both fitting between the environmental changes and explanatory curves (Lehmann et al., 2002) and more accurate ecological interpretation of the models (Austin, 1999; 2002). Since the relationships in ecosystems are complicated, this function of the GAM technique to explain the curvilinear relationships ensures fitting between the dependent variable and environmental variables and thus improves the prediction capability of the model (Yee and Mitchell 1991).

CONCLUSIONS
ZAKLJUČCI

Increasing the number of variables in the modelling techniques is as important as the use of different modelling techniques to improve the prediction capability of the target species (Moisen and Frescino, 2002). On the other hand, the adequacy of the sampling conducted for the tar-

get species (number of sample areas, type of data, ability to represent the species in the district, etc.) is also important in increasing the success of modelling techniques. In this study, the results of different models were significantly similar. However, the results of the correlative approaches were better than those of the MaxEnt method which is a profile technique. This is due to the reliability of the inventory data of the species (presence-absence data) (Elith et al., 2006; Phillips and Dudik, 2008). According to the GAM technique that had the best result; elevations from 950 m to 1500 m, hot and hotter aspects and areas with sandstone represented the most suitable potential distribution areas. This distribution map provides basic information that plays a crucial role in the decision-making process for the planning of Brutian pine species. It is also important to create the potential distribution maps of the other target species in the region using similar techniques for a holistic ecosystem planning. It should be noted that each technique has its unique advantages. It is important to consider various factors, such as data types of the target species, the number of sample areas, and explanatory variable selection, when selecting the most suitable modelling technique.

REFERENCES
LITERATURA

- Acarer, A. 2024: Brown bear (*Ursus arctos* L.) distribution model in Europe, *Sumarski List*, 148(5–6), 261-272.

- Aksaraylı, M., Ö. Saygin, 2011: Perceived service quality and determination of the effect on service preference with logistic regression analysis, *Dokuz Eylul University the Journal of Graduate School of Social Sciences,* 13(1), 21-37.

- Arbez, M. 1974: Distribution, ecology and variation of Pinus brutia in Turkey, *Forest Genetic Resources,* 1, 21-23.

- Armenteras, D., G. Rudas, N. Rodriguez, S. Sua, M. Romero, 2006: Patterns and causes of deforestation in the Colombian Amazon, *Ecological indicators*, 6(2), 353-368.

- Atalay, İ.; Efe, R.; Soykan, A. 2008: Mediterranean ecosystems of Turkey, ecology of the Taurus mountains. environment and culture in the Mediterranean Region Part I, Cambridge Scholars Publishing, pp, 337.

- Atalay, I., Efe, R., & Soykan, A. (2008). Mediterranean ecosystems of Turkey: ecology of the Taurus Mountains. In *Natural Environment and Culture in the Mediterranean Region* (Vol. 3,

No. 38, pp. 3-38). Cambridge Scholars Publishing in association with GSE Research.

- Atalay, İ., R. Efe, 2015: Türkiye Biocoğrafyası [In Turkish]. *Meta Basım Matbaacılık Hizmetleri İzmir*, pp, 536.

- Austin, M. 1999: A silent clash of paradigms, some inconsistencies in community ecology, *Oikos,* 86(1), 170-178.

- Austin, M. P. 2002: Spatial prediction of species distribution: an interface between ecological theory and statistical modelling, *Ecological modelling*, 157(2-3), 101-118.

- Austin, M. 2007: Species distribution models and ecological theory: a critical assessment and some possible new approaches, *Ecological modelling*, 200(1-2), 1-19.

- Babar, S., G. Amarnath, C. S. Reddy, A. Jentsch, S. Sudhakar, 2012: Species distribution models: ecological explanation and prediction of an endemic and endangered plant species (Pterocarpus santalinus Lf), *Current Science*, 102(8), 1157-1165.

- Baldwin, R. A. 2009: Use of maximum entropy modelling in wildlife research, *Entropy*, 11(4), 854-866.

- Beerling, D. J., B. Huntley, J. P. Bailey, 1995: Climate and the distribution of Fallopia japonica: use of an introduced species to test the predictive capacity of response surfaces, *Journal of Vegetation Science*, 6(2), 269-282.

- Blanco, J. F., E. A. Estrada, L. F. Ortiz, L. E. Urrego, 2012: Ecosystem-wide impacts of deforestation in mangroves: the Urabá Gulf (Colombian Caribbean) case study, *International Scholarly Research Notices (ISRN Ecology),* 1-4.

- Breiman, L. 2001: Random forests, *Machine Learning,* 45(1), 5-32.

- Breiman, L., J. H. Friedman, R. A. Olshen, C. J. Stone, 1984: Classification and regression trees, *Belmont,* CA, Wadsworth, International Group, 432, 151-166.

- Brotons, L., W. Thuiller, M. B. Araújo, A. H. Hirzel, 2004: Presence-absence versus presence-only modelling methods for predicting bird habitat suitability, *Ecography*, 27(4), 437-448.

- Caithness, N., 1995: Pattern, process and the evolution of the African antelope (Mammalia, Bovidae). In PhD Thesis.

- Carus, S., Y. Çatal, 2012: Basic Characteristics of Forest Enterprise in Burdur Region, Problems and Solutions, *The Journal of Graduate School of Natural and Applied Sciences of Mehmet Akif Ersoy University*, 3(1), 53-59.

- Chu, C. M., B. W. Tsai, K. T. Chang, 2009: Integrating decision tree and spatial cluster analysis for landslide susceptibility zonation, *International Journal of Geological and Environmental Engineering*, 3(11), 389-393.

- Clements, G. R., D. M. Rayan, S. A. Aziz, K. Kawanishi, C. Traeholt, D. Magintan, R. Tingley, 2012: Predicting the distribution of the Asian tapir in Peninsular Malaysia using maximum entropy modelling, *Integrative Zoology*, 7(4), 400-406.

- Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, J. J. Lawler, 2007: Random forests for classification in ecology, *Ecology*, 88(11), 2783-2792.

- De’ath, G., K. E. Fabricius, 2000: Classification and regression trees: a powerful yet simple technique for ecological data analysis, *Ecology*, 81(11), 3178-3192.

- Elitez, I., C. Yaltirak, G. Sunal, 2018: A new chronostratigraphy (40Ar-39Ar and U-Pb dating) for the middle section of the Burdur-Fethiye Shear Zone, SW Turkey (eastern Mediterranean), *Turkish Journal of Earth Sciences*, 27(5), 405-420.

- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, N. E. Zimmermann, 2006: Novel methods improve prediction of species' distributions from occurrence data, *Ecography*, 29(2), 129-151.

- Evans, J. S., S. A. Cushman, 2009: Gradient modeling of conifer species using random forests, *Landscape Ecology*, 24, 673-683.

- Evans, J. S., M. A. Murphy, Z. A. Holden, S. A. Cushman, 2010: Modeling species distribution and change using random forest, *In Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications,* New York, NY: Springer New York, pp, 139-159.

- Fajardo, J., J. Lessmann, E. Bonaccorso, C. Devenish, J. Muñoz, 2014: Combined use of systematic conservation planning, species distribution modelling, and connectivity analysis reveals severe conservation gaps in a megadiverse country (Peru), *PloS One*, 9(12), e114367.

- Ferrier, S., G. Watson, J. Pearce, M. Drielsma, 2002: Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling, *Biodiversity & Conservation*, 11(12), 2275-2307.

- Filz, K. J., T. Schmitt, 2015: Niche overlap and host specificity in parasitic Maculinea butterflies (Lepidoptera: Lycaenidae) as a measure for potential extinction risks under climate change, *Organisms Diversity & Evolution*, 15, 555-565.

- Franklin, J., F. W. Davis, M. Ikegami, A. D. Syphard, L. E. Flint, A. L. Flint, L. Hannah, 2013: Modelling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biology*, 19(2), 473-483.

- Garcia, K., R. Lasco, A. Ines, B. Lyon, F. Pulhin, 2013: Predicting geographic distribution and habitat suitability due to climate change of selected threatened forest tree species in the Philippines, *Applied Geography*, 44, 12-22.

- Grossmann, E., J. Ohmann, J. Kagan, H. May, M. Gregory, 2010: Mapping ecological systems with a random forest model: Trade-offs between errors and bias, *Gap Analysis Bulletin*, 17(1), 16-22.

- Guisan, A., T. C. Edwards Jr, T. Hastie, 2002: Generalized linear and generalized additive models in studies of species distributions: setting the scene, *Ecological modelling*, 157(2-3), 89-100.

- Gülsoy, S., K. Özkan, 2013: Determination of environmental factors and indicator plant species for site suitability assessment of Crimean Juniper in the Acipayam District, Turkey, *Sains Malaysiana*, 42(10), 1439-1447.

- Hamann, A., S. N. Aitken, 2013: Conservation planning under climate change: accounting for adaptive potential and migration capacity in species distribution models, *Diversity and Distributions*, 19(3), 268-280.

- Hernandez, P. A., C. H. Graham, L. L. Master, D. L. Albert, 2006: The effect of sample size and species characteristics on performance of different species distribution modelling methods, *Ecography*, 29(5), 773-785.

- Hobbs, R. J. 1994: Can revegetation assist in the conservation of biodiversity in agricultural areas? *Pacific Conservation Biology*, 1(1), 29-38.

- Holden, J. E., W. H. Finch, K. Kelley, 2011: A comparison of two-group classification methods, *Educational and psychological measurement*, 71(5), 870-901.

- Hosmer, D. W., S. Lemeshow, R. X. Sturdivant, 2000: Introduction to the logistic regression model, *Applied Logistic Regression*, 2, 1-30.

- Jackson, L. L., N. Lopoukhine, D. Hillyard, 1995: Ecological restoration: a definition and comments, *Restoration Ecology*, 3(2), 71-75.

- Jenness, J. 2006: Topographic Position Index (TPI) extension for ArcView 3.x, v. 1.2. Jenness Enterprises.

- Jones, T. G., H. R. Ratsimba, L. Ravaoarinorotsihoarana, G. Cripps, A. Bey, 2014: Ecological variability and carbon stock estimates of mangrove ecosystems in northwestern Madagascar, *Forests*, 5(1), 177-205.

- Kandemir, G. E., Z. Kaya, F. Temel, S. Önde, 2010: Genetic variation in cold hardiness and phenology between and within Turkish red pine (*Pinus brutia* Ten.) populations: implications for seed transfer, *Silvae Genetica*, 59(1-6), 49-57.

- Kılıç, M., Ş. T. Güner, 2000: Gölhisar Red pine stands, *Orman Mühendisliği*, 37(5), 18-21.

- Kıraç, A., A. Mert, 2019: Will Danford's lizard become extinct in the future, *Polish Journal of Environmental Studies*, 28(3), 1741-1748.

- Kramer-Schadt, S., J. Niedballa, J. D. Pilgrim, B. Schröder, J. Lindenborn, V. Reinfelder, A. Wilting, 2013: The importance of correcting for sampling bias in MaxEnt species distribution models, *Diversity and distributions*, 19(11), 1366-1379.

- Kumar, P. 2012: Assessment of impact of climate change on Rhododendrons in Sikkim Himalayas using Maxent modelling: limitations and challenges, *Biodiversity and Conservation,* 21, 1251-1266.

- Kumar, S., T. J. Stohlgren, 2009: Maxent modelling for predicting suitable habitat for threatened and endangered tree Canacomyrica monticola in New Caledonia, *Journal of Ecology and Natural Environment*, 1(4), 94-98.

- Lehmann, A., J. M. Overton, J. R. Leathwick, 2002: GRASP: generalized regression analysis and spatial prediction, *Ecological Modelling*, 157(2-3), 189-207.

- Ma, S. M., Y. B. Nie, Q. L. Geng, R. X. Wang, 2014: Impact of climate change on suitable distribution range and spatial pattern in Amygdalus mongolica, *Chinese Journal of Plant Ecology*, 38(3), 262-269.

- Malhi, Y., J. T. Roberts, R. A. Betts, T. J. Killeen, W. Li, C. A. Nobre, 2008: Climate change, deforestation, and the fate of the Amazon, *Science*, 319(5860), 169-172.

- Mann, H. B., D. R. Whitney, 1947: On a test of whether one of two random variables is stochastically larger than the other, *The Annals of Mathematical Statistics*, 18(1), 50-60.

- Marcer, A., L. Sáez, R. Molowny-Horas, X. Pons, J. Pino, 2013: Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation, *Biological Conservation*, 166, 221-230.

- Mateo, R. G., T. B. Croat, Á. M. Felicísimo, J. Muñoz, 2010: Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections, *Diversity and Distributions*, 16(1), 84-94.

- Matthews, E., R. Payne, M. Rohweder, S. Murray, 2000: *Forest Ecosystems*. World Resources Institute.

- Mert, A., A. Kıraç, 2019: GIS as a tool to map habitat suitability for two lizard species using environmental factors, *Fresenius Environmental Bulletin*, 28(2A), 1330-1336.

- Mert, A., K. Özkan, Ö. Şentürk, M. G. Negiz, 2016: Changing the potential distribution of Turkey Oak (*Quercus cerris* L.) under climate change in Turkey, *Polish Journal of Environmental Studies*, 25(4), 1633-1638.

- Moisen, G. G., E. A. Freeman, J. A. Blackard, T. S. Frescino, N. E. Zimmermann, T. C. Edwards, 2006: Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods, *Ecological Modelling*, 199(2), 176-187.

- Moisen, G. G., T. S. Frescino, 2002: Comparing five modelling techniques for predicting forest characteristics, *Ecological modelling*, 157(2-3), 209-225.

- Olsson, P. A., L. M. Martensson, H. H. Bruun, 2009: Acidification of sandy grasslands–consequences for plant diversity, *Applied Vegetation Science*, 12(3), 350-361.

- Ortega-Huerta, M. A., A. Townsend Peterson, 2008: Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods, *Revista Mexicana de Biodiversidad*, 79(1), 205-216.

- Özdemir, S. 2018: Potential distribution modelling and mapping using Random Forest method: an example of Yukarıgökdere district, *Turkish Journal of Forestry*, 19(1), 51-56.

- Özkan, K. 2012a: Identification of driving factors for presence of endemic species in the Aglasun district of the Mediterranean Region, Turkey, *Polish Journal of Ecology,* 60(2), 291-300.

- Özkan, K. 2012b: Modelling ecological data using classification and regression tree technique (CART), *Turkish Journal of Forestry*, 13(1), 1-4.

- Özkan, K. 2013: Using the non-parametric classifier CART to model Lebanon cedar (*Cedrus libani* A. Rich) distribution in a mountain Mediterranean Forest district, *Polish Journal of Environmental Studies*, 22(2), 495-501.

- Özkan, K., E. Kuzugüdenli, 2010: The relations between site index of brutian pine (*Pinus brutia* Ten.) and ecological site factors in Sütçüler district from the Mediterranean region, *Turkish Journal of Forestry*, 11(1), 16-29.

- Özkan, K., Ö. Şentürk, A. Mert, M. G. Negiz, 2015: Modeling and mapping potential distribution of Crimean juniper (*Juniperus excelsa* Bieb.) using correlative approaches, *Journal of Environmental Biology*, 36(1), 9-15.

- Özüdoğru, Ş. 2018: Cibyra in late antiquity, *Cedrus*, 6, 13 – 64.

- Parker, K. C. 1988: Environmental relationships and vegetation associates of columnar cacti in the northern Sonoran Desert, *Vegetatio*, 78, 125-140.

- Pearson, K. 1990: On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157-175.

- Pearson, R. G., C. J. Raxworthy, M. Nakamura, A. Townsend Peterson, 2007: Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar, *Journal of Biogeography*, 34(1), 102-117.

- Peng, C. Y. J., K. L. Lee, G. M. Ingersoll, 2002: An introduction to logistic regression analysis and reporting, *The Journal of Educational Research*, 96(1), 3-14.

- Phillips, S. J., R. P. Anderson, R. E. Schapire, 2006: Maximum entropy modelling of species geographic distributions, *Ecological modelling*, 190(3-4), 231-259.

- Phillips, S. J., M. Dudík, 2008: Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation, *Ecography*, 31(2), 161-175.

- Phillips, S. J., M. Dudík, R. E. Schapire, R. E. 2004: A maximum entropy approach to species distribution modelling, *In Proceedings of the twenty-first international conference on Machine learning,* pp, 83.

- Rameshprabu, N., P. S. Swamy, 2015: Prediction of environmental suitability for invasion of Mikania micrantha in India by species distribution modelling, J*ournal of Environmental Biology*, 36(3), 565.

- Ramirez-Villegas, J., F. Cuesta, C. Devenish, M. Peralvo, A. Jarvis, C. A. Arnillas, 2014: Using species distributions models for designing conservation strategies of Tropical Andean biodiversity under climate change, *Journal for Nature Conservation*, *22*(5), 391-404.

- Ray, R., K. V. Gururaja, T. V. Ramchandra, 2011: Predictive distribution modelling for rare Himalayan medicinal plant Berberis aristate DC, *Journal of Environmental Biology*, 32(6), 725.

- Roberts, D. W., S. V. Cooper, 1989: Concepts and techniques of vegetation mapping, *Land Classifications Based on Vegetation: Applications for Resource Management*, 90-96.

- Robertson, M. P., C. I. Peter, M. H. Villet, B. S. Ripley, 2003: Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modelling techniques, *Ecological Modelling*, 164(23), 153-167,

- Rodriguez-Galiano, V. F., B. Ghimire, J. Rogan, M. Chica-Olmo, J. P. Rigol-Sanchez, 2012: An assessment of the effectiveness of a random forest classifier for land-cover classification, *ISPRS Journal of Photogrammetry and Remote Sensing,* 67, 93-104.

- Rubel, F., K. Brugger, K. Haslinger, I. Auer, 2017: The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800–2100, *Meteorologische Zeitschrift*, 26(2), 115-125.

- Sarmiento-Ramirez, J. M., E. Abella-Perez, A. D. Phillott, J. Sim, P. Van West, M. P. Martin, J. Dieguez-Uribeondo, 2014: Global distribution of two fungal pathogens threatening endangered sea turtles, *PloS One*, 9(1), e85853.

- Shabani, F., L. Kumar, M. Ahmadi, 2016: A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area, *Ecology and Evolution*, 6(16), 5973-5986.

- Süel, H., 2019: Brown bear (*Ursus arctos*) habitat suitability modelling and mapping, *Applied Ecology and Environmental Research,* 17(2), 4245-4255,

- Şenel, S., B. Alatli, 2014: A review of articles that used logistic regression analysis, *Journal of Measurement and Evaluation in Education and Psychology*, 5(1), 35-52.

- Şentürk, Ö. 2012: Modelling of potential distribution areas of the priority target forest tree species in Sütçüler district, Ph. D. Thesis, Suleyman Demirel University Graduate School of Applied and Natural Sciences Department of Forestry Engineering, 180 pp

• Şentürk, Ö., S. Gülsoy, İ. Tümer, 2019: Potential distribution modelling and mapping of Brutian pine stands in the inner parts of the Middle Black Sea Region in Turkey, *Polish Journal of Environmental Studies*, 28(1), 321-327.

• Varol, S. 2011: Hyrogeochemical and hydrogeological properties of Tefenni (Burdur) plain and evaluation for Medical Geology, Doctoral dissertation, Suleyman Demirel University, pp, 269.

• Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, A. Guisan, 2008: Effects of sample size on the performance of species distribution models, *Diversity and Distributions*, 14(5), 763-773.

• Yee, T. W., N. D. Mitchell, 1991: Generalized additive models in plant ecology, *Journal of Vegetation Science*, 2(5), 587-602.

• Yılmaz, E., I. Çiçek, 2016: Thornthwaite climate classification of Turkey, *Journal of Human Sciences*, 13(3), 3973-3994.

SAŽETAK

Cilj ovoga rada bio je razjasniti temeljne karakteristike i područja primjene različitih tehnika modeliranja koje se koriste u aktualnim istraživanjima ekološkog modeliranja. Koristeći pet različitih metoda modeliranja distribucije, provedeno je modeliranje i mapiranje distribucije brucijskog bora u okrugu Gölhisar. Uzorci su prikupljeni s brucijskog bora na 400 pokusnih ploha u tom području. U modelima su se koristile sljedeće varijable: nadmorska visina, nagib, orijentacija, indeks zračenja, toplinski indeks, indeks topografskog položaja i geološka podloga. Od metoda modeliranja distribucije vrsta korištene su logistička regresija, klasifikacijo stablo, slučajna šuma, generalizirani aditivni model i maksimalna entropija. Izrađene su ROC krivulje te je učinkovitost modela distribucije vrsta procijenjena pomoću AUC krivulje. Statističke analize otkrile su da su najbolji modeli redom generalizirani aditivni model, slučajna šuma, klasifikacijsko stablo, maksimalna entropija i logistička regresija. Nadmorska visine i vrsta geološke podloge imale su najveći utjecaj u modelima distribucije brucijskog bora. Mapirani su rezultati generaliziranog aditivnog modela koji su imali najveću AUC vrijednost. Utvrđene su neke ekološke i statističke razlike između modela te su prikazani njihovi razlozi. U usporedbi s metodama koje se obično koriste za modeliranje distribucije vrsta, metoda generaliziranog aditivnog modela ima specifičnu funkciju zaobljenosti koja osigurava usklađivanje između promjena okoliša i eksplanatornih krivulja te točniju ekološku interpretaciju dobivenih modela.

KLJUČNE RIJEČI: klasifikacijsko stablo, ekološko modeliranje, generalizirani aditivni model, logistička regresija, MaxEnt, slučajna šuma