DEVELOPMENT OF AN ENSEMBLE CLASSIFIER WITH DATA FROM DESCRIPTION SHEETS TO CLASSIFY FOREST STANDS IN SITE QUALITIES

RAZVOJ KOMPOZITNOG KLASIFIKATORA S PODACIMA IZ OPISNIH LISTOVA ZA KLASIFIKACIJU BONITETA ŠUMSKIH SASTOJINA

Kyriaki KITIKIDOU*[1], Elias MILIOS[2], Panagiota PALAVOUZI[3]

Abstract

Aim of study: In this work, we tested the technique of combining the predictions of classifiers for the development of a single, ensemble classifier, in order to classify forest stands in site qualities. **Area of study**: We used data of the forest stands of Dadia – Lefkimi – Soufli forest (north-eastern Greece). **Materials and methods**: The variables that we used as input were the altitude, slope, age and canopy density. For the ensemble classifier development we applied the boosting algorithm. **Main results**: The canopy density was the most important predictor; topography which replaced altitudes and slopes was the second important predictor, while the developed ensemble classifier gave a percentage of correct classification up to 98.59% (for the worst site quality). **Research essentials**: If we consider that the initial site classification comprised over 70% of the Dadia-Lefkimi –Soufli forest area in the worst site quality, then the usage of boosting method for creating a collective classifier for site qualities in the studied forest can be characterized as fully successful. The application of this method using these input parameters do not need background information regarding the tree age and (or) other difficult to access information. Moreover, in a quite high degree, this site classification is not influenced by disturbances. The boosting method for creating a collective classifier for site qualities obviously will give far more accurate classifications of site productivity, if a more sophisticated scheme of data collection is used.

KEY WORDS: ensemble classifiers; forest stands; site qualities.

INTRODUCTION
UVOD

The technique of combining the predictions of many classifiers, for the creation of a single, collective classifier (ensemble classifier), has preoccupied researchers (Breiman 1996, Clemen 1989, Perrone 1993, Wolpert 1992, Opitz and Shavlik 1996). A collective classifier is effective when it is more accurate than any classifier that participates in the

Department of Forestry and Management of the Environment and Natural Resources, Democritus University, Pantazidou 193, 68200, Orestiada, Greece
[1] Kyriaki Kitikidou, Assistant Profesor, kkitikid@fmenr.duth.gr
[2] Elias Milios, Associate Profesor, emilios@fmenr.duth.gr
[3] Panagiota Palavouzi, Ms, giota_pal@hotmail.com
* Corresponding author

formation of a group and consists of classifiers whose errors are at different points in the vector space of input. The two most popular methods of creating collective classifiers are the method of boosting and the method of bagging (Freund and Schapire 1997, Drucker 1997). These techniques are based on data sampling (training data) and result in different training data sets for each classifier of the total classification system (Opitz and Shavlik 1996).

In a typical scenario of supervised learning, a set of samples are available, which are called the training set. The classes of these samples are known and the goal is to construct a model that would classify new samples in classes. The learning algorithm that builds the model is called an inducer. The basic idea of collective classification is the weighting of different classifiers and combining them into one single classifier, which performs better than each of the individual classifiers. When making a decision, people follow the same technique, taking various opinions and then evaluating those views for making the final decision (Rokach 2009).

The productivity of a forest is described as the site's ability to produce timber or forest biomass (Skovsgaard and Vanclay 2013). Various approaches have been developed to site productivity assessment (Pokharel and Dech 2011). However, the typical approach for site quality assessment is based on the strong correlation between height growth and volume. Hence, the site index has become an important tool in assessing site productivity (Clutter et al. 1983). Heigh-age observations are plotted on a graph and are used in assessing site productivity (Laubhann et al. 2009). Site index curves are developed via three methods: the parameter prediction method, the guide curve method and the difference equation method (Clutter et al. 1983).

The purpose of this paper is to apply the boosting method for creating a collective classifier, which classifies forest stands in site qualities, with input of the altitude, slope, age and canopy density.

MATERIALS AND METHODS
MATERIJALI I METODE

Study area – Data collection – *Područje istraživanja – Prikupljanje podataka*

Data collection came from the management plan of the forest of Dadia-Lefkimi -Soufli (Consorzio Forestale Del Ticino 2005). Situated at the southeast end of the Rhodope mountain range in northeaster Greece, at the crossroads of two continents, the National Park of Dadia-Lefkimi-Soufli Forest is of exceptional ecological significance at the European level (UNESCO World Heritage Centre 2012-2015). The elevation ranges from 10 to 604 m. Soils are shallow to moderate deep, exhibiting various textures. The mean annual temperature is 14.3 °C and the mean annual precipitation is 652.9 mm (Consorzio Forestale Del Ticino 2005).

The main tree species of the forest are *Pinus brutia* and *Quercus* spp. These species create pure and mixed formations; there is also *Pinus nigra*, mainly in reforestations, and as individual trees (Consorzio Forestale Del Ticino 2005). From 403 description sheets (i.e. 403 records-cases), the minimum and maximum altitude (m), the minimum and maximum slope (%), the minimum and maximum age of trees (in years), the minimum and maximum canopy density and site quality (quality I, II, III) were used. Quality III included the stands where, as referred in the description sheets, the site quality was III, IV or V. Consequently, three training sets were selected: one for site quality I (25 stands), one for site quality II (94 stands), and one for site quality III (284 stands).

In the description sheets of each stand, more than one site quality categories appear. These categories are not referred to different species. The determination of these categories was based on the composition of each stand and on the degree of species mixture. The initial site characterization for each species was made using site index curves. These site index curves were developed for each of these species in forests of other Greek areas (Consorzio Forestale Del Ticino 2005).

The altitudes and slopes were replaced by one variable, named topography, applying the Anderson-Rubin (1956) method. This method adjusts regression least squares formula to produce factor scores, uncorrelated with other factors, and uncorrelated with each other. The vector of observed variables, i.e. altitudes and slopes, is multiplied by the inverse of a diagonal matrix of their variances. The resulting new variable has an average of 0 and a standard deviation of 1. Using the same method, the ages were replaced with one variable, named age and the canopy densities with the variable canopy density. The new variables topography, age and canopy density were used as predictors, i.e. as input, while the site quality was used as the target variable.

As canopy density is considered the ratio of the sum of the areas of canopy projections (if we put one projection next to the other) divided by the area that these trees occupy (Dafis 1992). Canopy density can take values from 0 to over 1. The more the crowns are tangled, the higher the canopy density becomes. Canopy density is different than ground cover, since ground cover can take values up to 1.

In this study, different site classification predictions could have been developed for pine, oak and pine – oak formations. Unfortunately, even though the area of the mixtures and the species are given in the description sheets, parameters used for site classification were presented for the total stand area. This is the reason why one site classification, regardless of the species, was developed. Moreover, for the same reason, the dominant site quality (occupying the largest forested area of the stand) was used for the site productivity characterization of the stand. There weren't any different topographic features or canopy density data for different site qualities in each description sheet.

Boosting algorithm – *Algoritam jačanja klasifikatora*

For the creation of a single collective classifier (ensemble classifier), that will classify sample trees and/or sample surfaces in site qualities, with the input of the topographical features (altitude and slope), age and canopy density, we implemented the boosting algorithm (Freund and Schapire 1997, Drucker 1997). The application of the algorithm was done with the statistical package SPSS v. 21.0 (IBM 2012).

The boosting method can be used with any type of model and can reduce variance and bias in the forecast, i.e. to increase the accuracy of the model. Boosting produces a sequence of components, namely the main (base) models, each of which shall be drawn up by the entire set of data. Before drawing each successive component, records-cases are weighted based on the errors (residuals) of the previous component. The cases with large residuals relatively weight more, so the next component focuses on better predictions of these cases. All together the components-models compose a single model (ensemble model). The single model provides values for new records using a combination rule, depending on the scale of measurement of the target variable, i.e. the dependent variable (analog or categorical).

Boosting model measures – *Mjere modela jačanja klasifikatora*

Accuracy is calculated for the naïve model, the reference model, i.e. the simple linear model, without application of boosting and bagging, for the ensemble model and for the basic models.

For categorical target variables, the accuracy is (IBM 2012):

$$\frac{1}{N}\sum_{k=1}^{K} f_k \mathbf{II}\left(y_k == \hat{y}_k^m\right)$$

where:

N Total number of records. $N = \sum_{k=1}^{K} f_k$.

K Number of records-cases in the training dataset.

$\mathbf{II}(\pi)$ For any condition π, $\mathbf{II}(\pi)$ is 1 if π hold, and 0 otherwise.

f_k Frequency for the k-th record.

y_k Target value for the k-th record.

$\hat{y}_k^m = T^m(X_k)$ Predicted target value of the k-th record of the m-th bootstrap sample.

T^m Model for the m-th bootstrap sample.

X_k Predictions for the k-th record.

For the naïve model, $\hat{y}_k$ is the mode for categorical target variables (IBM 2012).

Table 1. Descriptive statistics of the variables.
Tablica 1. Opisna statistika analiziranih varijabli.

	Site – *Stanište*					
	I (25 stands) – *I (25 sastojina)*		**II** (94 stands) – *II (94 sastojina)*		**III** (284 stands) – *III (284 sastojina)*	
	Mean *Aritmetička sredina*	Standard Deviation *Standardna devijacija*	Mean Aritmetička sredina	Standard Deviation *Standardna devijacija*	Mean *Aritmetička sredina*	Standard Deviation *Standardna devijacija*
min_alt (m) *min_visina (m)*	164	113	161	75	179	82
max_alt (m) *max_visina (m)*	284	143	291	95	326	102
min_slope (%) *min_nagib (%)*	14	16	18	14	21	16
max_slope (%) *max_nagib (%)*	37	19	41	18	45	19
topography *topografija*	−0.39990	1.21886	−0.22293	0.89790	0.10899	0.99348
min_age (yrs) *min_starost (god)*	36	12	35	13	35	16
max_age (yrs) *max_ starost (god)*	81	21	81	23	80	21
age *starost*	0.07639	0.80742	0.02792	0.94724	−0.01597	1.03408
min_canopy density *min_gustoća krošnje*	0.7	0.1	0.7	0.1	0.6	0.1
max_canopy density *max_ gustoća krošnje*	1.1	0.1	1.0	0.1	1.0	0.1
canopy_density *gustoća krošnje*	0.52265	0.89964	0.01442	1.05272	−0.05078	0.98025

RESULTS
REZULTATI

Exploratory Data Analysis – *Istraživačka analiza podataka*

In table 1 the basic descriptive statistics of the variables of all data are given.

Importance of the estimators – *Važnost procjenitelja*

The term importance refers to the **relative importance** of each estimator in the forecasting model for the **given accuracy** of the model. The importance can take values from 0 and 1. The rank of estimators (the predictors) as to their relative importance is given in figure 1. If we focus on the accuracy of the classification of the stands in site qualities (the meaning of the boosting algorithm), most important, with importance equal to 0.79 is the canopy density, followed by the topography (importance = 0.21) and age (materiality < 0.01) (Figure 1).

The correct classification rate reached **98.59%**, in site quality III (properly classified the 280 of the 284 stands). The correct classification of stand in site qualities I and II was 4% and 77%, respectively. For a more analytical presentation of classification, one can refer to Table 2, which is derived from the confusion matrix (Table 3).

DISCUSSION
RASPRAVA

The main tool for the evaluation of site productivity or site quality is site index (Van Laar and Akca 2007). In site index models, the height growth of dominant, codominant or of trees having certain characteristics are used (Van Laar and Akca 2007). There is a huge number of publications regard-
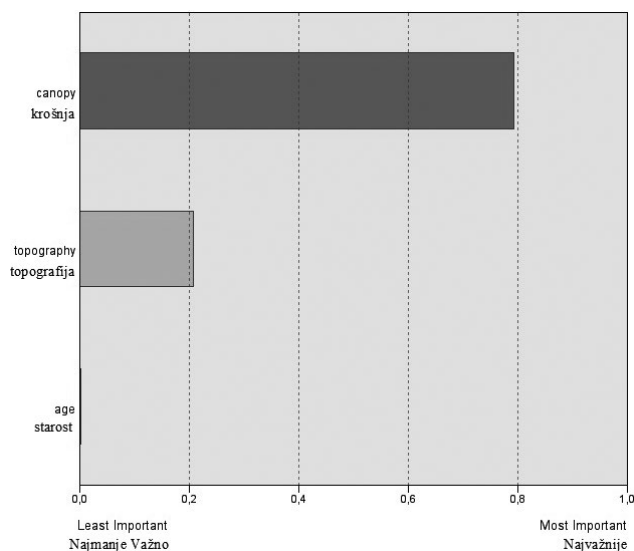

Figure 1. Importance of the predictors (boosting algorithm).
Slika 1. Važnost prediktora (algoritam jačanja).

ing site index models (Huang 1997; Krumland and Eng 2005). However, site quality is related to topographic features (see Dafis 1986; Vanclay 1992; Barnes et al. 1998; Papalexandris and Milios 2010; Milios et al. 2012) as well as to other environmental indicators like understory vegetation, soil physical and chemical properties, foliar nutrients, climatic data, soil properties etc. (see Kayahara et al. 1995, McKenney and Pedlar 2003, Bontemps and Bouriaud 2014, Watt et al. 2015). In Greece, soil depth is strongly related to site productivity, since it acts as a water storing reservoir during dry period (Dafis 1986; Hatzistathis & Dafis 1989) and it has been used in research studies for the site quality determination (Papalexandris and Milios 2010; Stambou-lidis et al 2013). Adamopoulos et al. (2009) and Adamo-

Table 2. Correct classification percentages.
Tablica 2. Točni postotci klasifikacije.

	Site – *Stanište*		
	I (25 stands) – *I (25 sastojina)*	II (94 stands) – *II (94 sastojina)*	III (284 stands) – *III (284 sastojina)*
Accuracy $= \dfrac{\text{True Positive} + \text{True Negative}}{\text{total}}$ Točnost $= \dfrac{\text{Istinito Pozitivno} + \text{Istinito Negativno}}{\text{ukupno}}$	0.25	23.33	69.48
Recall $= \dfrac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$ Opoziv $= \dfrac{\text{Istinito Pozitivno}}{\text{Istinito Pozitivno} + \text{Lažno Negativno}}$	4.00	77.05	98.59
Precision $= \dfrac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$ Preciznost $= \dfrac{\text{Istinito Pozitivno}}{\text{Istinito Pozitivno} + \text{Lažno Negativno}}$	100.00	100.00	100.00

Table 3. Confusion matrix.
Tablica 2. Matrica konfuzije.

	Site I – *Stanište I*	
	Predicted Positive *Predviđeno pozitivno*	Predicted Negative *Predviđeno negativno*
Condition Positive *Uvjet pozitivno*	1	24
Condition Negative *Uvjet negativno*	0	0
	Site II – *Stanište II*	
	Predicted Positive *Predviđeno pozitivno*	Predicted Negative *Predviđeno negativno*
Condition Positive *Uvjet pozitivno*	94	28
Condition Negative *Uvjet negativno*	0	0
	Site III – *Stanište III*	
	Predicted Positive *Predviđeno pozitivno*	Predicted Negative *Predviđeno negativno*
Condition Positive *Uvjet pozitivno*	280	4
Condition Negative *Uvjet negativno*	0	0

poulos et al. (2012) have used soil depth in combination to several other parameters, for the site quality distinction in reforestations found in the study area of the present study.

Similarly, topography is the second more important variable (importance = 0.21) in the classification of the stands in site qualities using boosting method. Topography incorporates slope, and slope is related to soil depth. Usually higher slope leads to lower soil depth (see Dafis 1986). In the present study, low productivity sites (site quality III) exhibit greater lowest and highest slope, compared to productive and medium productivity sites (site qualities I and II).

The most important variable (importance = 0.79) in the classification of the stands in site qualities is canopy density. A higher value of canopy density implies higher amount of photosynthetic tissues for a given area, being an indication of greater production ability. Canopy photosynthetic capacity is related to site productivity (Coops et al. 1998). According to Dafis (1986), in productive sites plants can endure heavier shade. As a result, in productive areas a higher canopy density is expected, compared to that of medium and low productive locations, since under the crowns of overstory more trees can survive. Moreover, in poor sites trees' crowns are small, as a result of slow or lack of differentiation (Oliver and Larson 1996). In the present study, low productivity sites (site quality III) exhibit lower lowest canopy density, compared to sites of high and medium productivity (site qualities I and II). However, canopy density is influenced by many factors, such as silvicultural treatments or other disturbances, stand development stage and species ecology. In particular, in the present study cuttings from the local residents and grazing strongly affected vegetation

structure in the past, while silvicultural interventions in the frame of forest management are taking place (Consorzio Forestale Del Ticino 2005). Moreover, as it is referred in the description sheets of the stands, pine formations (mainly) of *Pinus brutia*, oak formations as well as mixed formations of pines and oaks occur in the study area (Consorzio Forestale Del Ticino 2005). The different shade tolerance of the species, with their different position in the stand stories (overstory, middlestorey or understory) undoubtedly determines canopy density.

Even though canopy density is estimated in the form of lowest and highest values, and the influence of species mixture and stand vertical structure are not weighted up in the model, the correct classification of stand in site qualities I and II was 4% and 77%, respectively, while in site quality III was 99%. If we consider that the initial site classification comprised over 70% of the Dadia-Lefkimi –Soufli forest area in site quality III (Consorzio Forestale Del Ticino 2005), then the usage of the boosting method for creating a collective classifier for site qualities in the studied forest can be characterized as fully successful.

The development of alternative methods of site classification, using topographic features instead of growth of trees having certain characteristics, will be notably useful in bare and intensively degraded areas with sparse distribution of plants (see Smith et al. 1997). In this case, the usage of three site productivity categories (site qualities) instead of five or six is more practical and feasible in restoration of former forest lands. Furthermore, in areas were microtopography is highly variable, resulting in chances of site condition in rather small spatial scales (see also Papalexandris and Milios 2010), topographic features in combination with canopy density estimations can be used for site classification. In this case, two or three site productivity categories are preferable, since the frequent changes of site conditions make difficult to apply a more complex system. On the other hand, the adoption of canopy density as a variable for site classification models is very efficient, in cases where the vegetation is disturbed and the growth of most trees has been affected by these disturbances. Site index models based on dominant and codominant trees assume that these trees where always in that status and their growth did not retarded from disturbance factors (Smith et al. 1997, Raulier et al. 2003). Highest and lowest canopy density seems to be affected in a lower degree than the growth of dominant and codominant trees by disturbances (including silvicultural treatments). Only in extreme cases where trees' density in the whole area has been determined by intense disturbances, then the two extreme values of canopy density is affected. On the other hand, mean canopy density is a parameter that is more sensitive to disturbances that influence trees' density in any area of the stand.

In this study, site classification is not based on soil analysis (Bravo and Montero 2001; McKenney and Pedlar 2003) or

phytosociological surveys and (or) other vegetation analysis approaches that are strongly affected by disturbances (Barnes et al. 1998). Apart other deficits, the collection of this kind of data and analysis are expensive (see Smith et al. 1997), compared to the parameters used in the present study, since no background information regarding the tree age or other difficult to access information are needed. Moreover, in a quite high degree, the site classification using these parameters is not influenced by disturbances.

The boosting method for creating a collective classifier for site qualities, obviously will give far more accurate classifications of site productivity if a more sophisticated scheme of data collection is used. Since site productivity may present significant spatial variation (Skovsgaard and Vanclay 2013), a stand can exhibit more than one site productivity categories. Consequently, stand area must be divided in parts, regarding some characteristics that strongly influence productivity. Topography is such a factor. Thus, in the first step, a stand is divided in rather homogenous areas regarding the shape of the terrain (conceive, convex, inclined plane and other), the exposure or another characteristic that is considered crucial for site productivity determination. In a second step, inside these areas, easily measured or estimated parameters for the site classification, as in the present case, can be used for the estimation of site productivity, such as highest, lowest and mean slope, highest and lowest canopy density. The mean canopy density can be used in areas that are not strongly influenced by disturbances. Moreover, other parameters can be used for the estimation of site productivity; thickness or other characteristics of organic layer of forest floor are some of them (see Laamrani et al. 2014).

Some of the above – mentioned parameters can be easily measured using remote sensing data, but for others like canopy density, field observations are obligatory for their estimation. Pinno et al. (2009) refer that field measurements are needed in order to have precision in the prediction of site productivity within a forest management unit for *Populus tremuloides* in the Boreal Shield of Quebec.

In the case of mixed stands, either in groups or in a tree to tree basis site productivity, classification has to be conducted for each species and tree to tree mixture (Aertsen et al. 2012; Skovsgaard and Vanclay 2013). So, for each species or mixture, the above mentioned two-step process have to be applied.

CONCLUSIONS
ZAKLJUČCI

The usage of boosting method for creating a collective classifier for site qualities in the studied forest can be characterized as fully successful. The application of this method using altitude, slope, age and canopy density as input, do not need background information regarding the tree age and

(or) other information that is difficult to access. Moreover, in a quite high degree, this site classification is not influenced by disturbances. The boosting method for creating a collective classifier for site qualities, obviously will give far more accurate classifications of site productivity, if a more sophisticated scheme of data collection is used.

REFERENCES
LITERATURA

- Adamopoulos, S, Milios, E, Doganos, D, Bistinas, I, 2009: Ring width, latewood proportion and dry density in stems of Pinus brutia Ten. European Journal of Wood Products 67(4), pp: 471-477.
- Adamopoulos, S, Wimmer, R, Milios, E, 2012: Tracheid length – growth relationships of young Pinus brutia Ten. grown on reforestation sites. IAWA Journal 33(1), pp: 39-49.
- Aertsen, W, Kint, V, Von Wilpert, K, Zirlewagen, D, Muys, B, Van Orshoven, J, 2012: Comparison of location-based, attribute-based and hybrid regionalization techniques for mapping forest site productivity. Forestry 85(4), pp: 539-550.
- Anderson, R. Rubin, H, 1956: Statistical inference in factor analysis. Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability 5, pp: 111-150.
- Barnes, B, Zak, D, Denton, S, Spurr S, 1998: Forest Ecology. 4th ed. New York: John Willey & Sons Inc. pp: 792.
- Bontemps, J, Bouriaud, O, 2014: Predictive approaches to forest site productivity: Recent trends, challenges and future perspectives. Forestry 87(1), pp: 109-128.
- Bravo, F, Montero, G, 2001: Site index estimation in Scots pine (Pinus sylvestris L.) stands in the High Ebro Basin (northern Spain) using soil attributes. Forestry 74(4), pp: 395-406.
- Breiman, L, 1996: Bagging predictors. Machine Learning, 24(2), pp:123-140.
- Clemen, R, 1989: Combining forecasts: A review and annotated bibliography. International Journal of Forecasting, 5(4), pp:559-583.
- Clutter, J, Fortson, J, Pienaar, L, Brister, G, Bailey, R, 1983: Timber management: a quantitative approach. 1st ed. Wiley, New York. 333 pp.
- Consorzio Forestale Del Ticino, 2005: Study of protection and management of the public forest Dadia – Lefkimi – Soufli, for the period 2005-2014.
- Coops, N, Waring R, Landsberg, J, 1998: Assessing forest productivity in Australia and New Zealand using a physiologically-based model driven with averaged monthly weather data and satellite-derived estimates of canopy photosynthetic capacity. Forest Ecology and Management 104(1-3), pp: 113-127.
- Dafis, S, 1986: Forest Ecology (in Greek). Thessaloniki: Giahoudi Giapouli.
- Dafis, S, 1992: Applied silviculture (in Greek). Giahoudi-Giapoulis, Thessaloniki, Greece.
- Drucker, H, 1997: Improving regressor using boosting techniques. In: Proceedings of the 14th International Conferences on Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp: 107-115.
- Freund, Y, Schapire, R, 1997: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences, 55(1), pp: 119-139.

- Hatzistathis, A, Dafis, S, 1989: Reforestations – Forest Nurseries. (in Greek). Thessaloniki: Giahoudi Giapouli.
- Huang, S, 1997: Development of compatible height and site index models for young and mature stands within an ecosystem-based management framework. In: Empirical and process based models for forest tree and stand growth simulation; Amaro A, Tomé M (eds). pp: 61-98, Lisboa.
- IBM Corp. 2012: IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.
- Kayahara, G, Carter, R, Klinka, K, 1995: Site index of western hemlock (Tsuga heterophylla) in relation to soil nutrient and foliar chemical measures. Forest Ecology and Management 74(1-3), pp: 161-169.
- Krumland, B, Eng, H, 2005: Site index systems for major young-growth forest and woodland species in northern California (No. 4). California Department of Forestry & Fire Protection.
- Laamrani, A, Valeria, O, Bergeron, Y, Fenton, N, Cheng, L, Anyomi, K, 2014: Effects of topography and thickness of organic layer on productivity of black spruce boreal forests of the canadian clay belt region. Forest Ecology and Management 330, pp: 144-157.
- Laubhann, D, Sterba, H, Reinds, G, De Vries, W, 2009: The impact of atmospheric deposition and climate on forest growth in European monitoring plots: An individual tree growth model. Forest Ecology and Management 258(8), pp: 1751-1761.
- McKenney, D, Pedlar, J, 2003: Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada. Forest Ecology and Management 175(1-3), pp: 497-507.
- Milios, E, Pipinis, E, Petrou, P, Akritidou, S, Kitikidou, K, Smiris, P, 2012: The Influence of Position and Site on the Height Growth of Young Populus tremula L. Ramets in Low Elevation Formations in Northeastern Greece. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 40(2), pp: 302-307.
- Oliver, C, Larson, B, 1996: Forest Stand Dynamics. New York: John Wiley & Sons, Inc. 509 pp.
- Opitz, D, Shavlik, J, 1996: Actively Searching for an Effective Neural Network Ensemble. Connection Science, 8(3-4), pp: 337-354.
- Papalexandris, C, Milios, E, 2010: Analysis of natural Fagus sylvatica L. s.l. regeneration in low elevation stands located in the central part of Evros region in the Northeast of Greece. Is sprout origin regeneration significant for the species maintenance? Plant Biosystems 144(4), pp: 784-792.
- Perrone, M. 1993: Improving regression estimation: averaging methods for variance reduction with extension to general convex measure optimization. Ph.D. thesis, Brown University, Providence RI.
- Pokharel, B., Dech, P. 2011: An ecological land classification approach to modelling the production of forest biomass. The Forestry Chronicle 87(1), pp. 23-32.
- Pinno B, Paré, D, Guindon, L, Bélanger, N, 2009: Predicting productivity of trembling aspen in the Boreal Shield ecozone of Quebec using different sources of soil and site information. Forest Ecology and Management 289(3), pp. 782-783.
- Raulier, F, Lambert, M, Pothier, D, Ung, C, 2003: Impact of dominant tree dynamics on site index curves. Forest ecology and management, 184(1), pp: 65-78.
- Rokach, L. 2009: Ensemble-based classifiers. Artificial Intelligence Review, 33(1-2), pp: 1-39.
- Skovsgaard, J, Vanclay, J, 2013: Forest site productivity: A review of spatial and temporal variability in natural site conditions. Forestry 86(3), pp: 305-315.
- Skovsgaard, J, Vanclay, J, 2008: Forest site productivity: a review of the evolution of dendrometric concepts for evenaged stands. Forestry 81(1), pp: 13-31.
- Smith, D, Larson, B, Kelty, M, Ashton, P, Mark S, 1997: The practice of silviculture. Applied Forest Ecology. New York: John Willey & Sons, Inc. 560 pp.
- Stampoulidis, A, Milios, E, Kitikidou, K, 2013: The regeneration of pure Juniperus excelsa Bieb. Stands in Pespa National Park in Greece. Šumarski list 137(3-4), pp: 163-172.
- UNESCO World Heritage Centre 2012-2015. 2015: http://whc.unesco.org/en/tentativelists/5856/
- Van Laar, A, Akca, A, 2007: Forest mensuration (Vol. 13). Springer Science & Business Media. 385 pp.
- Vanclay, J, 1992: Assessing site productivity in tropical moist forests: a review..Forest Ecology and Management 54(1-4), pp: 257-287.
- Watt, M, Dash, J, Bhandari, S, Watt, P, 2015: Comparing parametric and non-parametric methods of predicting Site Index for radiata pine using combinations of data derived from environmental surfaces, satellite imagery and airborne laser scanning. Forest Ecology and Management 357(1), pp: 1-9.
- Wolpert, D, 1992: Stacked generalization. Neural Networks, 5(2), pp: 241-259.

Sažetak

Cilj rada: U ovome smo radu testirali tehniku kombiniranja predikcija klasifikatora za razvoj jednog kompozitnog klasifikatora, kako bi se klasificirao bonitet šumskih sastojina. Područje istraživanja: Koristili smo podatke šumskih sastojina šuma Dadia – Lefkimi – Soufli (sjeveroistočna Grčka). Materijali i metode: Varijable koje smo koristili kao ulazne su visina, nagib, starost i gustoća krošnje. Za razvoj kompozitnog klasifikatora primijenili smo algoritam jačanja klasifikatora. Glavni rezultati: Gustoća krošnje je najvažniji predskazatelj; topografija koja zamjenjuje visinu i nagib je drugi važan prediktor, dok je razvijeni kompozitni klasifikator dao postotak točne klasifikacije do 98,59% (za najgoru kvalitetu staništa). Osnove istraživanja: Ako uzmemo u obzir da je početna klasifikacija staništa obuhvatila više od 70% šumskog područja Dadia-Lefkimi-Soufli najgore kvalitete staništa, onda se korištenje metode jačanja za stvaranje kolektivnog klasifikatora za kvalitetu staništa kod proučavanih šuma, može okarakterizirati kao potpuno uspješno. Primjena ove metode pomoću navedenih ulaznih parametara ne zahtijeva pozadinske informacije u vezi sa starosti stabla i (ili) drugih teško dostupnih informacija. Štoviše, u prilično visokom stupnju, ova klasifikacija staništa nije pod utjecajem poremećaja. Metoda jačanja za stvaranje kolektivnog klasifikatora za kvalitetu staništa, očito će dati puno preciznije klasifikacije produktivnosti staništa ako se koristi sofisticiranija shema prikupljanja podataka.

KLJUČNE RIJEČI: kompozitni klasifikator; šumska sastojina; kvaliteta staništa.