

Assessing the dominant height of oriental beech (*Fagus orientalis* L.) in relation to edaphic and physiographic variables in the Hyrcanian Forests of Iran

Seyed Jalil Alavi ⁽¹⁾, Kouros Ahmadi ^(1,3), Carsten F. Dormann ⁽²⁾, Josep M. Serra-Diaz ^(3,4,5), Zahra Nouri ⁽⁶⁾

⁽¹⁾ Tarbiat Modares University. Faculty of Natural Resources and Marine Sciences. Department of Forestry. Tehran (Iran).

E-mail: j.alavi@modares.ac.ir

⁽²⁾ University of Freiburg. Department of Biometry and Environmental System Analysis. Tennenbacher Straße, 4. D-79106 Freiburg (Germany).

⁽³⁾ Aarhus University. Department of Biology. Section for Ecoinformatics and Biodiversity. Ny Munkegade, 114. DK-8000 Aarhus (Denmark).

⁽⁴⁾ Aarhus University. Center for Biodiversity Dynamics in a Changing World (BIOCHANGE). Department of Biology. Ny Munkegade, 114. DK-8000 Aarhus (Denmark).

⁽⁵⁾ Université de Lorraine. AgroParisTech. INRAE. Silva. F-54000 Nancy (France).

⁽⁶⁾ University of Applied Science and Technology (UAST). Branch of Mazandaran Wood & Paper Industries. PO. BOX 4814936361. Mazandaran, Sari (Iran).

Received 5 February 2020, accepted 9 November 2020, available online 10 December 2020.

This article is distributed under the terms and conditions of the CC-BY License (<http://creativecommons.org/licenses/by/4.0>)

Description of the subject. This study evaluates the application of Boosted Regression Trees (BRT) for predicting beech dominant height in the Hyrcanian forests of Iran, inscribed as a UNESCO's World Heritage due to its remarkable biodiversity.

Objectives. It is widely accepted that tree growth can be influenced by a wide variety of factors such as climate, topography, soil conditions and competition for resources. The early dominant height of trees modelling studies used the multiple linear regression. The development of more advanced non-parametric and machine learning methods provided opportunities to overcome the nonlinear relationships in forest ecosystems.

Method. In this study, boosted regression trees was evaluated to model the dominant height of *Fagus orientalis* as the most important tree species in the Hyrcanian forest, Iran. Dominant height was related to soil and topographical variables, which are available for 190 sample plots covering all importance environmental gradients in the research area.

Results. The results indicated BRT were found to outperform for modelling beech dominant height. This technique showed that phosphorus, percentage nitrogen, magnesium and percentage sand were among the most important variables.

Conclusions. This study demonstrates the ability of BRT to accurately model the dominant height of oriental beech in relation to environmental predictors, and encourages its use in forest ecology.

Keywords. Site quality, forest modelling, response curves, machine learning, mixed forest.

Évaluation de la hauteur dominante du hêtre oriental (*Fagus orientalis* L.) par rapport aux variables édaphiques et physiographiques dans les forêts hyrcaniennes d'Iran

Description du sujet. Cette étude évalue l'application des arbres à régression stimulée (BRT) pour prédire la hauteur dominante des hêtres dans les forêts hyrcaniennes en Iran, inscrites au patrimoine mondial de l'UNESCO en raison de leur remarquable biodiversité.

Objectifs. Il est largement admis que la croissance des arbres peut être influencée par une grande variété de facteurs tels que le climat, la topographie, les conditions du sol et la concurrence pour les ressources. Les premières études de modélisation de la hauteur dominante des arbres ont utilisé la régression linéaire multiple. Le développement de méthodes non paramétriques et d'apprentissage machine plus avancées a permis de surmonter les relations non linéaires dans les écosystèmes forestiers.

Méthode. Dans cette étude, des arbres à régression amplifiée ont été évalués pour modéliser la hauteur dominante de *Fagus orientalis* en tant qu'espèce d'arbre la plus importante dans la forêt hyrcanienne, en Iran. La hauteur dominante était liée aux variables pédologiques et topographiques, qui sont disponibles pour 190 placettes échantillons couvrant tous les gradients environnementaux importants dans la zone de recherche.

Résultats. Les résultats ont montré que les BRT étaient plus performants pour la modélisation de la hauteur dominante des hêtres. Cette technique a montré que le phosphore, le pourcentage d'azote, le magnésium et le pourcentage de sable étaient parmi les variables les plus importantes.

Conclusions. Cette étude démontre la capacité des BRT à modéliser avec précision la hauteur dominante du hêtre oriental par rapport aux prédicteurs environnementaux, et encourage son utilisation en écologie forestière.

Mots-clés. Qualité de la station, modélisation forestière, courbes de réponse, apprentissage machine, forêt mélangée.

1. INTRODUCTION

Northern forests of Iran, called Hyrcanian or Caspian forests, cover a relatively narrow strip in the north of Iran, which are among the most important and valuable ecosystems inscribed in United Nations Educational, Scientific, and Cultural Organization (UNESCO) World Heritage List. Hyrcanian forests are important sources of genetic variation, biodiversity, commercial woody products, and of various environmental services (Ahmadi et al., 2013). Covering an area of about 1.85 million ha, these forests account for 15% of the total Iranian forests and 1.1% of the country's area. These forests range from sea level up to an altitude of 2,800 m and comprise various forest types, harboring approximately 80 woody species (trees and shrubs). Oriental beech (*Fagus orientalis* Lipsky), chestnut-leaved oak (*Quercus castaneifolia* C.A.Mey.), velvet maple (*Acer velutinum* Boiss.), hornbeam (*Carpinus betulus* L.) and Caucasian alder (*Alnus subcordata* C.A.Mey.) are among the main tree species in these forests. Hyrcanian forests, along with similar North American and East Asian forest communities, are nowadays seen as remnants of contiguous Tertiary deciduous belt (Sagheb Talebi et al., 2014), and hence one of the world's oldest extant forests. Today, these forests are regularly harvested, but their management is rarely based on assessments of growth, standing biomass or specific target forest composition and must, in places, be considered unsustainable.

Sustainable forest management requires a reliable estimation of wood quantity and quality. For the range of stand densities usually targeted in forest management, dominant height is not dependent on stand density and responds little to forest thinning, so that it reflects site growing conditions better than mean height or stand basal area growth. Dominant height is thus considered an important variable for specifying stand development and predicting the growth potential of a site (von Gadow & Hui, 1999; Nunes et al., 2011). More specifically, dominant height, as proxy for growth potential, is affected by the soil's chemical and physical characteristics, the topography and the climate at the stand level (Nambiar et al., 2004). Indeed, Remy de Perthuis de Laillevault in the 18th century proposed the application of height growth for assessing site quality in forest stands, leading to the development of a site index (Vanclay & Skovsgaard, 1997; Batho & Garcia,

2006). Unlike diameter at breast height, mean height and volume, dominant height usually does not show sensitivity to silvicultural treatments and variations in the stand density (Curtis & Reukema, 1970; Hogg & Nester, 1991). The dominant height, due to close relationship with volume, is considered as a good site productivity index (Carman, 1975; Hagglund, 1981; Clutter et al., 1983). A further important characteristic of dominant height is that its variability within sample plots which are used in site index studies is relatively low, but it shows high variability between sample plots; this latter component of variability is that which is expected to be linked to variations in the productivity of sites (Herrera et al., 1999).

It is widely accepted that the growth capacity of forest species can be influenced by a wide variety of complex interacting factors including climate, topography, soil conditions and competition for resources (Assmann, 1970; Oliver & Larson, 1996). In both forest ecology (Coomes & Allen, 2007) and sustainable forest management (Pretzsch, 2009), an understanding of the variation in tree growth is important. Numerous studies have focused on climate, geologic, topographic and soil factors (Ung et al., 2001; Palahı et al., 2004; Seynave et al., 2005) or used indicator plants for site quality assessment and classification (La Roi et al., 1988; Berger & Walther, 2006).

Understanding the main factors affecting on high and low production efficiency in forest areas is essential for both science and industry (Pretzsch et al., 2015). A better understanding also may contribute to design of forest production systems which are very important in sustainable forest management (Forrester, 2014). In many of the early studies, assessment of forest dominant height was related to environmental variables using multiple linear regression, mostly without considering non-linear effects and interactions among predictors; similarly, heterogeneous variances and distributional assumptions were regularly violated (Aertsens et al., 2010). The number and complexity of modelling techniques to cope with the inherent complexity of ecological problems has increased markedly over the past years (Moisen et al., 2006; Aertsens et al., 2010; Hegel et al., 2010). The development of advanced nonparametric and machine learning techniques and the growing availability to geodatasets with high spatial resolution have increased the accuracy for predicting forest characteristics (Aertsens et al., 2010).

Many of machine learning algorithms have been used in different fields of forest science including classification and regression trees (CART), random forest (RF), boosted regression tree (BRT), artificial neural networks (ANN), support vector machine (SVM), cubist (Cubist) and multivariate adaptive regression splines (MARS). These data-driven methods have already been successfully applied to ecology and remote sensing to perform tasks such as species distribution (Alavi et al., 2019; Ahmadi et al., 2020a). However, the applications in forest growth and yield prediction are still limited. Boosted regression trees (BRT) approach is a promising machine-learning technique, which is currently valuable tools for ecological modelling, and seems to be powerful tools for analyzing large datasets and identifying non-linear relationships (Drew et al., 2011). The performance of this model has been investigated by several authors in recent years, pointing to an overall suitability for a wide range of applications and data sets (Moisen & Frescino, 2002; McKenney & Pedlar, 2003; Segurado & Araujo, 2004; Elith et al., 2006; Elith & Graham, 2009; Aertsen et al., 2010; Opper et al., 2012; França & Cabral, 2015).

Oriental beech is an important late-successional and shade-tolerant species that occurs in mild mountainous or marine climates with high humidity. Although growing on every soil, beech trees are rare in moist and heavy soils (Tabari et al., 2005). The contents of loam play an important role in beech growth and height, so that the best beech forests in terms of diameter and height growth are located on semi-heavy and well-drained soils with sufficient moisture (Habibi Kaseb, 1974). Other studies found the importance of such characteristics as soil depth, silt, phosphorus, pH, bulk density and topographic variables as the main factors in distribution of beech communities (Salehi et al., 2005; Mataji et al., 2009). In terms of nutrients requirements of the oriental beech tree, there are contrasting studies. Regarding to the effects of physiography on beech trees, Marvi Mohajer (1976) found beech trees located at elevation between 900 to 1,500 m a.s.l. had the best status based on tree height (Marvi Mohajer, 1976). He emphasized the beech forests located in mid-lands of Caspian forests have higher forest site productivity. Beech trees prefer north-facing slopes and demand for atmosphere moisture and cool environment (Gorgi Bahri & Sagheb-Talebi, 1992; Sagheb-Talebi, 1996; Alavi et al., 2012).

About 18% of the total forest area, 30% of the standing volume, and 24% of the stem number in Hyrcanian forests in Iran are occupied by beech forests. Because of the importance of oriental beech as the most valuable wood-producing species of Hyrcanian forests, in this study, we evaluated the dominant height of this species as an important indicator of forest

productivity in relation to physiographic and edaphic variables using a well-known machine learning method (boosted regression tree). The results of present study allow forest managers to identify the main drivers of dominant height of beech trees obtained from machine learning models and have sustainable planning for the future of these forests.

2. MATERIALS AND METHODS

2.1. Study area

This research was conducted in Kheyroud Forest of Mazandaran Province, Iran. The study area is managed by the Natural Resources Faculty of Tehran University, Iran. This forest is located on the Northern slopes of the Alborz Mountains, about 7 km east of the little sea-port of Nowshar, Caspian Sea (N 36.6, E 51.8). It measures about 10,000 ha and extends from 0 to 2,200 m above sea level. This forest is divided into seven districts and the present research has been conducted in Patom, Namkhaneh, Gorazbon, Chelir and Baharbon districts in an area about 6,500 ha (**Figure 1**). Mean annual rainfall is 1,368 mm with maximum and minimum falls in October and June, respectively. Temperatures vary from a mean monthly minimum of 2.6 °C in January and February to a maximum of 29.2 °C in June and July, with an annual mean of 16.2 °C (Alavi et al., 2012). This region has a semi-moist climate with cold winters according to the De Martonne climagram. The study area's forests are mixed and uneven-aged, dominated by *Fagus orientalis* associated with *Carpinus betulus*, *Acer velutinum*, *Parrotia persica* C.A.Mey., *Sorbus torminalis* (L.) Crantz, *Quercus castaneifolia*, *Alnus subcordata* C.A.Mey., *Acer laetum* C.A.Mey., *Prunus avium* (L.) L., *Ulmus glabra* Huds. and *Tilia begoniifolia* Steven species. These forests are managed as close-to-nature with single tree harvesting methods.

2.2. Data collection

A stratified sampling method based on landform extracted from DEM used to locate 190 circular sample plots of 0.1 ha in beech forests in the study area. Plots were established in sites without evidence of anthropogenic disturbances including forest harvesting. All trees with DBH > 7 cm were measured for diameter at breast height (DBH) using caliper to the nearest millimeter and total height of all beech trees was using Vertex IV (Haglöf, Sweden). Dominant tree height is defined as the average height of five highest trees within the sample plot.

The summary of environmental variables are presented in **Table 1**. Slope (in %) of the plots in the study area was recorded by using a standard

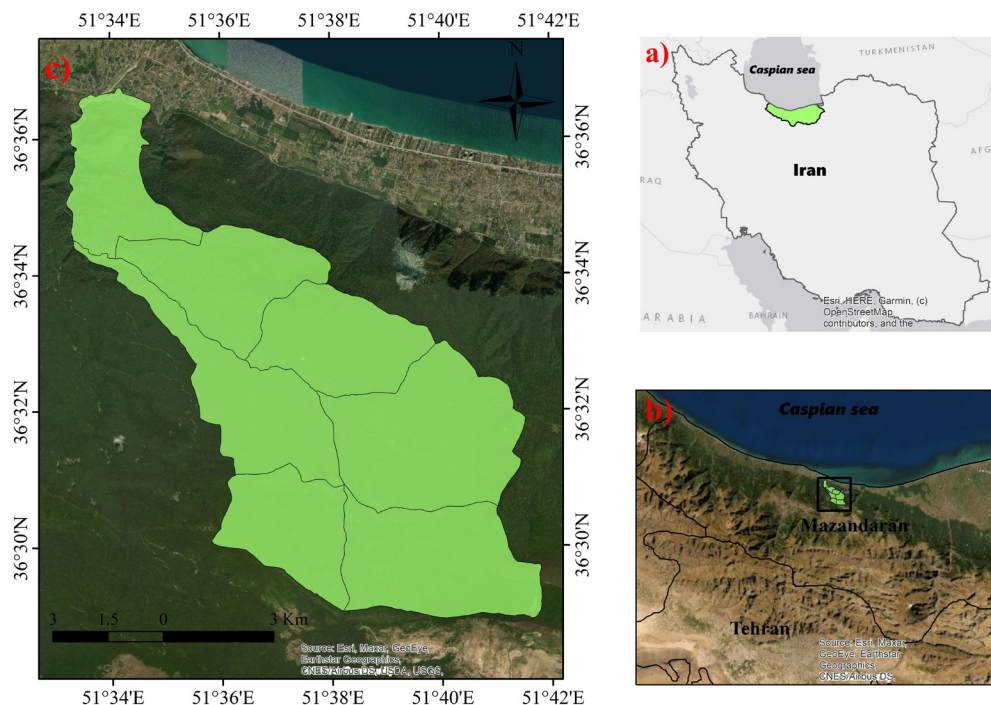


Figure 1. General location of study area — *Localisation générale de la zone d'étude.*

a. Iran; **b.** location of the study area in the north of Iran — *zone d'étude dans le nord de l'Iran*; **c.** study area in the Forest Research Station of Tehran University, Iran — *zone d'étude dans la station de recherche forestière de l'Université de Téhéran, Iran.*

Table 1. Summary of the site characteristics — *Résumé des caractéristiques du site.*

Variable	Minimum	Maximum	Mean	Standard Deviation
Dominant height (m)	36.27	47.60	42.71	2.82
Altitude (m a.s.l.)	470	1992	1223	299.44
Slope (%)	2	80	31	14
TRASP	0	1	0.51	0.36
Sand (%)	6.50	45.42	22.75	8.02
Clay (%)	25.51	65.51	39.71	6.59
Silt (%)	18.42	51.14	37.50	7.72
Bulk density ($\text{g}\cdot\text{cm}^{-3}$)	1.30	2.10	1.62	0.15
Nitrogen (%)	0.12	0.62	0.30	0.09
CN	9.77	20.07	12.83	1.82
pH	4.03	6.86	5.15	0.60
Lime (%)	0.63	2.29	1.15	0.24
Carbon (%)	1.57	9.36	3.86	1.45
OM (%)	2.70	16.10	6.65	2.49
Saturation (%)	43.93	55.58	49.44	1.91
P ($\text{mg}\cdot\text{kg}^{-1}$)	0.01	23.81	5.13	4.25
Ca ($\text{mg}\cdot\text{kg}^{-1}$)	24.00	608.00	166.18	124.49
Mg ($\text{mg}\cdot\text{kg}^{-1}$)	4.80	350.40	136.08	73.09
K ($\text{mg}\cdot\text{kg}^{-1}$)	151.70	754.27	345.37	117.69

clinometer. Aspect, as the azimuth measured from true north, was transformed to a topographic radiation index using the equation $TRASP = [1 - \cos((\pi/180)(\theta - 30))]/2$. This criterion assigns a value of zero to north-northeast facing slopes (typically the coolest and wettest orientation) and a value of one to the hotter and drier south-southwesterly slopes (Moisen & Frescino, 2002; Ahmadi et al., 2020b).

For quantifying nutrient availability, five topsoil samples (0-20 cm) were randomly taken within each plot. Soil samples were mixed and analyzed in the laboratory. Roots, shoots and pebbles were separated by hand and discarded and the air-dried soil samples were sieved. Soil variables including physical and chemical properties were determined by the following methods: bulk density (by clod method; Plaster, 2013), texture by Bouyoucos hydrometer method (Bouyoucos, 1962), pH in water (soil: water ratio 1:2.5), total organic C by Walkley and Black method (Allison, 1965), total N by Kjeldahl method (Bremner & Mulvaney, 1982), the available P by using the Olsen method (Homer & Pratt, 1961), available K, Ca and Mg by a flame atomic absorption spectrophotometer (AA500F, PG Instruments Ltd, China), the proportion of $CaCO_3$ (total lime) by the Calcimeter method (Allison & Moodie, 1965).

2.3. Statistical methods

The number of candidate variables included in the final model were firstly reduced by removing highly

correlated variables. Collinearity among environmental predictors were tested by hierarchical cluster analysis using squared Spearman correlations with the Hmisc package (Harrell Jr et al., 2018) in the R statistical software (R Core Team, 2018). In case of collinearity problems, one of any pair of predictors showing such problems should be discarded for modelling purposes (Draper & Smith, 1998; Dormann et al., 2013). The variables percentage saturation, percentage carbon and percentage organic matters were hence removed from the set of predictors. We then applied develop predictive model for variation of beech dominant height using boosted regression trees (BRT) (Figure 2).

The linear model was simplified using backward stepwise and bayesian information criterion (BIC), which considers both the goodness-of-fit and model complexity and penalizes model complexity more than the AIC (Burnham & Anderson, 2002) and has been argued to select for optimal explanation rather than optimal prediction. Classification and regression trees were created with the *rpart* package (Therneau & Atkinson, 2018) in the R statistical software. The resulted decision trees were pruned based on 10-fold cross-validation (McKenney & Pedlar, 2003).

Boosted regression trees were fitted using the *dismo* package (Hijmans et al., 2017) with a fixed tree complexity of 3 (according to recommendations by Elith et al. [2008] for small datasets) and a bag fraction of 0.75 and learning rate 0.001 and Gaussian response type. The BRT model was then simplified by reducing the number of independent variables (Elith

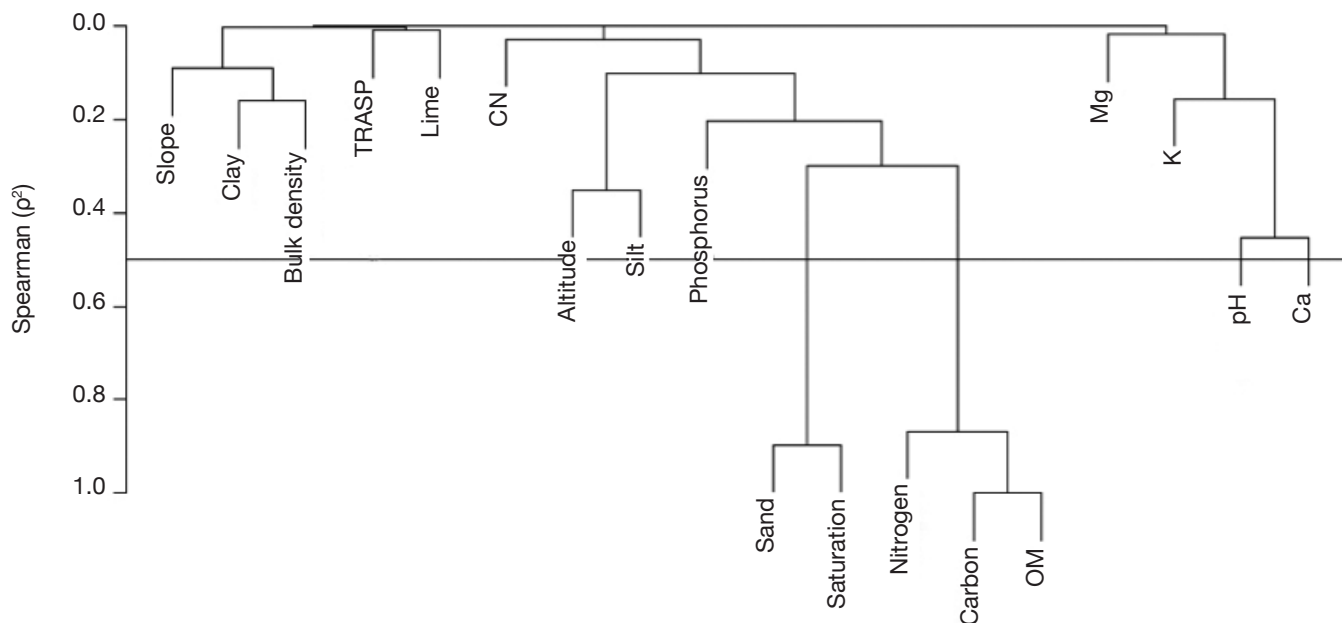


Figure 2. Hierarchical clustering using squared Spearman correlation (ρ^2) of environmental variables as similarities — *Regroupement hiérarchique utilisant la corrélation de Spearman au carré (ρ^2) des variables environnementales comme similarités.*

et al., 2008). Although the predictive importance of a variable can often be very insightful, most scholars are interested in how the variable is related to the outcome.

2.4. Evaluation of predictive performance of boosted regression trees

Since there is no independent data to evaluate the predictive performance of the models, 10-fold cross-validation method was used to measure their performances. In 10-fold cross-validation, the data are split into 10 random subsets of equal size. The modelling technique is then applied 10 times; each time one of the subsets is left out and the prediction accuracy is calculated by using that subset. The procedure was replicated 100 times. Predictive performance of derived models is quantified by calculating model evaluation measures on the predicted values for cross-validation. The performance of models was measured in terms of coefficient of determination (R^2) and the root mean squared error (RMSE). Finally, our ecological interpretation of optimal model relies on the assessment of the relative importance of the explanatory variables and their partial plots.

3. RESULTS

In the present study, we evaluated the dominant height of oriental beech, which is one of the most abundant species in the Hyrcanian forests of Iran, using a boosted regression tree model and edaphic and topographic variables. The results showed that 12 variables were influential, but phosphorus, percentage nitrogen, Mg and percentage sand had the highest effect on predicting the beech dominant height. Calcium and percentage clay were the least important variables in BRT. Large positive values indicate the variable is predictive, whereas zero or negative importance values identify variables not predictive (Ishwaran, 2007). Only few of the descriptors contributed noticeably to the estimation of beech dominant height, namely phosphorous, percentage of nitrogen, percentage of sand and Mg. The predictive performance in terms of coefficient of determination (R^2) values of BRT model is in **table 2**. This study

showed an improvement of ensemble method (BRT) with 58% in R^2 (**Figure 3**).

The partial dependency plots of the most important variables in BRT technique in beech dominant height variation showed that the higher dominant height is obtained when phosphorus and organic matter increase. Partial dependence plots representing the marginal effect of single variables using BRT model on estimates of beech dominant height are shown in **Figure 4**.

Table 2. Performance indices of boosted regression trees (BRT) — *Indices de performance des arbres à régression stimulée (BRT)*.

Statistical index	Boosted Regression Trees	
	Calibration	Validation
R^2	0.85	0.54 (0.15)
RMSE (m)	1.17	1.91 (0.31)

The numbers in the parentheses are standard deviation — *les nombres entre parenthèses sont les écarts-types*.

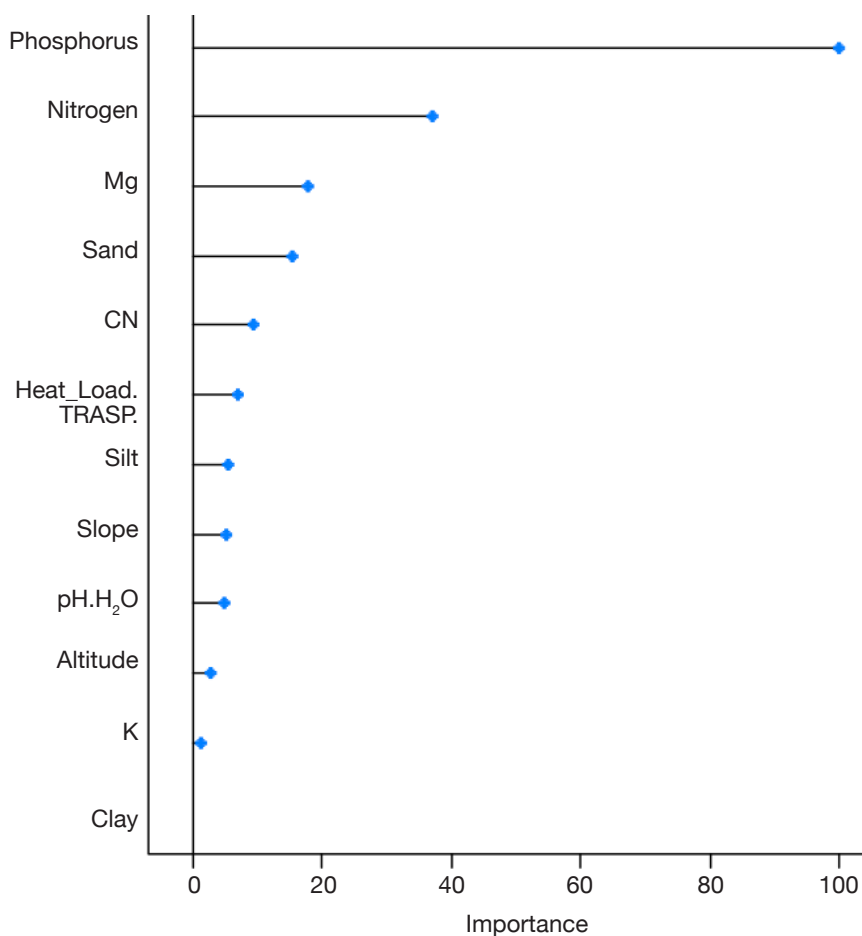


Figure 3. Variable importance plot generated by the BRT model — *Graphique d'importance variable généré par le modèle BRT*.

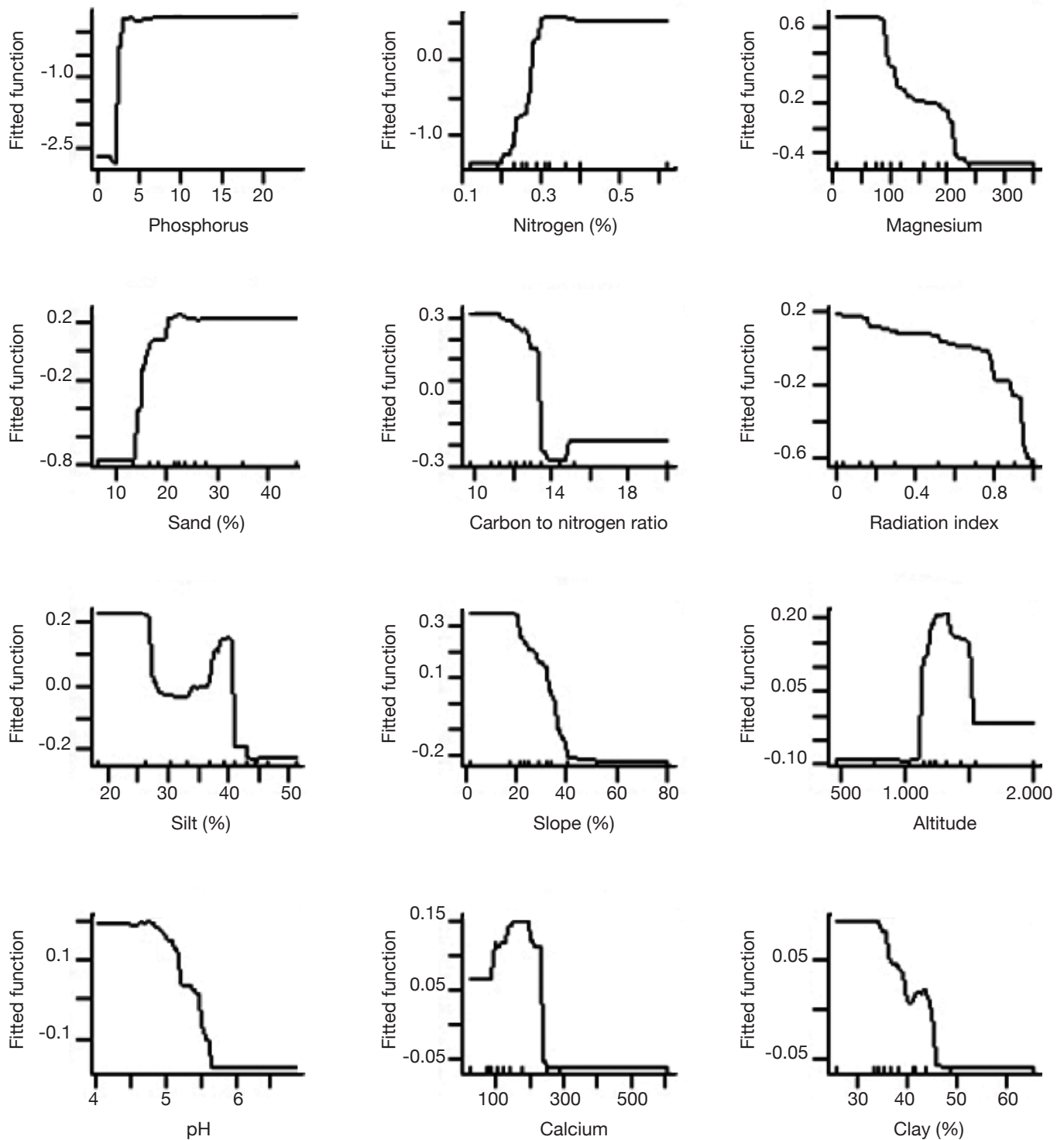


Figure 4. Partial plots of variables included in the BRT model in order of importance on estimates of dominant height of *Fagus orientalis* Lipsky — *Tracés partiels des variables incluses dans le modèle BRT par ordre d'importance sur les estimations de la hauteur dominante de Fagus orientalis Lipsky.*

4. DISCUSSION

In the present study, we evaluated the dominant height of oriental beech, which is one of the most abundant

species in the Hyrcanian forests of Iran, using a boosted regression tree model and edaphic and topographic variables. Like other studies (Lawler et al., 2006; Leathwick et al., 2006; Moisen et al., 2006; Prasad

et al., 2006; Benito Garzón et al., 2008; Lawler et al., 2009; Pittman et al., 2009; Aertsen et al., 2010; Knudby et al., 2010; Leclere et al., 2011; Vincenzi et al., 2011; Opiel et al., 2012; França & Cabral, 2015), in this research, application of ensemble technique (boosted regression trees) provides an effective methodology for predicting the beech dominant height in Hyrcanian forests. Elith et al. (2008) highlighted some main advantages of BRT approach including strong predictive performance, reliable identification of relevant variables and interactions. The increasing growth of BRT application in ecological studies is a witness for its efficiency (Elith et al., 2006; Leathwick et al., 2006; Elith et al., 2008; Pittman et al., 2009; Aertsen et al., 2010; Froeschke & Froeschke, 2011; Aertsen et al., 2012; Kint et al., 2012). Differences in model performances may be attributed to their inherent properties. There are a certain number of assumptions in linear regression including normality, homoscedasticity, independence of variables and model linearity that are rarely met by these models (Zuur et al., 2009), but tree-based approaches may overcome these difficulties. Contrary to linear models, in machine learning techniques an algorithm is used in order to learn the relationship between the response and explanatory variables, then dominant patterns in data are found *via* the inputs and response observations, not *a priori*, and finally the model structure is developed as a direct function of that particular dataset (Miller & Franklin, 2002; Elith et al., 2008; França & Cabral, 2015). Although classification and regression trees are inherently simple and interpretable, they have a major drawback; a small change in the data can often prompt to extensive changes in the form of the fitted tree. Therefore, it is somewhat difficult to interpret these trees reliably. This is the downside of such a simple model structure, thus ensemble methods could be used for solving these problems (Simpson & Birks, 2012). The results of this study also showed ensemble methods are preferred and superior techniques for predicting response variables and their superiority is attributed to the incorporation of nonlinearity and interaction effects, because they are important features and lead to lower prediction errors.

The application of boosted regression trees model for prediction purposes is particular useful when there are complex interactions between predictors and response variable (in our case the dominant height of *F. orientalis*) and the possibility of highly correlated predictor variables. Despite the fact that there were some systematic differences in performance among methods, the BRT method shows relatively similar and consistent patterns in predicting *F. orientalis* dominant height. Thus, we conclude that the dominant height of *F. orientalis* in the Hyrcanian forests can be successfully predicted using BRT nonlinear modelling techniques. Therefore, we use an ensemble approaches for further investigation.

An important aspect in ecological modelling involves the evaluation and interpretation of the results. Phosphorus and nitrogen were the most important variables affecting the dominant height of beech tree species. Soil nitrogen and phosphorus are the most common macronutrients which limit the growth of plants under natural conditions (Liu et al., 2014). Phosphorus is of particular importance in accelerating the root growth, cell division, and growth of meristem tissues, its limitation is associated with a sharp decline in tree growth. As a result, phosphorus deficiency will slow down or stop the growth of above- and underground parts of the forest trees. The results of this study showed that with increasing nitrogen content, the dominant height of beech tree species also increases. On the other hand, beech tree has a decreasing behavior relative to the C/N variable. The ratio of carbon to nitrogen is one of the important indices of mineralization and soil fertility. Increasing nitrogen contents and decreasing carbon to nitrogen ratios increase the activity of soil microorganisms and accelerate the litter decomposition (Habibi Kaseb, 1992; Shabani et al., 2012); as a result, the growth of beech tree species increases. In the analysis of the response curve of beech tree, Alavi et al. (2017) concluded that NPK and C/N variables are effective indices on tree growth (Alavi et al., 2017).

The performance of species along elevation gradient is governed by a series of interacting biological, climatic, and historical factors (Colwell & Lees, 2000). Further, elevation represents a complex gradient along which many environmental variables change simultaneously (Austin et al., 1996). Beech tree has the best performance at altitudes from 1,200 to 1,700 m a.s.l. which is consistent with Marvi Mohajer (1976). It seems these altitude ranges have optimum humidity conditions and high productivity which resulted from optimal combination resource availability (Rahbek, 1995; Rosenzweig, 1995). Having demands for cool climates, beech tree avoids lower altitudes, since they are warmer and drier. The major decline in beech performance at higher altitudes could be due in part to ecophysiological constrains, such as reduced growing season, low temperature and low ecosystem productivity in high elevation (Körner, 1998).

It is evident that beech trees have better performance on north-northeast directions (typically the coolest and wettest orientations). In the Northern Hemisphere, south-facing slopes may receive as much as six times more solar radiation than north-facing slopes. Thus, the south-facing slopes have a more xeric environment, that is, warmer, drier and a more variable microclimate, than the mesic north-facing slopes (Nevo, 1997; Nevo, 2001). Beech tree has the best performance in gentle slopes. The slope degree is an essential feature of topography in relation to runoff and soil erosion. The

runoff and soil loss intensity can vary with different slope gradients. Since the slope gradient is the main factor for controlling soil erosion, with increasing slope gradient, the amount of soil loss increases significantly (Koulouri & Giourga, 2007). Several authors have also confirmed the exponential influence of slope gradient on soil loss (Lal, 1976; Roose, 1977).

Soil texture as an important soil characteristic determines the rates of water intake, water storage in the soil, the amount of aeration (vital to root growth) and influences the soil fertility. The effects of the textural properties of soils are frequently reflected in the composition and growth rate of forest vegetation (Sharma et al., 2010). The result of this study indicated percentage sand was a more important factor than clay in affecting the beech dominant height. The partial dependence plots for these two variables are also interesting. The partial plots for these variables show that the beech tree species has the best performances in low clay and medium sand contents, respectively which is consistent with the results of Tabari et al. (2005). While soils high in clay are difficult to manage because of their great strength and sticky nature, an intermediate amount of clay in a soil improves its capacity to hold water and plant nutrient ions. A balanced combination of sand, silt and clay composition makes loamy soil, which is the best one for the plant growth (Pidwirny, 2004) and support the luxuriant vegetation.

This research also shows that the high pH values are a limiting factor on the dominant height of beech tree. pH of the soil directly or indirectly affects the growth of the tree species. The most important role of soil pH is controlling the solubility of nutrients in the soil. In other words, the ability to absorb nutrients is highly dependent on soil pH. Nutrients have different solubility at different pH values. Usually, by increasing the pH values, the solubility of the essential nutrients for the plant growth is reduced, and the deficiency of nutrients such as phosphorus, iron, zinc and manganese in the plant can be observed (Salardini, 2011; Alavi et al., 2017).

5. CONCLUSIONS

Boosted regression tree model was used for evaluation of beech dominant height in Hyrcanian forests. Our assessment of the mentioned techniques based on two measures of accuracy showed that boosted regression tree model has high performance in reducing the prediction error of dominant height variable in these forests. The high flexibility of this model is attributed to its ability for incorporating the nonlinear and interaction effects. The boosted regression trees method with low RMSE has a good accuracy for predicting beech dominant height in the Hyrcanian forests. By

reviewing the literatures and the response curves resulted from the different modelling approaches, the response curves of BRT also have better ecological interpretability and rationality. Boosted regression tree technique indicated that phosphorus, percentage nitrogen, magnesium and percentage sand were among the most important variables in predicting the dominant height of beech tree species in the study area. There are other modelling techniques, for example, Multiple Adaptive Regression Splines and Artificial Neural Networks that should also be considered for future evaluations of model performance, since their applications in ecological modelling studies have been gradually increasing. Nonetheless, in these techniques, the scientists have not much control over model fitting, and it is difficult to assess the relative importance of individual explanatory variable. Overall, machine learning techniques help researchers gain more insights, both in terms of ecological relationships and species spatial distributions specially in terms of forest modelling and yield tables. They are now considered as a valuable tool for improving ecological management and biodiversity conservation.

Bibliography

- Aertsen W. et al., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol. Modell.*, **221**, 1119-1130, doi: 10.1016/j.ecolmodel.2010.01.007
- Aertsen W. et al., 2012. Predicting forest site productivity in temperate lowland from forest floor, soil and litterfall characteristics using boosted regression trees. *Plant Soil*, **354**, 157-172, doi: 10.1007/s11104-011-1052-z
- Ahmadi K., Alavi S.J., Tabari Kouchaksaraei M. & Aertsen W., 2013. Non-linear height-diameter models for oriental beech (*Fagus orientalis* Lipsky) in the Hyrcanian forests, Iran. *Biotechnol. Agron. Soc. Environ.*, **17**, 431-440.
- Ahmadi K. et al., 2020a. The potential impact of future climate on the distribution of European yew (*Taxus baccata* L.) in the Hyrcanian forest region (Iran). *Int. J. Biometeorol.*, **64**, 1451-1462, doi.org/10.1007/s00484-020-01922-z
- Ahmadi K. et al., 2020b. Patterns of density and structure of natural populations of *Taxus baccata* in the Hyrcanian forests of Iran. *Nord. J. Bot.*, **38**(3), doi.org/10.1111/njb.02598.
- Alavi S.J. et al., 2012. Extracting ecological optimum and amplitude of *Fagus orientalis* along environmental gradients in Kheyroud forest, Nowshahr. *J. Nat. Environ.*, **64**, 399-415.
- Alavi S.J., Nouri Z. & Zahedi Amiri G., 2017. The response curve of beech tree (*Fagus Orientalis* Lipsky) in relation to environmental variables using generalized additive model in Khayroud forest, Nowshahr.

- J. Wood For. Sci. Technol.*, **24**, 29-59, doi: 10.22069/jwfst.2017.11630.1617
- Alavi S.J. et al., 2019. The response of English yew (*Taxus baccata* L.) to climate change in the Caspian Hyrcanian mixed forest ecoregion. *Reg. Environ. Change*, **19**(5), 1495-1506, doi.org/10.1007/s10113-019-01483-x
- Allison L.E., 1965. Organic carbon. In: Norman A.G., ed. *Methods of soil analysis. Part 2: chemical and microbiological properties*. Madison, WI, USA: American Society of Agronomy, 1367-1378.
- Allison L.E. & Moodie C.D., 1965. Carbonate. In: Norman A.G., ed. *Methods of soil analysis. Part 2: chemical and microbiological properties*. Madison, WI, USA: American Society of Agronomy, 1379-1396.
- Assmann E., 1970. *The principles of forest yield study: studies in the organic production, structure, increment, and yield of forest stands*. Oxford, UK: Pergamon Press.
- Austin M.P., Pausas J.G. & Nicholls A.O., 1996. Patterns of tree species richness in relation to environment in south-eastern New South Wales, Australia. *Aust. J. Ecol.*, **21**, 154-164, doi.org/10.1111/j.1442-9993.1996.tb00596.x
- Batho A. & Garcia O., 2006. De Perthuis and the origins of site index: a historical note. *FBMIS*, **1**, 1-10.
- Benito Garzón M., Sánchez De Dios R. & Sainz Ollero H., 2008. Effects of climate change on the distribution of Iberian tree species. *Appl. Veg. Sci.*, **11**, 169-178, doi: 10.3170/2008-7-18348
- Berger S. & Walther G.R., 2006. Distribution of evergreen broad-leaved woody species in Insubria in relation to bedrock and precipitation. *Bot. Helv.*, **116**, 65-77, doi: 10.1007/s00035-006-0749-8
- Bouyoucos G.J., 1962. Hydrometer method improved for making particle size analyses of soils. *Agron. J.*, **54**, 464-465.
- Bremner J.M. & Mulvaney C.S., 1982. Nitrogen—Total 1. In: Page A.L., Miller R.H. & Keeney D.R., eds. *Methods of soil analysis. Part 2. Chemical and microbiological properties*. Madison, WI, USA: American Society of Agronomy, Soil Science Society of America, 595-624.
- Burnham K.P. & Anderson D.R., 2002. *Model selection and multimodel inference*. 2nd ed. New York, NY, USA: Springer-Verlag.
- Carmean W.H., 1975. Forest site quality evaluation in the United States. In: Brady N.C., ed. *Advances in agronomy*. Vol. 27. New York, NY, USA: Academic Press, 209-269.
- Clutter J.L. et al., 1983. *Timber management: a quantitative approach*. John Wiley & Sons, Inc.
- Colwell R.K. & Lees D.C., 2000. The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol. Evol.*, **15**(2), 70-76, doi.org/10.1016/s0169-5347(99)01767-x
- Coomes D.A. & Allen R.B., 2007. Effects of size, competition and altitude on tree growth. *J. Ecol.*, **95**, 1084-1097, doi: 10.1111/j.1365-2745.2007.01280.x
- Curtis R.O. & Reukema D.L., 1970. Crown development and site estimates in a Douglas-fir plantation spacing test. *For. Sci.*, **16**, 287-301.
- Dormann C.F. et al., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27-46, doi.org/10.1111/j.1600-0587.2012.07348.x
- Draper N.R. & Smith H., 1998. *Applied regression analysis*. Wiley.
- Drew C.A., Wiersma Y.F. & Huettmann F., 2011. *Predictive species and habitat modeling in landscape ecology: concepts and applications*. New York, NY, USA: Springer.
- Elith J.L. et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.
- Elith J.L., Leathwick J.R. & Hastie T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.*, **77**, 802-813.
- Elith J.L. & Graham C.H., 2009. Do they? How do they? WHY do they differ? on finding reasons for differing performances of species distribution models. *Ecography*, **32**, 66-77, doi: 10.1111/j.1600-0587.2008.05505.x
- Forrester D.I., 2014. The spatial and temporal dynamics of species interactions in mixed-species forests: from pattern to process. *For. Ecol. Manage.*, **312**, 282-292, doi.org/10.1016/j.foreco.2013.10.003
- França S. & Cabral H.N., 2015. Predicting fish species richness in estuaries: which modelling technique to use? *Environ. Modell. Software*, **66**, 17-26, doi.org/10.1016/j.envsoft.2014.12.010
- Froeschke J.T. & Froeschke B.F., 2011. Spatio-temporal predictive model based on environmental factors for juvenile spotted seatrout in Texas estuaries using boosted regression trees. *Fish. Res.*, **111**, 131-138, doi.org/10.1016/j.fishres.2011.07.008
- Gorgi Bahri Y. & Sagheb-Talebi K., 1992. The effects of late spring frosts on beech forests of Golband (Noshar). *Pazhouhesh Sazandegi J.*, **15**, 18-21.
- Habibi Kaseb H., 1974. Investigation of influence of soil texture on beech trees' growth rate in Iran. *J. Iran Nat. Resour.*, **31**, 61-69.
- Habibi Kaseb H., 1992. *Fundamentals of forest soil science*. Tehran: Tehran University Press.
- Hagglund B., 1981. Evaluation of forest site productivity. *For. Abstr.*, **42**(11), 515-527.
- Harrell Jr F.E. et al., 2018. *Hmisc: Harrell Miscellaneous. R package version 4.1-1*, <https://CRAN.R-project.org/package=Hmisc>, (19/11/2020).
- Hegel T.M., Cushman S.A., Evans J. & Huettmann F., 2010. Current state of the art for statistical modelling of species distributions. In: *Spatial complexity, informatics, and wildlife conservation*. Springer, 273-311.
- Herrera B., Campos J.J., Finegan B. & Alvarado A., 1999. Factors affecting site productivity of a Costa Rican secondary rain forest in relation to *Vochysia ferruginea*,

- a commercially valuable canopy tree species. *For. Ecol. Manage.*, **118**, 73-81, doi.org/10.1016/s0378-1127(98)00489-7
- Hijmans R.J., Phillips S., Leathwick J. & Elith J., 2017. *Dismo: Species Distribution Modeling. R package version 1.1-4*, <https://CRAN.R-project.org/package=dismo>, (19/11/2020).
- Hogg B.W. & Nester M.R., 1991. Effect of stocking rate on predominant height of young Caribbean pine plantations in coastal Queensland. *Aust. For.*, **54**, 134-138, doi.org/10.1080/00049158.1991.10674569
- Homer D.C. & Pratt P.F., 1961. *Methods of analysis for soils, plants and waters*. Berkeley, CA, USA: University of California, 150-196.
- Ishwaran H., 2007. Variable importance in binary regression trees and forests. *Electron. J. Stat.*, **1**, 519-537, doi.org/10.1214/07-ejs039
- Kint V. et al., 2012. Radial growth change of temperate tree species in response to altered regional climate and air quality in the period 1901-2008. *Clim. Change*, **115**, 343-363, doi.org/10.1007/s10584-012-0465-x.
- Knudby A., Brenning A. & LeDrew E., 2010. New approaches to modelling fish-habitat relationships. *Ecol. Modell.*, **221**, 503-511, doi.org/10.1016/j.ecolmodel.2009.11.008
- Körner C., 1998. A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, **115**, 445-459, doi: 10.1007/s004420050540
- Koulouri M. & Giourga C., 2007. Land abandonment and slope gradient as key factors of soil erosion in Mediterranean terraced lands. *Catena*, **69**, 274-281, doi.org/10.1016/j.catena.2006.07.001
- La Roi G.H., Strong W.L. & Pluth D.J., 1988. Understorey plant community classifications as predictors of forest site quality for lodgepole pine and white spruce in west-central Alberta. *Can. J. For. Res.*, **18**, 875-887, doi.org/10.1139/x88-135
- Lal R., 1976. Soil erosion problems on an alfisol in Western Nigeria and their control. *Geoderma*, **25**, 215-230, doi.org/10.1016/0016-7061(81)90037-9
- Lawler J.J., White D., Neilson R.P. & Blaustein A.R., 2006. Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biol.*, **12**, 1568-1584, doi.org/10.1111/j.1365-2486.2006.01191.x
- Lawler J.J. et al., 2009. Projected climate-induced faunal change in the Western Hemisphere RID C-7190-2009 RID E-4643-2011. *Ecology*, **90**, 588-597, doi: 10.1890/08-0823.1
- Leathwick J.R. et al., 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Mar. Ecol. Prog. Ser.*, **321**, 267-281, doi.org/10.3354/meps321267
- Leclere J., Oberdorff T., Belliard J. & Leprieur F., 2011. A comparison of modeling techniques to predict juvenile 0+ fish species occurrences in a large river system. *Ecol. Inf.*, **6**, 276-285, doi.org/10.1016/j.ecoinf.2011.05.001
- Liu X. et al., 2014. Available phosphorus in forest soil increases with soil nitrogen but not total phosphorus: evidence from subtropical forests and a pot experiment. *PLoS One*, **9**, 1-8, doi: 10.1371/journal.pone.0088070
- Marvi Mohajer M., 1976. Some qualitative characteristics of Iranian beech forests. *Iran J. Nat. Resour.*, **34**, 77-96.
- Mataji A., Zahedi Amiri G. & Asri Y., 2009. Vegetation analysis based on plant associations and soil properties in natural forests. *Iran J. For. Poplar Res.*, **17**, 85-98.
- McKenney D.W. & Pedlar J.H., 2003. Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada. *For. Ecol. Manage.*, **175**, 497-507, doi.org/10.1016/s0378-1127(02)00186-x
- Miller J. & Franklin J., 2002. Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecol. Modell.*, **157**, 227-247, doi: 10.1016/S0304-3800(02)00196-5
- Moisen G.G. & Frescino T.S., 2002. Comparing five modelling techniques for predicting forest characteristics. *Ecol. Modell.*, **157**, 209-225, doi.org/10.1016/s0304-3800(02)00197-7
- Moisen G.G. et al., 2006. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecol. Modell.*, **199**, 176-187, doi: 10.1016/j.ecolmodel.2006.05.021
- Nambiar E.K.S., Ranger J., Tiarks A. & Toma T., 2004. Site management and productivity in tropical plantation forests. In: *Proceedings of workshops in Congo, July 2001 and China, February 2003*. Jakarta: CIFOR.
- Nevo E., 1997. Evolution in action across phylogeny caused by microclimatic stresses at "Evolution Canyon". *Theor. Popul. Biol.*, **3**, 231-243, doi.org/10.1006/tpbi.1997.1330
- Nevo E., 2001. Evolution of genome-phenome diversity under environmental stress. *Proc. Natl. Acad. Sci.*, **98**, 6233-6240, doi.org/10.1073/pnas.101109298
- Nunes L., Patrício M., Tomé J. & Tomé M., 2011. Modeling dominant height growth of maritime pine in Portugal using GADA methodology with parameters depending on soil and climate variables. *Ann. For. Sci.*, **68**, 311-323, doi.org/10.1007/s13595-011-0036-8
- Oliver C.D. & Larson B.C., 1996. *Forest stand dynamics*. New York, NY, USA: Wiley.
- Oppel S. et al., 2012. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biol. Conserv.*, **156**, 94-104, doi: 10.1016/j.biocon.2011.11.013
- Palahí M. et al., 2004. Site index model for *Pinus sylvestris* in north-east Spain. *For. Ecol. Manage.*, **187**, 35-47, doi.org/10.1016/s0378-1127(03)00312-8
- Pidwirny M., 2004. *Introduction to environmental issues-The soil resource*. Kelowna, Canada: Okanagan College.
- Pittman S.J., Costa B.M. & Battista T.A., 2009. Using lidar bathymetry and boosted regression trees to predict the

- diversity and abundance of fish and corals. *J. Coastal Res.*, **10053**, 27-38, doi.org/10.2112/si53-004.1
- Plaster E., 2013. *Soil science and management*. Farmington Hills, MI, USA: Cengage Learning.
- Prasad A.M., Iverson L.R. & Liaw A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, **9**, 181-199.
- Pretzsch H., 2009. *Forest dynamics, growth, and yield*. Springer.
- Pretzsch H. et al., 2015. Growth and yield of mixed *versus* pure stands of Scots pine (*Pinus sylvestris* L.) and European beech (*Fagus sylvatica* L.) analysed along a productivity gradient through Europe. *Eur. J. For. Res.*, **134**(5), 927-947, doi.org/10.1007/s10342-015-0900-4
- R Core Team, 2018. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Rahbek C., 1995. The elevational gradient of species richness: a uniform pattern? *Ecography*, **18**, 200-205, doi.org/10.1111/j.1600-0587.1995.tb00341.x
- Roose E.J., 1977. Use of the universal soil loss equation to predict erosion in West Africa. In: *Soil erosion: prediction and control*. Ankeny, IA, USA: Soil Conservation Society of America, 60-74.
- Rosenzweig M.L., 1995. *Species diversity in space and time*. Cambridge, UK: Cambridge University Press.
- Sagheb-Talebi K., 1996. Quantitative and qualitative characteristics of beech sapling (*Fagus sylvatica* L.) growing under various site conditions with emphasis on light. *Beih. Schweiz. Z. Forstwesen*, **78**, 219.
- Sagheb Talebi K. et al., 2014. *Forests of Iran: a treasure from the past, a hope for the future*. Dordrecht, The Netherlands: Springer
- Salardini A.A., 2011. *Soil fertility*. 9th ed. Tehran: Tehran University Press.
- Salehi A., Zarrinkafsh M., Zahedi Amiri G. & Marvie Mohadjer M.R., 2005. A study of soil physical and chemical properties in relation to tree ecological groups in Nam-Khaneh district of Kheirood-kenar forest. *Iran J. Nat. Resour.*, **58**, 567-577.
- Segurado P. & Araujo M.B., 2004. An evaluation of methods for modelling species distributions. *J. Biogeogr.*, **31**, 1555-1568.
- Seynave I. et al., 2005. *Picea abies* site index prediction by environmental factors and understorey vegetation: a two-scale approach based on survey databases. *Can. J. For. Res.*, **35**, 1669-1678, doi.org/10.1139/x05-088
- Shabani S., Akbarinia M., Jalali G. & Aliarab A., 2012. Relationship between soil characteristics and beech regeneration density in canopy gaps with different sizes. *J. Wood For. Sci. Technol.*, **18**, 63-78.
- Sharma C.M., Gairola S., Ghildiyal S.K. & Suyal S., 2010. Physical properties of soils in relation to forest composition in moist temperate valley slopes of the Central Western Himalaya. *J. For. Environ. Sci.*, **26**, 117-129.
- Simpson G.L. & Birks H.J.B., 2012. Statistical learning in palaeolimnology. In: *Tracking environmental change using lake sediments*. Springer, 249-327.
- Tabari M. et al., 2005. Response of oriental beech (*Fagus orientalis* Lipsky) seedlings to canopy gap size. *Forestry*, **78**, 443-450.
- Therneau T.M. & Atkinson B., 2018. *rpart: recursive partitioning and regression trees. R package version 4.1-13*, <https://CRAN.R-project.org/package=rpart>, (20/11/2020).
- Ung C.-H. et al., 2001. Biophysical site indices for shade tolerant and intolerant boreal species. *For. Sci.*, **47**, 83-95.
- Vanclay J.K. & Skovsgaard J.P., 1997. Evaluating forest growth models. *Ecol. Modell.*, **98**, 1-12, doi.org/10.1016/s0304-3800(96)01932-1
- Vincenzi S. et al., 2011. Application of a random forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. *Ecol. Modell.*, **222**, 1471-1478, doi: 10.1016/j.ecolmodel.2011.02.007
- von Gadow K. & Hui G., 1999. *Modelling forest development*. Dordrecht, The Netherlands: Kluwer Academic Publishers, doi.org/10.1007/978-94-011-4816-0
- Zuur A.F. et al., 2009. Zero-truncated and zero-inflated models for count data. In: Zuur A.F. et al. *Mixed effects models and extensions in ecology with R*. Springer, 261-293.

(93 ref.)