NUSANTARA BIOSCIENCE
Vol. 6, No. 1, pp. 57-62
May 2014

ISSN: 2087-3948
E-ISSN: 2087-3956
DOI: 10.13057/nusbiosci/n060110

Using and comparing two nonparametric methods (CART and RF) and SPOT-HRG satellite data to predictive tree diversity distribution

SIAVASH KALBI♥, ASGHAR FALLAH, SEYED MOHAMMAD HOJJATI
Department of Forestry, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University, P.O.Box: 578, Sari, Mazandaran, Iran. Tel./Fax. +98 151 3822715, ♥email: s.kalbi63@gmail.com

Manuscript received: 13 September 2013. Revision accepted: 19 December 2013.

Abstract. *Kalbi S, Fallah A, Hojjati SM. 2014. Using and comparing two nonparametric methods (CART and RF) and SPOT-HRG satellite data to predictive tree diversity distribution. Nusantara Bioscience 6: 57-62.* The prediction of spatial distributions of tree species by means of survey data has recently been used for conservation planning. Numerous methods have been developed for building species habitat suitability models. The present study was carried out to find the possible proper relationships between tree species diversity indices and SPOT-HRG reflectance values in Hyrcanian forests, North of Iran. Two different modeling techniques, Classification and Regression Trees (CART) and Random Forest (RF), was fitted to the data in order to find the most successful model. Simpson, Shannon diversity and the reciprocal of Simpson indices were used for estimating tree diversity. After collecting terrestrial information on trees in the 100 samples, the tree diversity indices were calculated in each plot. RF with determinate coefficient and RMSE from 56.3 to 63.9 and RMSE from 0.15 to 0.84 has better results than CART algorithms with determinate coefficient 42.3 to 63.3 and RMSE from 0.188 to 0.88. Overall the results showed that the SPOT-HRG satellite data and nonparametric regression could be useful for estimating tree diversity in Hyrcanian forests, North of Iran.

Keywords: Tree diversity, random forest, classification, regression tree

INTRODUCTION

Forest management and farming, along with natural disturbances like wildfire, storms, and floods have caused widespread land use changes and landscape fragmentation (Ramezani and Holm 2010). These processes may have resulted in biodiversity losses, environmental functions and ecological processes which generate and maintain soil, convert solar energy into plant tissue, regulate climatic parameters and provide multiple forest products (Isik et al. 1997).

Hyrcanian forests are the individual natural ecosystem that enjoys the highest plants and animals diversity comparing with other ecosystems in Iran. They are being destroyed by degradation and conversion to other land uses. Under pressure to make informed management decisions rapidly, conservation practitioners must increasingly rely on predictive models to provide them with information on species distributions (Loiselle et al. 2003; Saatchi et al. 2000). The most accurate ways to collect biographical data on species distributions are intensive ground surveys or inventories of species in the field. However, remote sensing offers a cost-efficient means for deriving complete spatial coverage of environmental information for large areas in a consistent manner. Recent studies have indicated that remote sensing may be able to provide useful information on biodiversity (Hernandez-Stefanoni and Dupuy 2007; Mohammadi and Shataee 2010).

Dogan and Dogan (2006) tested the predictability of several biodiversity indices such as Shannon's diversity, Simpson and richness using spatial predictor variables. These variables are topography, geology, soil, climate, normalized difference vegetation index (NDVI), and land cover. They offered three models for Shannon's diversity, Simpson, and richness indices. Mohammadi and Shataee (2010) investigated the possibility of estimation of tree diversity using Landsat ETM+ data in the Hyrcanian forests, North of Iran.

The models for tree species richness and the reciprocal of the Simpson index were obtained with reasonable accuracy. Bawa et al. (2002) reported that there is a statistically significant relation between the species diversity and NDVI of IRS 1C imagery and NDVI may be used to characterize areas of high and low tree species richness in tropical forests where biodiversity losses are significant. The regression analysis approach has broadly been applied in ecological surveys (Lehmann et al. 2002). Linear regression is a commonly used statistical technique for modeling biodiversity because of its easy use and direct interpretability (Curt et al. 2001; Seynave et al. 2005). The development of advanced nonparametric and machine learning techniques are opening up plenty of opportunities for modeling biodiversity with greater accuracy and may be better fitted to address the mentioned problems compared with linear regression (Aertsen et al. 2010).

Generalized linear models (McCullagh and Nelder 1989) and generalized additive models (Hastie and Tibshirani 1990) using presence-absence survey data have been taken much more attention recently. Moisen and Frescino (2002) investigated the performance of non-parametric techniques as CART, generalized additive models (GAM) and artificial neural networks (ANN) compared to parametric techniques for the prediction of several species independent forest characteristics in the interior Western United States. MARS and ANN worked best to simulated data, but less suitable for real data, in which case a LM approach often provided comparable results. Shataee et al. (2012) compared three nonparametric models include k-nearest neighbor (k-NN), support vector machine regression (SVR) and tree regression based on random forest (RF), for estimation forest structure characteristic using ASTER satellite data. Overall, they showed RF produced has better results than SVR and k-NN.

The aim of this study was to compare and evaluate two statistical non-parametric (CART, RF) for modeling tree species diversity. It is also intended to investigate the relationship between the properties of satellite image spectral bands and tree species diversity; in order to predict the distribution of plant species diversity using new nonparametric methods over the study area.

MATERIALS AND METHODS

Study area

The study area is located in the Hyrcanian forests, the district 1 of Darabkola's forests, Sari, North of Iran (Figure 1). The boundary of this area is located at 36° 28´-36° 33´ N and 53° 16´-53° 20´ 30´´ W. The Darabkola's forestry plan, with about 2600 ha area, consists of natural temperate and uneven-aged stands. The main tree species are *Quercus Castaneifolia* (chestnut-leaved oak), *Carpinus betulus* (hornbeam), *Acer velutinum* (velvet maple), *Alnus subcordata* (Caucasian alder), *Tilia begonifolia* (linden tree), *Parrotia persica* (Persian ironwood), *Ulmus glabra* (*elm*), *Acer platanoides* (Norway maple), *Diospyros lotus* (date pulm), *Zelkova carpinifolia* (Siberian elm), *Fagus orientalis* (Oriental beech) and *Acer cappadocicum* (coliseum maple).

Field data

Species richness and diversity indices are dependent on the size of the sample plot. Phytosociological data were collected based on a systematic sampling method from 5th June to 15th July 2010. The size and the number of quadrats were determined using the species-area curve (Misra 1968). Choosing the sample size, the number of sampling units to select and measure, is a key part of planning a survey. 100 sample plots (quadrate shape) were placed using a stratified random sampling design 450 × 500 m. The sample plot size was 60 × 60 m and characteristics of trees with DBH more than 7.5 cm were measured. The geographical center of each plot was registered using a GPS Oregon 550.

Diversity indices

A large number of diversity indices can be used to characterize tree size diversity within a stand (Smith et al. 1992; Varga et al. 2005; Ozdemir et al. 2008). Two common approaches for measuring alpha diversity are species richness and evenness/ heterogeneity (Ojo and Ola-Adams 1996). Species richness simply refers to the number of species in the community while evenness/ heterogeneity refers to the distribution of individuals among the species. In this study, species richness wasn't considered. For the measurement of evenness/ heterogeneity, Simpson, Shannon diversity indices and the reciprocal of the Simpson index were computed for each of the sites. The more uncertainty one has about the species of an individual, the higher the diversity of the community. The proportion of a species has been based on a variety of variables to represent frequency, including the number of individuals (Niese and Strong 1992; Condit et al. 1996), basal area (Harrington and Edwards 1995; LeMay et al. 1997), stems per ha (McMinn 1992; Harrington and Edwards 1995); and biomass (Swindel et al. 1984). In this study, the proportion of basal area species is used in this index.

Satellite data

The SPOT-HRG data were orthorectified using 23 GCPs and DEM. The total root mean square errors (RMSE) were obtained about 0.67 for visible and near infra bands and 0.5 for the middle infra band. Pixel size of middle band was resized to 10m using nearest-neighbor resampling method. The geometric precision of the images was also verified using road vector layer and unused collected GPS control points and proved the accuracy of geometric rectification. In order to atmospheric correction, the COST general method was used for decreasing of effect of attenuation and scattering in the visible and near-infrared bands. The DNs of images were converted to radiance and then to reflectance values. The reflectance of the haze number was determined through the histogram evaluation.

Image processing techniques

After geometric rectification and atmospheric corrections, the most used vegetation indices were generated for probabilistic capabilities of these indices in regression modeling (Table 1). Also used of mean and variance each four bands and principal component analyses for all bands and three bands.

Spectral signature extraction of the plot

The pixel sizes of all used images were aggregated to 60 meters according to size of field plots (60×60) and their spectral values were averaged. Then the averaged values of main and processed images of SPOT-HRG were extracted in place of each plot.

Statistical models

Classification and regression tree

Classification and regression tree, a statistical procedure introduced by Breiman et al. (1984), is primarily used as a classification tool, where the objective is to classify an object into two or more populations (Lee et al. 2006). Regression trees, while effective at incorporating disparate data types, non-normal distributions, and non-linear relationships, do not allow for tree optimization, and accuracy may suffer in the presence of outliers and non-balanced datasets (Lawrence et al. 2004; Barrett et al. 2010). Regression trees are hierarchical structures, where the internal nodes contain tests on the input attributes. Each

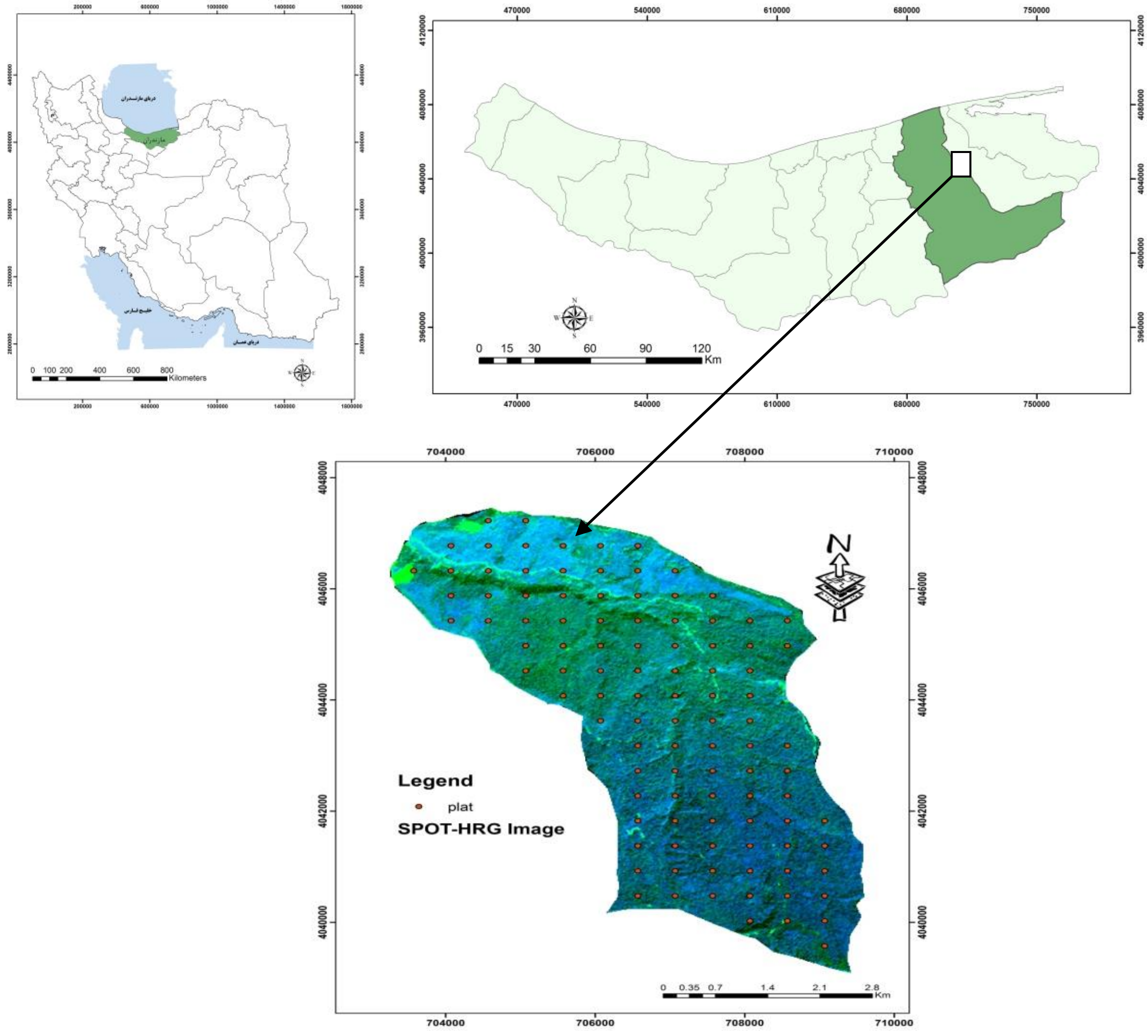

Figure 1. Location of the study area in the Mazandaran Province (a) and allocation of sample plots (b) in the study area.

Table 1. Most importance Spectral vegetation indices examined in this study.

Index	**Equation**	**Reference**
Normalized Ratio (NR)	Red-NIR	Mohammadi et al. (2010
Simple Ratio (SR)	$\frac{NIR}{Red}$	Birth and McVey (1968)
Difference Vegetation Index (DVI)	NIR-Red	Tucker (1979)
Modified Soil Adjusted Vegetation Index (MSAVI2)	NIR + 0.5 $\sqrt{(NIR+0.5)^2-2(NIR-Red)}$	Qi et al. (1994)
Normalized difference vegetation index(NDVI)	(NIR-RED) /(NIR+RED)	Rouse et al. (1973)
Short wave infrared to visible ratio (SVR)	SWIR/[(RED+GRN)/2]	Wolter et al. (2008)
Moisture stress index (MSI)	SWIR/NIR	Rock et al. (1986)
Reduced Simple Ratio (RSR)	$\frac{NIR}{Red}(1-\frac{SWIR-SWIR_{min}}{SWIR_{max}-SWIR_{min}}$	Brown et al. (2000)
Renormalized Difference Vegetation Index (RDVI)	$\sqrt{NDVI * DVI}$	Roujean and Breon (1995)
Normalized difference water index (NDWI)	NIR-SWIR/NIR+SWIR	Gao (1996)
Global environmental monitoring index (GEMI)	$\eta(1\text{-}0.25\eta)\ \frac{Red-0.125}{1-Red}$ $\eta=\frac{2(NIR^2-Red^2)+1.5*NIR+0.5*Red}{NIR+Red+0.5}$	Pinty and Verstraete (1992)

Note: SWIRmin and SWIRmax are the minimum and maximum reflectance values observed in the corresponding pixels in field plots.

branch of an internal test corresponds to an outcome of the test, and the prediction for the value of the target attribute is stored in a leaf. Each leaf of a regression tree contains a constant value as a prediction for the target variable (Kocev et al. 2009). The resulting prediction of the tree is taken from the leaf at the end of the path.

Random forest

Random forest is a novel ensemble classifier; it uses a similar but improved method of bootstrap as bagging (Zhang et al. 2009). It uses the strategy of a random selection of a subset of predictors to grow each tree, where each tree is grown on a bootstrap sample of the training set. This number, m, is used to split the nodes and is much smaller than the total number of variables available for analysis (Breiman 2001). In training, the random forest algorithm creates multiple CART-like trees (Breiman et al. 1984), each trained on a bootstrapped sample of the original training data, and searches only across a randomly selected subset of the input variables to determine a split (for each node). Random forests for regression are formed by growing trees depending on a random vector such that the tree predictor takes on numerical values. However, when constructing a tree, random forest searches for only a random subset of the input features (bands) at each splitting node and the tree is allowed to grow fully without pruning (Chan and Paelinckx 2008). The random forests predictor is formed by taking the average over a number of the trees specified by the user (Lariviere and van den Poel 2005).

The number of predictors used to find the best split at each node is a randomly chosen subset of the total number of predictors (Prasad et al. 2006). One of the main parameters which should be determined in RF is a k predictor (independent variables) in each node for predicting dependent values (response). The response of each tree depends on a set of predictor values, which is independently chosen with replacement and with the same distribution of all trees in the forest, which is a subset of the predictor values of the original data set. The simplest choosing way k is calculation of root square of total independent variables ($k \leq \sqrt{m}$, m is the number of input variables).

Model evaluation and performance assessment

Data were randomly split into two data sets, 70% of the data for modeling and 30% for testing. For each model that was tested, four statistics are reported; these are the squared coefficient of determination (R^2) (Pearson, 1896) and adjusted coefficient of determination (adjusted R^2). The validity of performances was examined using regression diagnostics metrics, i.e., root means square error (RMSE), relative RMSE, bias, and relative bias, and using the independent and unused 30 samples. In addition to, some common graphical diagnostic tools (McRoberts 2009) were used to illustrate the quality of performances.

$$RMSe = \sqrt{\sum_{i=1}^{m}\left(est_i - obs_i\right)} \Big/ m$$

$$RMSe\% = \frac{\sqrt{\sum_{i=1}^{m}\left(est_i - obs_i\right)} \Big/ m}{\sum_{i=1}^{m}\left(obs_i\right)/m} * 100$$

$$\text{Bias} = \sum_{i=1}^{m}\left(est_i - obs_i\right) \Big/ m$$

$$\text{Bias}\% = \frac{\sum_{i=1}^{m}\left(est_i - obs_i\right) \Big/ m}{\sum_{i=1}^{m}\left(obs_i\right)/m} * 100$$

Where *est* is estimated values from implementation of algorithms in *m* validation samples, *obs* is observation values and m is the number of validation samples.

RESULTS AND DISCUSSION

Descriptive statistics of indices

Simpson, Shannon’s diversity indices and the reciprocal of Simpson index descriptive statistics for the proportion of basal area species is provided in Table 2. The value of Simpson, Shannon’s indices, and the reciprocal of the Simpson index ranged from 0.105 to 0.86, 0.12 to 2.89 and 1.105 to 4.02, respectively. It indicates a wide range of tree species diversity in the study area (Table 2).

All models were critically investigated for confounding factors and checked for all basic assumptions (Table 3). The number of predictor variables entering the models is ranging from three to five, while the predictor variables selected by each technique are not identical.

The measures of performance are summarized for each model in Table 4. The best model performance was realized with highest R^2, adjusted R^2 and lowest RMSE, $RMSE_r$ Bias and $Bias_r$ values. In the total cases, the best goodness-of-fit, i.e., lowest values for RMSE and Bias and the highest adjusted R^2, was obtained from the RF models.

Discussion

Hyrcanian forests comprise a diverse vegetation cover in the north of Iran and are increasingly degraded and converted to other land uses (Mohammadi et al. 2008). In this study, assessing utility SPOT5-HRG satellite images data and two different regression techniques for modeling tree diversity in Hyrcanian forest. These results are similar to those obtained in other studies (Foody and Cutler 2007; Hernandez-Stefanoni and Dupuy 2007; Mohammadi and Shataee 2010) where researchers demonstrated that satellite data can identify broad patterns of tree species diversity.

In this study, the infrared index was determined to be very important to estimate the species diversity of trees and this wavelength was used due to the high reflection in the infrared spectrum (Bawa et al. 2002). Correlation coefficients between species diversity and range of values in different bands corresponding positive and reflects the increasing range of different wavelengths, tree and shrub diversity also increased. The dense masses, in which there is more species diversity, reflect a large amount of the infrared spectral range, but in sparse masses, where the species diversity is low, reflected infrared is decreased because the red wavelength enters into the forest and spreads, which influences its absorption and ultimately reduces its reflection.

Increasing the diversity and density of the canopy tree increases the rate of reflection in this range. With adjusted R^2 values for the best models ranging from 42.8 to 64.5, the results look satisfactory compared to other studies (Mohammadi and Shataee 2010; Gillespie et al. 2009; Dogan and Dogan 2006). In recent years the RF algorithm has gained popularity as an effective regression method in the remote sensing domain (Shataee et al. 2012). The results of the present study confirm that the RF algorithm is a robust and accurate method for the modeling of satellite data.

The robustness of the RF algorithm can be explained by the ability of the modeling and classification algorithm to exploit the noise in the dataset to create a more diverse classifier (Breiman 2001). In all cases, CART model has shown a poor result for modeling biodiversity. The results of this study have been consistent with some of previous studies (Moisen and Frescino 2002; Aertsen et al. 2010) that reported that CART models performed worst than nonparametric regressions. This may be owed to the fact that CART models produce a stepwise response function. In case of a rather smooth relationship between predictors and response, this can lead to low performance.

Table 2. Descriptive statistics of model and validation samples for indices

	Training samples					Evaluation samples				
Index	N	Mean	Min	Max	S.D	N	Mean	Min	Max	S.D
Simpson	70	0.474	0.105	0.86	0.19	30	0.514	0.122	0.752	0.148
Shannon's diversity	70	1.22	0.12	2.66	0.547	30	1.38	0.154	2.89	0.519
Reciprocal	70	2.07	1.10	4.02	0.708	30	2.27	1.22	4.02	0.735

Table 3. Overview of the predictor variables selected by the tree biodiversity models developed with two techniques.

Index	Modeling technique	Variable(s) selected by the model
Simpson	CART	SVR, MSI, NR, Mean NIR
	RF	Mean SWIR, Variance NIR, Mean Red, NR
Shannon's	CART	Mean SWIR, MSI, Mean NIR, Variance Green
diversity	RF	RDVI, Mean SWIR, pci3, Mean NIR, Variance NIR
Reciprocal	CART	PCI4, MSI, Mean NIR, SVR, Variance swir
	RF	SVR, Mean swir, Mean NIR

Table 4. Performance indices of all SI-models for the three tree species and two modeling techniques. Best model performance for every evaluation measure, is highlighted in bold.

Index	Modelling technique	R^2	R^2_{adj}	RMSE	$RMSE_r$	Bias	$Bias_r$
Simpson	CART	43.6	42.8	0.188	36.9	0.038	7.4
	RF	56.8	56.2	0.15	30.1	0.038	7.4
Shannon-Wiener	CART	53.7	53	0.68	53.7	-0.23	-18.6
	RF	61.2	60.6	0.56	43.9	-0.19	-14.9
Reciprocal	CART	63.8	63.3	0.88	41.6	-0.15	-7.4
	RF	64.5	63.9	0.84	40.06	-0.22	-10.5

CONCLUSION

Tree diversity is one of the important properties that determine the vegetation needed to field measurements, the limits of its own and must determine which tools and methods to use auxiliary data such as satellite images data is used. Overall the results show that the SPOT-HRG data could be useful for estimating tree diversity and therefore can be employed to assess and monitor the status of tree diversity in the northeastern forests of Iran.

REFERENCES

Aertsen W, Kint V, Van Orshoven J, Ozkan K, Muys B. 2010. Comparison and ranking of different modeling techniques for prediction of site index in Mediterranean mountain forests. Ecol Model 221: 1119-1130.

Barrett K, Kasischke ES, McGuire AD, Turetsky MR, Kane ES. 2010. Modeling fire severity in black spruce stands in the Alaskan boreal forest using spectral and non-spectral geospatial data. Remote Sens Environ 114: 1494-1503.

Bawa K, Rose J, Ganeshaiah KN, Barve N, Kiran MC, Umashaanker R. 2002. Assessing biodiversity from space: an example from the Western Ghats, India. Conserv Ecol 6 (2): 7.

Birth GS, McVey GR. 1968. Measuring the color of growing turf with a reflectance spectrophotometer. Agron J 60: 640-643.

Breiman L, Friedman JH, Olshen RA. Stone CJ. 1984. Classification and Regression Trees. Wadsworth International Group, Belmont.

Breiman L. 2001. Random forests. Machine Learning 45: 5-32.

Brown L, Chen JM, Leblanc SG, Cihlar J. 2000. A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: an image and model analysis. Remote Sens Environ 71: 16-25.

Chan JCW, Paelinckx D. 2008. Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens Environ 112: 2999-3011.

Condit R, Hubbell SP, LaFrankie JV, Sukumar R, Manokaran N, Foster RB, Ashton PS. 1996. Species-area and species-individual relationships for tropical trees: a comparison of three 50-ha plots. J Ecol 84: 549-62.

Curt T, Bouchaud M, Agrech G. 2001. Predicting site index of Douglas-fir plantations from ecological variables in the Massif Central area of France. For Ecol Manag 149: 61-74.

Dogan HM, Dogan M. 2006. A new approach to diversity indices — modeling and mapping plant biodiversity of Nallihan (A3-Ankara/Turkey) forest ecosystem in frame of geographic information systems. Biodiv Conserv 15: 855-878.

Foody GM, Cutler MEJ. 2007. Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. Ecol Mod 19 (5): 37-42.

Gao BC. 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens Environ 58 (3): 257-266.

Gillespie TW, Foody GM, Rocchini D, Giorgi AP, Saatchi S. 2008.. Measuring and modelling biodiversity from space. Progr Phys Geogr 32 (2): 203-221.

Harrington TB, Edwards MB. 1995. Structure of mixed pine and hardwood stands 12 years after various methods and intensities of site preparation in the Georgia Piedmont. Can J For Res 26: 1490-1500.

Hastie TJ, Tibshirani RJ. 1990. Generalized additive models. Chapman and Hall, New York.

Hernandez-Stefanoni JL, Dupuy J. 2007. Mapping species density of trees, shrubs and vines in a tropical forest, using field measurements, satellite multispectral imagery and spatial interpolation. Biodiv Conserv 16: 3817-3833.

Isik K, Yaltikik F, Akesen A. 1997. The interrelationship of forests, biological diversity and the maintenance of natural resources. Unasylva FAO 48: 190-191.

Kocev D, Džeroski S, White MD, Newell GR, Griffoens P. 2009. Using single- and multi-target regression trees and ensembles to model a compound index of vegetation condition. Ecol Mod 220: 1159-1168.

Larivière B, Van den Poel D. 2005. Predicting customer retention and profitability by using random forests and regression forests techniques. Expert Syst Appl 29 (2): 472-484

Lawrence R, Bunn A, Powell S, Zambon M. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sens Environ 90: 331-336.

Lee TS, Chiu CC, Chou YC, Lu CJ. 2006. Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. Comput Stat Data Anal 50: 1113-1130.

Lehmann A, Overton JMcC, Austin MP. 2002. Regression models for spatial prediction: their role for biodiversity and conservation. Biodiv Conserv 11: 2085-2092.

LeMay V, Morgan D, Söderberg U. 1997. Biodiversity measures: What could be obtained from forest inventory data. Paper presented at the Canadian forest inventory committee annual meeting. Fort Simpson, NWT.

Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH. 2003. Avoiding pitfalls of using species distribution models in conservation planning. Conserv Biol. 17: 1591-1600.

McCullagh P, Nelder JA. 1989. Generalized Linear Models. Chapman and Hall, New York.

McMinn JW. 1992. Diversity of woody species 10 years after four harvesting treatments in the oak-pine type. Can J For Res 22: 1179-1183.

McRoberts RE. 2009. Diagnostic tools for nearest neighbors techniques when used with satellite imagery. Remote Sens Environ 113: 489-499.

Misra R. 1968. Ecology Workbook. Oxford and IBH Publishing Co., New Delhi, India.

Mohammadi J, Shataee S. 2010. Possibility investigation of tree diversity mapping using Landsat ETM+ data in the Hyrcanian forests of Iran. Remote Sens Environ 114: 1504-1512.

Mohammadi J, Shataee Sh, Habashi H, Amiri M. 2008. The effect of shelterwood logging on the diversity of tree species in the Loveh's Forest, Gorgan. Iranian J For Popular Res 16 (2): 241-250.

Moisen GG, Frescino TS. 2002. Comparing five modeling techniques for predicting forest characteristics. Ecol Model 157: 209-225.

Niese JN, Strong TF. 1992. Economic and tree diversity trade-offs in managed northern hardwoods. Can J For Res 22: 1807-1813.

Ojo LO, Ola-Adams BA. 1996. Measurement of tree diversity in the Nigerian rainforest. Biodiv Conserv 5: 1253-1270.

Ozdemir I, Norton DA, Ozkan UY, Mert A, Senturk O. 2008. Estimation of tree size diversity using object-oriented texture analysis and aster imagery. Sensors 8: 4709-4724.

Pinty B, Verstraete MM. 1992. GEMI: A non-linear index to monitor global vegetation from satellites. Vegetatio 101: 15-20.

Prasad AM, Iverson IR, Liaw A. 2006. Newer classification and regression tree techniques: Bagging and random forest for ecological prediction. Ecosystem 9: 181-199

Qi J, Chehbouni A, Huete AR, Kerr YH, Sorooshian S. 1994. A modified soil adjusted vegetation index. Remote Sens Environ 48: 119-126.

Ramezani H, Holm S. 2010. A distance dependent contagion functions for vectorbased data. Environ Monit Assess (available online). doi:10.1007/s10651-009-0123-2.

Rock BN, Vogelmann JE, Williams DL, Vogelmann AF, Hoshizaki T. 1986. Remote detection of forest damage. BioScience 36 (7): 439-445.

Roujean JL, Breon FM. 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens Environ 51: 375-384.

Rouse JW, Haas RH, Schell JA, Deering DW. 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: Third Earth Resources Technology Satellite-1 Symposium, 10-14 December 1973, Washington, DC (Washington, DC: NASA). 309-317.

Saatchi S, Buermann W, Hans ter Steege HT, Mori S, Smith TB. 2000. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. Remote Sens Environ 112 (18): 2000-2017.

Seynave I, Gegout JC, Herve JC, Dhote JF, Drapier J, Bruno E, Dume G. 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.

Shataee S, Kalbi S, Fallah A, Pelz D. 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. Intl J Remote Sens 33 (19): 6254-6280.

Smith WR, Farrar RM, Murphy PA, Yeiser JL, Meldahl RS, Kush JS. 1992. Crown area and basal area relationships for open-grown southern pines for modeling competition and growth. Can J For Res 22: 341-347.

Swindel BF, Conde LF, Smith JE. 1984. Species diversity: concept, measurement, and response to clearcutting and site-preparation. For Ecol Manag 8: 11-22.

Tucker CJ. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8: 127-150.

Varga P, Chen HYH, Klinka K. 2005. Tree-size diversity between single- and mixed-species stands in three forest types in western Canada. Can J For Res 35: 593-601.

Wolter PT, Townsend PA, Kingdon CC, Sturtevant BR. 2008. Remote sensing of the distribution and abundance of host species for spruce budworm in northern Minnesota and Ontario. Remote Sens Environ 112: 3971-3982.