Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/forestsystems
http://dx.doi.org/10.5424/fs/2013223-03874

Forest Systems 2013 22(3): 484-496
ISSN: 2171-5068
eISSN: 2171-9845

Forest attributes estimation using aerial laser scanner and TM data

Sh. Shataee

Forestry Department. Gorgan University of Agricultural Sciences and Natural Resources. 386 Gorgan, Iran

Abstract

Aim of study: The aim of this study was performance of four non-parametric algorithms including the *k*-NN, SVR, RF and ANN to estimate forest volume and basal area attributes using combination of Aerial Laser Scanner and Landsat-TM data.

Area of study: Data in small part of a mixed managed forest in the Waldkirch region, Germany.

Material and methods: The volume/ha and basal area/ha in the 411 circular plots were estimated based on DBH and height of trees using volume functions of study area. The low density ALS raw data as first and last pulses were prepared and automatically classified into vegetation and ground returns to generate two fine resolution digital terrain and surface models after noise removing. Plot-based height and density metrics were extracted from ALS data and used both separated and combined with orthorectified and processed TM bands. The algorithms implemented with different options including *k*-NN with different distance measures, SVR with the best regularized parameters for four kernel types, RF with regularized decision tree parameters and ANN with different types of networks. The algorithm performances were validated using computing absolute and percentage RMSe and bias on unused test samples.

Main results: Results showed that among four methods, SVR using the RBF kernel could better estimate volume/ha with lower RMSe and bias (156.02 m^3 ha^{-1} and 0.48, respectively) compared to others. In basal area/ha, *k*-NN could generate results with similar RMSe (11.79 m^3 ha^{-1}) but unbiased (0.03) compared to SVR with RMSe of 11.55 m^3 ha^{-1} but slightly biased (–1.04).

Research highlights: Results exposed that combining Lidar with TM data could improve estimations compared to using only Lidar or TM data.

Key words: forest attributes estimation; ALS; TM; non-parametric algorithms.

Introduction

Optical satellite imagery usually presents two-dimensional spectral information and reflectance responses of a canopy cover's surface. However, LiDAR data from an airborne laser scanner (ALS) provides semi three-dimensional data set relating to canopy cover. From the early 1980s so far, application of ALS-LiDAR data for forest attribute estimation has been tested in different forest areas in the world and in the stand, plot and tree levels. Since LiDAR data have tested for retrieval of forest structure biophysical data such as canopy or tree height (Naesset, 2004; Andersen *et al.,* 2003; Maltamo *et al.*, 2006a; Naesset, 1997; Chen and Hay, 2011), basal area/ha, volume/ha (Nilsson, 1996; Nelson *et al.,* 1997; Holmgren, 2004) and biomass (Lefsky *et al.,* 2001; Packalén and Maltamo, 2007; Latifi *et al.,* 2010). There are many studies that have proved the possibility of making accurate estimates of forest attributes using only Lidar data (Wulder, 1998; Lefsky *et al*., 1999; Naesset, 2002; Lim *et al.*, 2003; Holmgren, 2004; Maltamo *et al.,* 2006a,b; Dalponte *et al.,* 2009; Breidenbach *et al.,* 2010). However, many studies, have demonstrated that ALS data were not suitable to estimate forest attributes and classifications when used alone (Packalén, 2009; Latifi *et al.,* 2010; Chen and Hay, 2011). Therefore, it can be assumed that a combination of optical data and LiDAR data would be a successful combination for estimating forest attributes (Packalén, 2009; Straub *et al.,* 2010).

Forest attribute estimations using integration of LiDAR and optical remote sensing data can be traditionally done using regression approaches. The regression approaches can be generally divided into two groups: parametric and non-parametric approaches. For forest variables estimations using integration of LiDAR and optical data, some researches has been applied the parametric approaches (Chen and Hay, 2011; Popescu

* Corresponding author: shataee@yahoo.com
Received: 18-12-12. Accepted: 19-02-13.

and Wynne, 2004; He *et al.,* 2011; Lu *et al.,* 2012), and some others applied non-parametric approaches (Chen *et al.,* 2012; McInerney *et al.,* 2010; Stojanova *et al.,* 2010). The multiple linear or non-linear regressions are typical parametric approaches which define relationships between forest dependents variables such as tree volume and image spectral metrics and height and intensity metrics derived from LiDAR data. Although the parametric approaches have been widely used and easy to interpret, these empirical approaches often do not have the ability to characterize forest complexity, especially at fine spatial scales (Chen *et al.,* 2010) or in the mountainous and natural mixed deciduous forests. In addition, the main advantages of non-parametric methods are that they do not rely on any probability distribution. In parametric models, often unbiased estimators for the expected value and the parameters exist. However, parametric regression models may easily fail with respect to the practitioner's demand for simultaneous estimations of several response variables. This is often due to strong restrictions in model validity and as a large number of response variables may result in a small number of degrees of freedom (Breidenbach *et al.,* 2010).

Recently, the use of non-parametric algorithms with different theories have been explored for estimation of forest attributes due to their advantages algorithms such as flexibility and ability to describe non-linear dependencies (Franco-López *et al*., 2001; Makela and Pekkarinen, 2004; Hyvönen, 2007; Sironen *et al*., 2010). The non-parametric machine learning techniques have demonstrated superior performances over classic regression analysis for estimating forest attributes (Latifi *et al.,* 2010; Chen and Hay, 2011). Machine-learning algorithms consist of groups of data mining and non-parametric-based methods with different bases. These algorithms can be grouped into distance-based algorithms such as; *k* nearest neighbor (*k*-NN) and its variants; tree-structured or decision tree algorithms such as random forest (RF); statistical learning theory based algorithms such as support vector machine regression (SVR) and artificial intelligent based algorithms such as artificial neural network (ANN). Each of these algorithms has advantages and disadvantages in terms of their inherent capabilities and performances. So, regarding to their different theories, investigation on capability of performances of these algorithms can be a useful way to find the best algorithm to quantify forest attributes using different based remote sensing data.

Definition of used algorithms

The k nearest neighbour (k-NN) algorithm is the simplest machine-learning algorithm used for both classification and regression. For regression, *k*-NN simply assigns a property value for an object to be taken as the average of the values of *k*-NNs. The *k*-NN is commonly used to estimate forest attributes using various remote-sensing data (Franco-López *et al*., 2001; Makela and Pekkarinen, 2004; McRobert *et al*., 2007). Implementing *k*-NN algorithms takes three parameters including the value of *k* or number of NN, type of distance measured and weights for nearest neighbors. For continuous response variables, the *k*-NN prediction for the *i*th target set element is:

$$\bar{y} = \frac{1}{K} \sum_{i=1}^{k} w_i \, y_i \qquad [1]$$

where, $\bar{y}$ is the prediction (outcome) of the query point, $\{y_i, i = 1, 2,\ldots, k\}$ is the set of response variable observations for the *k* reference set elements that are nearest to the *i*th target set element in a feature space with respect to distance metrics, and w_i is the weight assigned to the *i*th nearest neighbour that can be obtained with equation [2] (McRoberts, 2009):

$$w_i = \sum_{i=1}^{k} w_{ij} \qquad [2]$$

The random forest (RF) means forest of decision trees and is an extension of classification and regression tree (CART) method (Breiman, 2001) to reduce instability of a decision tree. The RF can be used for regression-type problems to predict forest continuous dependent variables (Eskelson *et al.,* 2009; Breidenbach *et al.*, 2010) as well as classification problems to predict categorical dependent variables (Walton, 2008). In regression problems, RF is an arbitrary number (ensemble) of simple trees (subset from independent variables) used to vote responses that are combined (averaged) to obtain an estimate of dependent variables. Data and variables can be randomly sampled by iteratively bagging bootstrap sampling to generate a forest of regression tree. The predictions of the RF are taken to be the average of the predictions of the trees. Mean of errors are obtained trough difference between observed and tree responses (predictions).

The support vector regression (SVR) algorithm is regression branch of support vector machine (SVM) techniques based on statistical learning theory developed by Vapnik (1995). The SVR generally focus on

boundaries between classes and then map the input as space created by independent variables using non-linear transformation according to kernel functions. The most commonly used kernels are radial basis function (RBF), sigmoid and polynomial. In SVR, the algorithm is trying to find a hyper plane that can accurately predict the distribution of information. In this mapped high-dimensional space, an optimal linear separator or hyper plane maximizes the margin between classes. The SVR has been used to estimate forest biophysical variables in a few studies such as estimating forest canopy cover (Wang and Brenner, 2009); tree level estimation of biomass (Dalponte *et al.*, 2009) and for leaf area index (Durbha *et al.*, 2007).

The artificial neural network (ANN) algorithm is a powerful non-linear modeling technique that is free of traditional assumptions; well suited for complex nonlinear relationships and perfect for exploratory analyses where the goal is to establish if any relationship exists among a set of variables. It attempts to model nonlinear functions with large numbers of variables. The neural network designers therefore traditionally run training algorithms a number of times with a given "neural" network design. In ANN, the type of neural network, the number of input variables and hidden units and settings of various control parameters in training algorithms can all affect the final performance of a network. Radial basis function (RBF), generalized regression neural networks (GRNN) and multilayer perceptron (MLP) are the most commonly used types of neural networks. In recent years, several studies have used ANN for retrieving biophysical variables (Jin and Liu, 1997; Atzberger, 2004; Niska *et al.,* 2010).

Regarding to above mentioned notes, investigation on capability of performances of four different theories-based non-parametric algorithms using combination of ALS and TM data, is main objective of this study.

Materials and methods

Field data

This study was done in a small area according to LiDAR covered area on municipal and state forest enterprises around Waldkirch, 13 km north-east of Freiburg, Germany. The study area is a part of municipal forests of Waldkirch with 1,267 ha and 411 plots (Fig. 1). It was a managed forest, semi flat, and comprising hardwood and softwood mixed forest with dominant species as follows; European beech (*Fagus sylvatica* L.), Norway spruce [*Picea abies* (L.) Karst.], European silver fir (*Abies alba* Mill.), Douglas fir [*Pseudotsuga menziesii* (Mirb.) Franco.] and Sycamore maple (*Acer pseudoplatanus* L.) (Breidenbach *et al.*, 2010).

In this study, species-specific volumes of tree species for 411 sample plots were selected from the database of the Waldkirch forest that was compiled in 2002 by the Baden-Wurttemberg forest service. The spreadsheet also contained the species-specific aggregated total basal area and the average diameter at breast height (DBH). Forest characteristics were recorded in all plots consisting of four concentric circular plots with radii of 2, 3, 6 and 12 m; plots were positioned on the intersection of a 100 × 200 m sample grid. Individual tree volume was calculated using DBH, height parameters and volume functions of the Baden-Wurttemberg state forest service and only DBH in the case of individual basal area used to determine basal area. Four radius-specific weighting factors were then applied to volumes and basal areas to scale them to per-hectare values. These plot-level attributes were derived by aggregating the weighted individual tree volumes and basal areas.

ALS Data

Low resolution ALS data (return density of 0.5-1 m^{-2}) was acquired from the State Surveying Office of Baden-Wurttemberg using an Optech ALTM 1225 laser scanner from a flight height of approximately 900 m above ground level (AGL) between the years 2000 and 2005. The swath width was 500-600 m with 50% overlap and scan angle of ± 20°. The vertical accuracy and horizontal accuracy were determined about 0.15 and 0.45 m respectively. The ALS raw data as first and last pulse data were automatically classified into vegetation and ground returns. Correct classification of last pulse data was manually checked and corrected by experts from the state surveying office with the highest possible accuracy (Breidenbach *et al.*, 2010). The one meter resolution digital terrain model (DTM) and digital surface model (DSM) computed after removing the noise hits by TreesVis software (Weinacker *et al.,* 2004). The plot-level height percentiles (*i.e.* 1st, 5th, 10th, 20th, 25th, 30th, 40th, 50th, 60th, 70th, 75th, 80th, 90th, 95th, and 99th percentiles), density statistics metrics (such as variance, skewness, mean and etc.) and various ratios of returns as percentages and number of

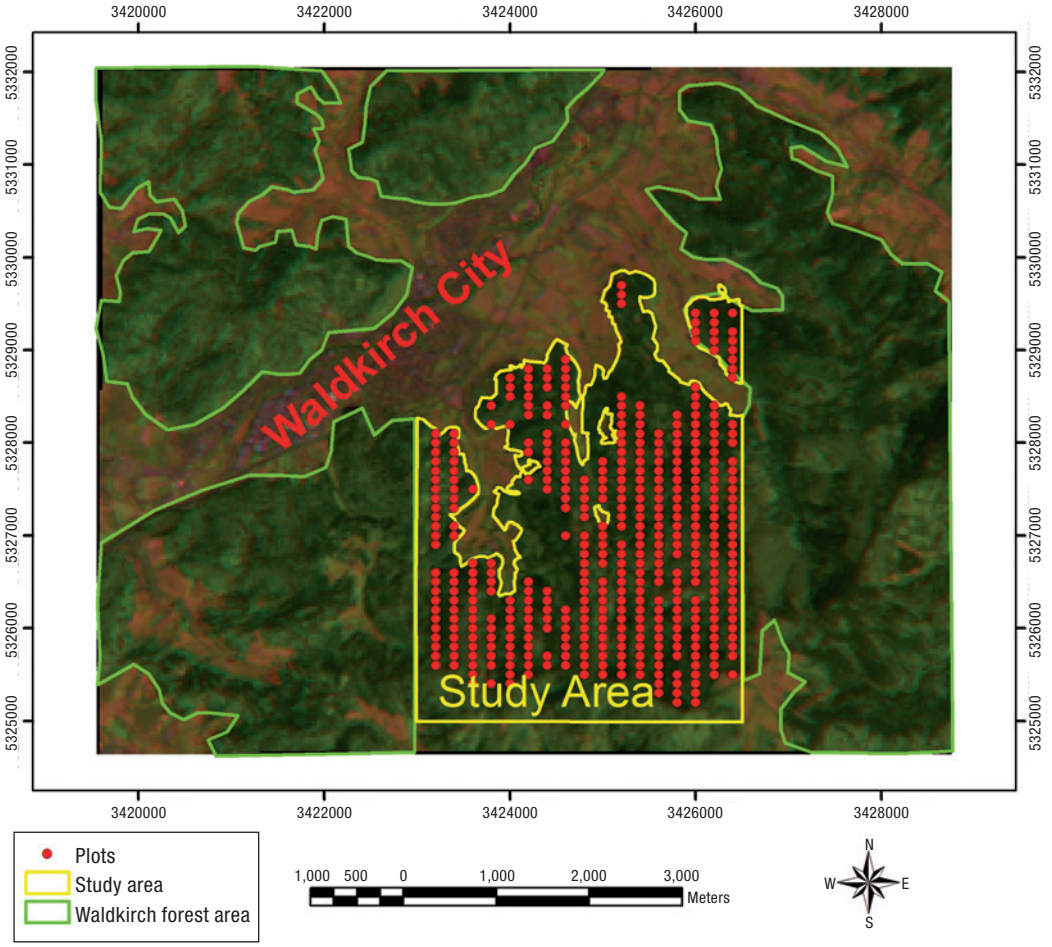

Figure 1. Spread of sample plots and position of study area in the municipal forest of Waldkirch region.

returns above 2 meters height break of ALS height data were computed by "*CloudMetrics*" syntax in FUSION/LDV software. These variables extracted for all sample plots that fell within study area with 900 square meters equal to plot size. In addition, a canopy cover map was produced by computing the number of returns greater than 2 m height break as canopy threshold height divided by the total number of first returns within each cell (30 m grid). For wall-to-wall mapping, the study area was tessellated into a 30 meter squared cells to contain sampled inventory plots with TM pixel size. For more details on mathematical equations and functions, please refer to FUSION manual (McGaughey, 2010).

Landsat5-TM data

A small window corresponding to the study area was selected from a Landsat5-TM image taken on 13 August 2003. These images were co-registered to an orthorectified true color composite of a digital aerial photo using tie points and DEM extracted from LiDAR data with RMSE of 0.79 pixels. Tasseled cap transformation, standardized principal components (PCA) and some famous and most used vegetation indices including SVI (Simple Ratio Vegetation Index), TVI (Transformed Vegetation Index), NDVI (Normalized Difference Vegetation Index), RVI (Ratio Vegetation Index), DVI (Differential Vegetation Index), NRVI (Normalized Ratio Vegetation Index), TTVI (Thiam's Transformed Vegetation Index), CTVI (Corrected Transformed Vegetation Index) and GNDVI (Green Normalized Difference Vegetation Index) were produced for quantifying and enhancing biophysical characteristics. Spectral values of main and artificial TM bands were extracted corresponding to locations of the 411 plots as observed units and other pixels as target units.

Table 1. Descriptive statistical analysis of validation, training and total sample sets

Descriptive	Volume/ha			Basal area/ha		
	Validation	Training	Total	Validation	Training	Total
Mean	418.23	414.36	415.33	34.77	32.74	33.25
Median	393	392	393	33	32	32
Standard deviation	226.37	221.93	222.77	14.99	14.46	14.61
Minimum	12	11	11	2	3	2
Maximum	1,089	1,087	1,089	68	71	71

Algorithms implementations

In all implementations, 75% of plots were randomly selected as training samples and the remaining 25% were selected as test or validation sample sets. Descriptive analysis of training, validation and total sample sets are showed in the Table 1.

In *k*-NN implementations, the number of *k*-NNs, the type of distance measurement and weight for nearest neighbours are considered as the three important parameters. To determine the optimal *k* (number of nearest neighbors), the *v*-fold and leave-one-out (LOO) cross-validation selection methods were used based on a *k* search range. Cross-validation is a well established technique that can be used to obtain estimates of model parameters that are unknown. In this method, by applying 10 folds and the given range of *k* (1 to 64) on each distance metrics, the algorithm will then calculate the sum square of errors in validation sample sets for each *k* value and then it finds the best *k* that has produced the lowest validation error (Fig. 2). For an efficiency comparison of distance measurements, the four distance measurements of Euclidean [3], squared Euclidean [4], Manhattan [5] and Chebychev [6] were individually tested as both weighted and non-weighted. In all implementations, the independent variables were standardized through a simple transformation from 0 to 1.

$$D(x,p) = \sqrt{(x-p)^2} \quad [3]$$

$$D(x,p) = (x-p)^2 \quad [4]$$

$$D(x,p) = \mathrm{Abs}(x-p) \quad [5]$$

$$D(x,p) = \mathrm{Max}(|x-p|) \quad [6]$$

In all formulas, *D* is the distance between a target and a reference unit, *x* is target unit and *p* is reference unit

In order to assessment of efficiency of used different data sources *i.e.* TM and its artificial bands, LiDAR

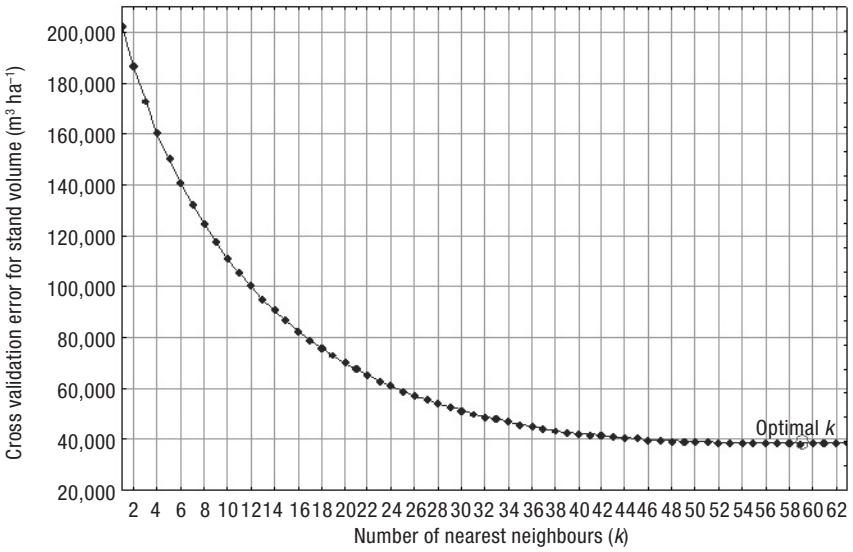

Figure 2. The graph of cross-validation error (sum of squared observed errors in training and test samples) in *k* number of nearest neighbors to find optimal *K* with lowest error.

height and density statistics metrics and integration of TM and LiDAR data for forest variables estimations, the different *k*-NN implementations with weighted squared Euclidean distance metric were done on these data sources.

In RF, good performance depends on the regularization of the decision tree and stopping parameters. For determination of an optimal number of trees, 400 initial trees were used to produce a graph, which shows the average squared error rates against each number of trees for training and test samples. One of the main parameters, which should be used to determinate RF, is *k* predictor (independent variables) in each node for predicting dependent values (response). The simplest way to determine *k* is calculating the square root of total independent variables ($k \leq \sqrt{m}$, *m* is number of independent variables). Default rates were also used to stop splitting parameters and conditions, which stop processes of growing trees when stopping conditions are reached. The used stopping parameters for all estimations consisted of minimum 1 child in a node and maximum 200 nodes in each tree to stop growing the trees. The RF implementation was repeated on the difference data sources by RBF kernel as the best kernel which is produced the higher results compared to other used kernels.

Performance of SVR is affected by three parameters; capacity (*C*), which presents a trade-off between the model's complexity and the amount to which deviations larger than (*C*) are tolerated; epsilon (ε), which controls the width of the δ-insensitive zone used to fit the training data, and gamma (γ), as a kernel function parameter. The kernel parameters can be selected by some means such as prior knowledge; user experience or it can be determined by fixing a parameter and controlling other parameters through the highest cross-validation accuracy. In this study, three different kernel models were examined; RBF [7], polynomial [8] and sigmoid [9] in a fixed 0.016 of gamma and based on "1/ (number of independent variables)" (Hsu *et al.* 2010).

$$K(X_i, X_j) = (\gamma\, X_i^{\mathrm{T}}\, X_j + r)^d, \gamma > 0 \qquad [7]$$

$$K(X_i, X_j) = \exp(-\gamma \left\| X_i - X_j \right\|^2), \gamma > 0 \qquad [8]$$

$$K(X_i, X_j) = \tanh(\gamma\, X_i^T X_j + r) \qquad [9]$$

Here, γ, *r*, and *d* are kernel parameters, X_i and X_j are observed and predicted values respectively.

Two other parameters were selected using 10-fold cross-validation with 1,000 iterations for minimizing the error function (Schölkopf *et al.*, 1998) through a specified grid search method (Hsu *et al.*, 2010) to determine rates for the best capacity and epsilon. The specified grid search included a capacity range from 1 to 64, which is equal to range of input variables (Mattera and Haykin, 1999) and epsilon values from 0.1 to 0.5. In order to applying different data sources, the SVR implementations were performed using each of data sources in different rate of gamma based on "1/(input variables)" and epsilon values from 0.1 to 0.5, but different capacities equal with number of input data source variables.

In ANN**,** the type of neural network, number of input variables and hidden units and settings of various control parameters in the training algorithms all have an effect on the performance of a network. In this study, three most used types of neural networks' kernels *i.e.* RBF, GRNN and MLP kernels were tested to find the network that will produce the best results and estimations. The number of neural networks was set at 1,000 in all performances. In addition, the ANN implementations were repeated on the different data sources by RBF kernel for volume/ha and GRNN kernel for basal area/ha estimations.

Validation and quality performance assessment

Validity of performances was examined using regression diagnostics metrics of root mean square error (RMSE) [10], percentage RMSE [11], bias [12] and percentage bias [13] using the 25% unused samples.

$$RMSe = \sqrt{\sum_{i=1}^{m} (E_i - O_i)^2 \Big/ m} \qquad [10]$$

$$\text{Bias} = \sum_{i=1}^{m} (E_i - O_i) \Big/ m \qquad [11]$$

$$Percentage\ RMSe = \frac{\sqrt{\sum_{i=1}^{m} \left(E_i - O_i\right)^2 \Big/ m}}{\sum_{i=1}^{m} \left(O_i\right) / m} \times 100 \qquad [12]$$

$$\text{Percentage Bias} = \frac{\sum_{i=1}^{m} \left(E_i - O_i\right) \Big/ m}{\sum_{i=1}^{m} \left(O_i\right) / m} \times 100 \qquad [13]$$

Where E is estimation values from implementation of algorithms in m validation samples, O is observation values and m is the number of validation samples

Results

Results of k-NN implementations (Table 2) showed that considering weight on distance provides a more effective estimate of target pixels or places, in places where estimates for volume and basal area are required. It means that giving progressively greater weight on references, those units that are closer or more similar to a target unit as squared could be more useful. Secondly, it was demonstrated that the squared Euclidean distance could better produce volume and basal area in target units with lower RMSE and bias compared to other metrics. In addition, using a larger k could reduce the bias and RMSE rates. Fig. 3 shows scatter graph of observed basal area values versus predicted in test samples. Results of using different data sources in k-NN performances with weighted squared Euclidean

Table 2. Results of k-NN implementations for plot-level volume/ha and basal area/ha estimations

Variables	k (optimal)	Distance measure	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	64(64)	Chebychev*	174.20	41.16	–14.43	–3.53
	64(14)	Chebychev	169.92	40.15	–0.15	–0.03
	64(38)	Manhattan*	168.94	39.91	2.22	0.52
	64(31)	Manhattan	164.89	38.96	–2.04	–0.48
	64(59)	Squared Euclidean*	161.48	38.15	–4.03	–0.96
	64(14)	Squared Euclidean	169.27	39.99	3.01	0.70
Basal area/ha (m^3 ha^{-1})	64(59)	Squared Euclidean*	11.79	34.80	0.03	0.09
	64(15)	Squared Euclidean	12.38	36.53	0.45	1.31
	64(14)	Chebychev	12.26	36.19	0.5	1.47
	64(48)	Chebychev*	12.18	35.97	–0.08	–0.24
	64(38)	Manhattan *	12.32	36.36	0.41	1.19
	64(31)	Manhattan	11.97	35.33	0.24	0.71

* Weighted.

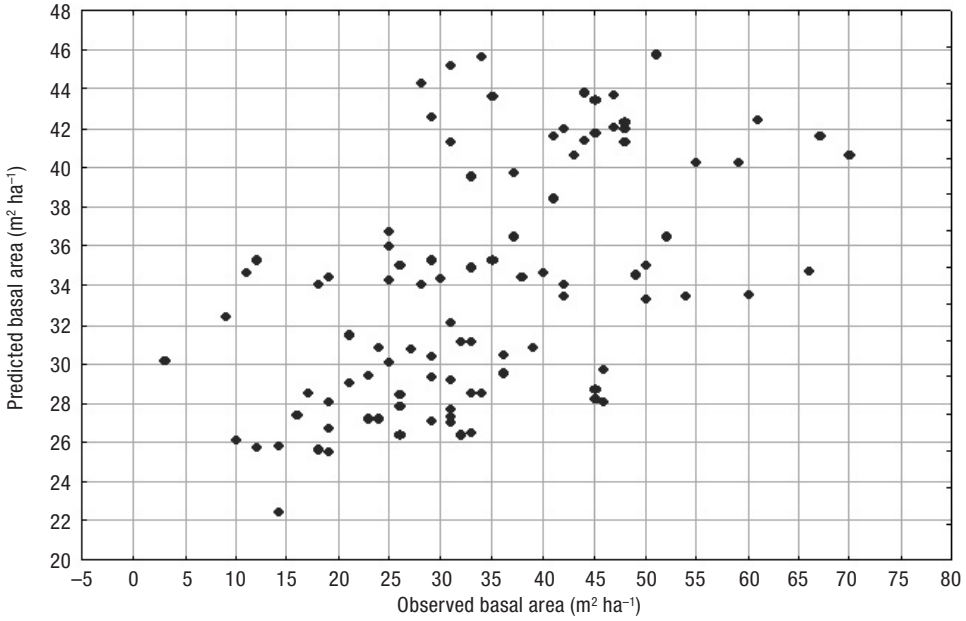

Figure 3. The scatter plots of predicted and observed basal area (m^2 ha^{-1}) in the test samples using k-NN with squared Euclidian distance metric.

Table 3. M Results of using different data sources in *k*-NN performances with weighted squared Euclidean distance

Variables	Data	*k*-NN parameters		Validity metrics			
		K (optimal)	Distance measure	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	TM	23(23)	Squared Euclidean*	196.07	46.47	–11.40	–2.76
	LiDAR	41(40)	Squared Euclidean*	162.68	38.44	–5.62	–1.34
	TM&LiDAR	64(59)	Squared Euclidean*	161.48	38.15	–4.03	–0.96
Basal area/ha (m^2 ha^{-1})	TM	23(23)	Squared Euclidean*	13.07	38.59	–0.56	–1.68
	LiDAR	41(40)	Squared Euclidean*	11.93	35.20	0.03	0.09
	TM&LiDAR	64(59)	Squared Euclidean*	11.79	34.80	0.03	0.09

distance showed that combination of LiDAR and TM data could produce estimations with lower absolute and percentage RMSE and Bias compared to using only LiDAR or TM data (Table 3).

Results of the implementation of SVR with different kernels and their best given parameters (Table 4), showed that using RBF kernel could produce lower RMSE and bias compared to other kernels. In volume estimation, the RMSE rate of applying polynomial function kernels is close to those obtained by RBF kernels; but results of performance of polynomial function kernels was very biased compared to those obtained by RBF. Fig. 4 shows the scatter graph of observed volume values *versus* predicted values in the test samples by implementation of SVR with RBF kernels and suitable hyper plane parameters. In basal area estimates, RMSE and bias rates obtained by using three kernels are very similar with only very slight difference. Table 5 shows the results of using different data sources in SVR implementation using RBF kernel as the best kernel that could produce estimations with lowest error. As it show and expected in advance, combination of LiDAR and TM data could produce best results.

Performance of RF depends on determining the number of trees and number of predictors in each node for producing a good response. Table 6 shows results of RF performances using the default number of predictors in each node and the optimal number of trees with different subsample portions in RFs. In volume/ha estimation, the performance results of RF showed that using a subsample portion of 60%, 7 predictors in each node and 400 initial trees could better predict volume with lower RMSE compared to a 50% percent subsample portion with 7 predictors, but the predictions were very biased. However, for basal area estimates, the best performance was produced using a 50% percent sub sample portion and 2 children in each node with the same 400 initial number of trees and 7 predictors. As Table 7 shows, combination of LiDAR and TM data for volume/ha estimation could produce better results compared to using only LiDAR or TM data. However, in basal area/ha estimation, using LiDAR data as alone and as combined to TM data produced slightly similar results.

In ANN, with an application of a fixed 1,000 number of neural networks, the results of volume estimation showed that, the RBF network compared to other types of networks could better test neural networks. However, in basal area, the GRNN type of network produced lower results for RMSE and bias

Table 4. Results of SVR implementations for plot-level volume/ha and basal area/ha estimations

Variables	Kernel	Hyper plane parameters			Validity metrics			
		Gamma	Capacity	Epsilon	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	RBF	0.016	6	0.26	156.02	36.86	0.48	0.11
	Polynomial	0.016	19	0.12	156.29	36.92	–19.90	–4.93
	Sigmoid	0.016	20	0.50	165.73	39.16	15.10	3.44
Basal area/ha (m^2 ha^{-1})	RBF	0.016	55	0.20	11.55	34.10	–1.04	–3.17
	Polynomial	0.016	20	0.28	12.05	35.57	–1.03	–3.15
	Sigmoid	0.016	20	0.18	12.05	35.57	–1.28	–3.93

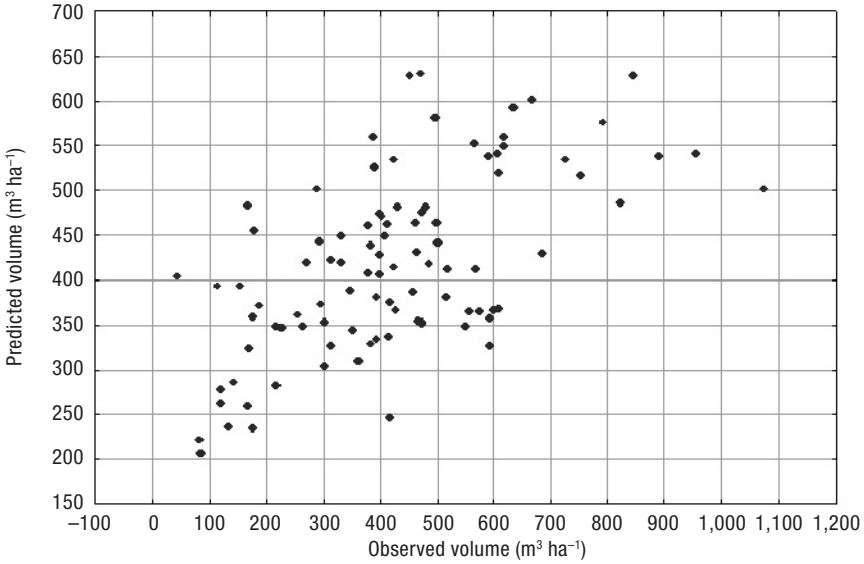

Figure 4. The scatter plots of predicted and observed volume (m^3 ha^{-1}) in the test samples using SVR with RBF kernel and suitable hyper plane parameters

(Table 8). Results of using different data sources are also given in Table 9.

Comparative results of implementations are given in Table 10. For volume estimation, results obtained by the SVR algorithm were better than those obtained by the other algorithms that were tested. This result is in agreement with that of Shataee *et al.* (2012), a study in which the algorithms SVR, RF, and *k*-NN were compared to estimate forest attributes using ASTER data in the Hyrcanian forest and reported that SVR performed well. In basal area estimation, regression methods yielded close results; however, implemen-

Table 5. Results of using different data sources in SVR implementation using RBF kernel

Variables	Data	Hyper plane parameters			Validity metrics				
		Kernel	Gamma	Capacity	Epsilon	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	TM	RBF	0.042	20	0.10	194.82	46.03	–33.41	–8.80
	LiDAR		0.025	40	0.50	170.36	40.25	16.27	3.70
	TM &LiDAR		0.016	56	0.10	156.02	36.86	0.48	0.11
Basal area/ha (m^2 ha^{-1})	TM	RBF	0.042	23	0.10	13.13	38.76	–2.06	–6.49
	LiDAR		0.025	38	0.20	12.07	35.64	–1.61	–5.00
	TM &LiDAR		0.016	55	0.20	11.55	34.10	–1.04	–3.17

Table 6. Results of RF implementations for plot-level volume/ha and basal area/ha estimations

Variables	*k* predictor	Decision tree and stopping parameters			Validity metrics			
		Number of trees	Subsample portion	Minimum number of children in each node	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	7	400	0.60	5	176.17	40.04	–27.2	–6.59
	7	400	0.50	5	185.20	46.38	–2.5	–0.63
Basal area/ha (m^2 ha^{-1})	7	400	0.60	5	12.87	36.79	–1.99	–6.05
	7	400	0.50	5	12.64	38.54	0.05	0.1

Table 7. Results of using different data sources in RF implementations

Variables	Data	Decision tree and stopping parameters				Validity metrics			
		k predictor	Number of tree	Subsample portion	Minimum number of child in each node	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	TM	5	400	0.6	5	197.07	50.72	22.45	5.40
	LiDAR	6	400	0.6	5	185.30	47.94	20.14	4.90
	TM&LiDAR	7	400	0.6	5	176.17	40.04	–27.2	–6.59
Basal area/ha (m^2 ha^{-1})	TM	5	400	0.6	5	13.59	43.87	1.69	5.11
	LiDAR	6	400	0.60	5	12.53	38.86	0.56	1.71
	TM&LiDAR	7	400	0.50	5	12.64	38.54	0.05	0.17

Table 8. Results of ANN implementations for plot-level volume/ha and basal area/ha estimations

Variables	Network parameters		Validity metrics			
	Number of networks	Type of network	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	1,000	RBF	185.73	44.64	–3.98	–0.96
Basal area/ha (m^2 ha^{-1})	1,000	GRNN	12.06	35.55	–0.60	–1.81

Table 9. Results of using different data sources in ANN implementations with RBF network for volume/ha and GRNN for basal area/ha estimation

Variable	Data	Network parameters		Validity metrics			
		Number of network	Type of network	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	TM	1,000	RBF	190.26	50.58	45.68	10.71
	LiDAR	1,000	RBF	186.68	43.64	–14.20	–3.43
	TM&LiDAR	1,000	RBF	185.73	44.64	–3.98	–0.96
Basal area /ha (m^2 ha^{-1})	TM	1,000	GRNN	13.67	43.08	2.05	6.08
	LiDAR	1,000	GRNN	13.94	44.09	2.00	5.97
	TM&LiDAR	1,000	GRNN	12.06	35.55	–0.60	–1.81

Table 10. Performance assessment of algorithms for volume/ha and basal area/ha estimation

Variables	Algorithm	RMSE	RMSE (%)	Bias	Bias (%)
Volume/ha (m^3 ha^{-1})	*k*-NN	161.48	38.15	–4.03	–0.96
	SVR	156.02	36.86	0.48	0.11
	RF	185.20	46.38	–2.50	–0.63
	ANN	185.73	44.64	–3.98	–0.96
Basal area/ha (m^2 ha^{-1})	*k*-NN	11.79	34.80	0.03	0.09
	SVR	11.55	34.10	–1.04	–3.17
	RF	12.64	38.54	0.05	0.17
	ANN	12.06	35.55	–0.60	–1.81

tation of *k*-NN using squared Euclidian distance metric had slightly higher results with lower RMSE and bias.

Discussion

In a case study and with integration of LiDAR and optical Landsat-TM data for forest attribute estimation using four different theory-based non-parametric algorithms including *k*-NN, SVM, RF and ANN, results showed that non-parametric algorithms have superiority in estimations compared to habitual statistical methods like multiple regressions. The primary and pretest results of analysis using multiple regression with backward stepwise model building showed that estimations are not acceptable compared to those which obtained by non-parametric algorithms (for instance, results of volume estimation assessment were the RMSE = 203 m^3 ha^{-1}; percentage RMSE = % 48.54; Bias = –18.37 m^3 ha^{-1} and percentage Bias = % –4.58).

The result of *k*-NN demonstrated that the squared Euclidean distance could better produce volume and basal area in target units with lower RMSE and bias compared to other metrics. It confirmed outcomes reported in other research by Kajisa *et al.* (2008) in that *k*-NN implementations with Euclidean distance had consistently smaller RMSE and percentage RMSE than those with other distances. In other studies, the squared Euclidean was the most used distance metric (Franco-López *et al.*, 2001; Sironen *et al.*, 2010) and could produce better results compared to other metrics (Shataee *et al.,* 2012). Some studies (Nilsson, 1997) have shown that applying more *k* may lead to biased results due to averaging the values of more pixels, but in this study, and confirmed by other research (Finely *et al.*, 2006; McRoberts, 2007), higher *k* values could reduce the bias and RMSE rates.

Results of the implementation of SVR with different kernels showed that using RBF kernel could produce lower RMSE and bias compared to other kernels. It means that the RBF kernel with its parameters, obtained by the specified grid search method, could prepare and map the best hyper plane or feature space for the SVR model. This is agreed with other studies (Durbha *et al.,* 2007; Cortez and Morais, 2007), the RBF is the most popular kernel in SVMs.

In both variables, the algorithms *k*-NN and SVR had better performance than RF and ANN algorithms. McInery and Nieuwenhuis (2009) also reported that *k*-NN compared to RF could produce lower RMSE and bias in plot level estimations of stand volume and stand basal area in Ireland. In addition, Shafri and Ramle (2009) indicated that SVM in comparison to decision tree algorithms had better accuracy in classification of satellite data. Niska *et al.* (2010) also reported that in a comparison with k-most similar neighbour and ANN models, SVR identified as best-suited method for accurate prediction.

Conclusions

Generally, results of study showed that non-parametric algorithms to predict forest parameters using airborne laser scanning and TM data had different performances particularly in terms of their capacity for volume estimation. This study exposed that the SVR algorithm, because of its robustness to dimensionality and ability to generalize as well as accounting non-linear relationships through regularization of kernel parameters can be suggested as the best non-parametric algorithm for regression-based application.

Although, results of this study are valuable and important for extracting and retrieving of forest quantities information, however, regarding to type of forest, spatial spread of study area and condition of the study of area, which is a managed forest and forest structure is comprised from one to two stories, then results and outcomes of this study perhaps can be applied and tested in similar forests elsewhere and/or adopted in other types of forest with same structure and composition.

Acknowledgements

I gratefully acknowledge Professore Dietter Pelz, the former head of forest biometry department, Frieburg University for providing the forest inventoried data and preparing some facilities in my sabatical duration. Also, I would like to say thanks to Dr. Holger Weinacker assistance research of department of remote sensing and landscape information systems, Frieburg University for providing the Lidar data and usefull scientific guides.

References

Andersen H, 2003. Estimation of critical forest structure metrics through the spatial analysis of airborne laser

scanner data. PhD dissertation. University of Washington, Seattle, WA, USA.

Andersen HE, Foster JR, Reutebuch SE, 2003. Estimating forest structure parameters within Fort Lewis Military Reservation using airborne laser scanner (LIDAR) data. In: Proceedings, 2nd Inter-national Precision Forestry Symposium. Seattle, Washington, University of Washington, College of Forest Resources. pp: 45-53.

Atzberger C, 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. Remote Sensing of Environment 93: 53-67.

Breidenbach J, Nothdurft A, Kandler G, 2010. Comparison of nearest neighbor approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. European Journal of Forest Resources.

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

Chen G, Hay GJ, Zhou Y, 2010. Estimation of forest height, biomass and volume using support vector regression and segmentation from Lidar transects and Quickbird imagery. 18th International Conference on Geoinformatics, 18-20 June 2010, Beijing, China.

Chen G, Hay GJ, 2011. A support vector regression approach to estimate forest biophysical parameters at the object level using airborne LiDAR transects and Quick Bird data. Photogrammetric Engineering & Remote Sensing 77(7): 733-741.

Chen G, Wulder MA, White JC, Hilker T, Coops NC, 2012. LiDAR calibration and validation for geometric-optical modelling with Landsat imagery. Remote Sensing of Environment 124: 384-393.

Dalponte M, Coops NC, Bruzzone L, Gianelle D, 2009. Analysis on the use of multiple returns Lidar data for the estimation of tree stems volume. IEEE journal of selected topics in applied earth observations and remote sensing 2(4): 310-318.

Durbha SS, King RL, Younan NH, 2007. Support vector machines regression for retrieval of leaf area index from multi angle imaging spectroradiometer. Remote Sensing of Environment 107: 348-361.

Eskelson BNI, Hailemariam T, Barrett TM, 2009. Estimating current forest attributes from panelled inventory data using plot-level imputation: a study from the Pacific Northwest. Forest Science 55: 64-71.

Finley AO, McRoberts RE, Ek AR, 2006. Applying an efficient k-nearest neighbour search to forest attribute imputation. Forest Science 52: 130-135.

Franco-López H, Ek AR, Bauer ME, 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors' method. Remote Sensing of Environment 77(3): 251-274.

Jin Y, Liu C, 1997. Biomass retrieval from high-dimensional active/passive remote sensing data by using artificial neural networks. International Journal of Remote Sensing 18(4): 971-979.

He Q, Xu H, Zhang Y, 2011. Estimation of forest biophysical parameters using small-footprint Lidar with low density in a coniferous forest, International Symposium on LiDAR and Radar Mapping 2011: technologies and applications, Nanjing, China, May 26, 2011.

Holmgren J, 2004. Prediction of tree height, basal area, and stem volume using airborne laser scanning, Scandinavian Journal Forest Resources 19: 543-553.

Hsu CW, Chang CC, Lin CJ, 2010. A practical guide to support vector classification. Department of computer science, National Taiwan University, Taipei, Taiwan. http://www.csie.ntu.edu.tw/~cjlin.

Hyvonen P, 2007. The updating of forest resource data for management planning for privately owned forests in Finland. Academic dissertation, Faculty of Forestry, University of Joensuu. 40 pp.

Kajisa T, Murakami T, Mizoue N, Kitahara F, Yoshida S, 2008. Estimation of stand volumes using the k -nearest neighbor method in Kyushu, Japan. Journal of Forest Research 13(4): 249-254.

Latifi H, Nothdurfet A, Koch B, 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: optimization of variable selection on optical/LiDAR-derived predictors. Forestry Journal 83: 395-407.

Lefsky MA, Cohen WB, Acker SA, Parker GG, Spies TA, Harding D, 1999b. LiDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment 70: 339-61.

Lefsky MA, Cohen WB, Harding DJ, Parker GG, Acker SA, Gower ST, 2001b. Lidar remote sensing of aboveground biomass in three biomes. In the international archives of the Photogrammetry, remote sensing and spatial information science, Vol XXXIV, Part 3/W4, Commission III, Annapolis MD, 22-24 October. pp: 155-60.

Lim K, Treitz P, Wulder M, St-Onge B, Flood M, 2003. LiDAR remote sensing of forest structure, Progress in Physical Geography 27(1): 88-106.

Lu D, Chen Q, Wang G, Moran E, Batistella M, Zhang M, Laurin GV, Saah D, 2012. Aboveground forest biomass estimation with Landsat and Lidar data and uncertainty analysis of the estimates. International Journal of Forestry Research.

Makela H, Pekkarinen A, 2004. Estimation of forest stands volumes by Landsat TM imagery and stand-level field-inventory data. Forest Ecology and Management 196: 245-255.

Maltamo M, Eerikainen K, Packalen P, Hyyppa J, 2006a. Estimation of stem volume using laser scanning based canopy height metrics. Forestry 79: 217-229.

Maltamo M, Malinen J, Packalen P, Suvanto A, Kangas J, 2006b. Non-parametric estimation of stem volume using laser scanning, aerial photography, and stand register data. Canadian Journal of Forest Resources 36: 426-436.

Mattera D, Haykin S, 1999. Support vector machines for dynamic reconstruction of a chaotic system. In: Advances in Kernel methods: support vector learning (Schölkopf B, Burges CJC, Smola AJ, eds). Cambridge, MA, MIT Press. pp: 211-242.

McGaughey RJ, 2010. Manual of FUSION/LDV: software for LIDAR data analysis and visualization. United State, Department of Agriculture, Forest services, research station. 153 pp.

McInerney D, Nieuwenhuis M, 2009. A comparative analysis of k-NN and decision tree methods for the Irish National Forest Inventory, International Journal of Remote Sensing 30: 4937-4955.

McInerney DO, Suárez-Mínguez J, Valbuena R, Nieuwenhuis M, 2010. Forest canopy height retrieval using lidar data, medium resolution satellite imagery and Knn estimation in Aberfoyle, Scotland. Forestry 83(2): 195-206.

McRoberts R, Tomppo E, Finley A, Heikkinen J, 2007. Estimating aerial means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. Remote Sensing of Environment 111: 466-480.

McRoberts R, 2009. Diagnostic tools for nearest neighbours techniques when used with satellite imagery. Remote Sensing of Environment 113: 489-499.

Mohammadi J, Shataee Sh, Yaghmaee F, Mahiny A, 2010. Modeling forest stand volume and tree density using Landsat ETM+ data. International Journal of Remote Sensing 31(11): 2959-2975.

Naesset E, 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 52: 49-56.

Naesset E, 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80: 88-99.

Naesset E, 2004. Practical large-scale forest stand inventory using a small airborne scanning laser. Scandinavian Journal Forest Resources 19: 164-179.

Nelson R, Oderwald R, Gregoire TG, 1997. Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. Remote Sensing of Environment 60: 311-26.

Nilsson M, 1996. Estimation of tree heights and stand volume using an airborne LiDAR system. Remote Sensing of Environment 56: 1-7.

Nilsson M, 1997. Estimation of forest variables using satellite image data and airborne Lidar. PhD thesis. Swedish University of Agricultural Sciences. Department of Forest Resource Management and Geomatics, Acta Universitatis Agriculturae Sueciae. Silvestria 17.

Niska H, Skon JP, Packalen P, Tokola T, Maltamo M, Kolehmainen M, 2010. Neural networks for the prediction of species-specific plot volumes using airborne laser scanning and aerial photographs, IEEE Trans. Geosciences of Remote Sensing 48(3): 1076-1085

Packalén P, 2009. Using airborne laser scanning data and digital aerial photographs to estimate growing stock by tree species. Academic dissertation, Faculty of Forest Sciences, University of Joensuu, Finland. 41 pp.

Packalén P, Maltamo M, 2007. The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. Remote Sensing Environ 109: 328-341.

Popescu SC, Wynne RH, 2004. Seeing the trees in the forest: using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering & Remote Sensing 70(5): 589-604.

Schölkopf B, Smola A, Müller KR, 1998. Nonlinear component analysis as a kernel Eigen value problem. Neural Computation 10: 1299-1319.

Shafri HZM, Ramle FSH, 2009. A comparison of support vector machine and decision tree classifications using satellite data of Langkawi Island. Information Technology Journal 8: 64-70.

Shataee Sh, Kalbi S, Fallah A, Pelz D, 2012. Forest attributes imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. International Journal of Remote Sensing 33: 6254-6280.

Sironen S, Kangas A, Maltamo M, 2010. Comparison of different non-parametric growth imputation methods in the presence of correlated observations. Forestry 83(1): 39-51.

Stojanova D, Panov P, Gjorgjioski V, Kobler A, Dzeroski S, 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. Ecological Informatics 5: 256-266.

Straub C, Weinacker H, Koch B, 2010. A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. European Journal of Remote Sensing 129: 1069-1080.

Vapnik V, 1995. The nature of statistical learning theory. New York, Springer-Verlag.

Weinacker H, Koch B, Heyder U, Weinacker R, 2004. Development of filtering, segmentation and modelling modules for lidar and multispectral data as a fundament of an automatic forest inventory system. In: ISPRS Working Group VIII/2 "Laser scanners for forest and landscape assessment". University of Freiburg, Freiburg, Germany.

Walton T, 2008. Sub pixel urban land cover estimation: comparing cubist, random forests, and support vector regression. Photogrammetric Engineering & Remote Sensing 74(10): 1213-1222.

Wang Z, Brenner A, 2009. An integrated method for forest canopy cover mapping using Landsat ETM+ imagery. ASPRS/MAPPS conference, Texas, USA.

Wulder M, 1998. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. Progress in Physical Geography 22: 449-76.

Wulder M, Magnussen S, Harding D, Boudewyn P, Seemann D, 2000. Stability of surface LiDAR height estimates on a point and polygon basis. Proceedings of the 22nd Annual Canadian Remote Sensing Symposium. Victoria, British Columbia, 21-25 August, 433-38.