ORIGINAL ARTICLE

Forest stands volume estimation by using Finnish Multi-Source National Forest Inventory in Stołowe Mountains National Park

Przemko Pachana

Eberswalde University for Sustainable Development, Faculty of Forest and Environment, Alfred-Möller-Strasse 1, 16225 Eberswalde, Germany

Warsaw University of Life Sciences, Faculty of Forestry, Department of Forest Management, Geomatics and Forest Economics, Nowoursynowska 159, 02-776 Warsaw, Poland, phone: +48791709242, e-mail: przemko.pachana@gmail.com

Abstract

The purpose of the present study was to convey to the reader the method and application of the Finnish Multi-Source National Forest Inventory (MS-NFI) that was devised in the Finnish Forest Research Institute. The study area concerned is Stołowe Mountains National Park, which is located in the south-western Poland, near the border with the Czech Republic. To accomplish the above mentioned aim, the following data have been applied: timber volume derived from field sample plots, satellite image, digital map data and digital elevation model. The Pearson correlation coefficient between independent and dependent variables has been verified. Furthermore, the non-parametric k-nearest neighbours (k-NN) technique and genetic algorithm have been used in order to estimate forest stands biomass at the pixel level. The error estimates have been obtained by leave-one-out cross-validation method. The main computed forest stands features were total and mean timber volume as well as maximum and minimum biomass occurring in the examined area. In the final step, timber volume map of the growing stock has been created.

Introduction

All decision-making demands information; in forestry, such information is collected by forest inventories that estimate forest characteristics over a defined area. Such characteristics include the volume of the growing stock or the area of a certain type of forest and, nowadays, also measure forest biodiversity, for example, vegeta-tion occurring in a given area or the volume of dead wood. However, it must be pointed out that measuring trees and ground plots is related to many methodological problems. Forest inventories cover different spatial scales, from the forest stand level to regional or country level and finally to global level. It is obvious that these varying scales also require different methodologies. Nevertheless, the main objective of forest inventory is

to provide statistically unbiased, reliable forest resource information for strategic planning, mainly by decision makers. Making full use of new technology is one of the most important challenges that forest management is facing today. Resource information has to be obtained with a limited budget, which requires better timing of resource assessment activities as well as improved use of multiple data source.

Nowadays, remote sensing is a unique technique for expanding ground plot level information to large area of continuous information. Remote sensing combined with geographic information systems provide a variety of techniques that can be used in forest management planning including prevention of damage, risk assessment and carrying out different activities in critical situations such as forest fires, insect invasions and wind damages. Satellite images provided possibilities for improving forest inventory efficiency at relatively small additional expenses. Therefore, the remote sensing data derived from space-borne sensors has been the key to a more efficient use of forest inventory. It has been proved that increasing availability of supplementary data such as satellite data has made it possible to augment the efficiency of forest inventories as well as satisfy the needs for information at reasonable costs (Franco-Lopez et al. 2001). In the first satellite image experiments directed to forest management inventory, regression and discriminant analysis methods were tested (Tomppo 1988), which resulted in necessity of making the aerial estimates of the parameters as well as pixel level predictions of the variable values separately. Generally, such approach was laborious in obtaining a sufficient degree of detailed information, and the dependence structure of the variables was lost when the predictions of different variables were done independently. In order to avoid this drawback, the Finnish Multi-Source National Forest Inventory (MS-NFI) has been devised. The development of the Finnish MS-NFI began in 1989, and the first effective results were computed in 1990 (Tomppo 1990). The main goal of the discussed method elaboration was the need to obtain forest source information for large areas with maintaining low work effort and overall expenditures. The Finnish MS-NFI conducted by the Finnish Forest Research Institute has used satellite images and digital map data, in addition to field measurements. Nevertheless, the essential data for MS-NFI are derived from ground plots and satellite images, whereas digital map data as well as

digital elevation model improve the accuracy of the prediction significantly. Generally, the method can be used in a straightforward way in different forest conditions and with different remote sensing materials.

The main aim of utilisation of satellite images into forest inventories has been concentrated on the estimation of the variables, such as diameter breast height (DBH), basal area, age and stand volume (Tokala et al. 1996). A nonparametric k-nearest neighbour (k-NN) method has made it possible to estimate all inventory variables at the same time. The nonparametric k-NN technique differs from the traditional estimation and classification methods. The k-NN algorithm as a nonparametric estimation approach can be used in a wide range of estimation and classification applications. Discussed technique classifies objects based on most similar or closest training samples in the feature space. An object is classified by a majority vote of its neighbours, where generally Euclidian distance is applied. The success of discussed technique is based on the fact that it allows simultaneous prediction of all inventory parameters for each pixel of the examined area. The k-NN technique produces both statistics for arbitrary units as well as wall-to-wall maps and, therefore, can be used for ecological and environmental studies (Pakkala et al. 2002).

Method similar to MS-NFI, which uses k-NN algorithm, has been developed in Sweden (Reese et al. 2002). Swedish use their method for a multitude of purposes but mainly in the official Swedish statistics as a basis for post-stratification. Multi-source method that uses k-NN technique has also been developed in the United States, where it serves primarily to post-stratified timber volume and forest area estimation (Franco-Lopez et al. 2001; McRoberts et al. 2002). Furthermore, multi-source inventory methods have also been tested in several other countries such as Austria (Koukal et al. 2005), China (Tomppo et al. 2001) or Italy (Maselli et al. 2005).

Study area

The Stołowe Mountains National Park as a part of the Sudetes range is situated in southwestern Poland, in Kłodzko County in Lower Silesian Voivodeship, near the border with the Czech Republic. In terms of geographical location, the area of interest is situated

between 50°24'30'' and 50°30'15'' north latitude, and 16°16'15'' and 16°26'24'' east longitude. The highest elevation is situated in western part of the study area and is equal to 919 m a.s.l., whereas the lowest elevation is located in northeastern part and is equal to 393 m a.s.l. The climate of the Stołowe Mountains National Park belongs to mountain climate. Precipitation increases with the altitude, exceeding 660 mm in the lowest part of the park and 1000 mm in the highest part. The highest intensity of precipitation falls in July, whereas the lowest falls in January, February and March. Snow remains in the highest part of the park for more than 120 days a year, whereas vegetation period persists around 195 days a year. The Stołowe Mountains landscape started to form 70 million years ago, and its unique shape is the result of erosion that has been affecting this region for hundreds of thousands of years.

National Park was established in 1993 and covers 6327.88 ha, where forest stands occupy 5803.24 ha and non-forest areas cover 524.64 ha. Moreover, in the examined area, strict protection zone has been created, and it is equal to 376 ha. Spruce is the dominant species that occupies total area equal to 4544.58 ha. Second species in terms of quantity occurrence in the study area concerns beech, which covers 660.95 ha. Subsequently occurring birch occupies area equals to 236.74 ha, and larch with total area 173.76 ha. Other species such as pine, fir, sycamore, ash, alder and oak occupy a total area equal to 194.90 ha (Borusiewicz 2009).

In terms of forest habitat type, mountain deciduous mixed forest has the largest share, which is equal to 62.4%. The second largest forest habitat type is mountain deciduous forest and mountain coniferous mixed forest with shares of 18.5% and 16.6%, respectively. Participation of the forest stands not compatible to the forest habitat type is equal to 50%, whereas compatible and partly compatible stands are equal to 15% and 35%, respectively (Borusiewicz 2009).

MATERIALS

Field data

In order to obtain reference data, the whole area of the Stołowe Mountains National Park has been covered uniformly by ground plots. The size of single-field sample plot was equal to 500 m², whereas the spacing between individual plots were 400 m × 400 m. Ground plots establishment and data collection have been done for the purposes of National Park conservation plan, including protection plan of forest ecosystems. Originally, 395 ground plots have been established; however, for the computation purposes, 62 plots have been discarded, because some of them were not located directly in the stand but, for instance, on the meadows or roads. Data from field sample plots were collected in August and September 2009, whereas in 2010, all ground plots have been checked and fixed by the Forest Faculty of SGGW employees (Warsaw University of Life Sciences). Measurements include DBH of all trees as well as the heights of selected trees. Height measurements were estimated by using Suunto devices, whereas DBH measurements were done by using caliper. Inventory teams measured DBH between intervals 8–12 cm for each tree, which was located in the core zone 50 m² of the sample plot, whilst in the whole plot of 500 m², the measurements were taken for each trees that exceeded 12 cm of the DBH. After collecting the necessary data, timber volume in cubic meter per hectare was calculated for each plot.

Satellite data

The satellite image concerning study area has been downloaded from USGS Global Visualization Viewer (http://glovis.usgs.gov), which provides free satellite images for the whole Earth. The image was taken in August 2007 by Landsat 7 satellite from the altitude of 705 km. Acquired data are stored in GeoTIFF format, and each band is delivered as a grayscale, uncompressed, 8-bit string of unsigned integers values. The metadata (MTL) file is included as a text file and provides general information about the image for all available bands. Owing to the fact that Landsat data files are stored in GeoTIFF, it is possible to reference raster image for a given geodetic model or map projection. The image that was derived from Landsat 7 satellite consisted of seven spectral bands with a spatial resolution equal to 30 m × 30 m for each pixel.

Digital map data

Digital map data in raster format with spatial resolution of 30 m × 30 m have been used in order to reduce the estimate errors. In order to obtain the above-mentioned aim, the examined area has been divided into forestry

land and non-forestry land, such as arable land, urban areas, build-up areas and roads. Forestry land occupies 92%, whilst non-forestry land covers 8% of the whole study area. The errors are reduced by using two statistical methods into computations (Katila et al. 2000; Katila and Tomppo 2002). The first method uses confusion matrix derived from land classes distribution on the basis of ground plots and map data, and the second uses stratification of the field sample plots on the basis of map data.

Digital Elevation Model

A digital elevation model (DEM) have been obtained by using Airborne Laser Scanners and was used for two purposes. The first concerns stratification on the basis of elevation data, and the second relies on correcting the spectral values by reference to the angle between solar illumination and the terrain normal. Stratification in this case uses the maximum vertical distance from each pixel to its nearest neighbours (Katilla and Tomppo 2001). Used DEM was stored in the raster format with horizontal spatial resolution of 30 m × 30 m and vertical resolution of 0.1 m.

METHODOLOGY

Data Pre-processing and Rectification

In order to perform calculations, it is necessary to install R Statistical Software with additional packages: RANN, genalg, raster and rgdal. Furthermore, two external packages, which were specially devised for the MS-NFI computation purposes, are necessary. Digital map data were created from vector data and rasterised to the pixel size of 30 m × 30 m by using ArcGIS Software. DEM also had to be aggregated to the same pixel level of 30 m × 30 m, because all raster data that have been pulled into calculations must have the same pixel size. Landsat satellite image has been processed by using ERDAS IMAGINE 2014 Software. In order to obtain the above-mentioned aim, it was necessary to use layer stack process that combines image bands into single multi-band image file. Subsequently, subset image had to be created, which allows for the extraction of the AOI. Furthermore, all raster images as well as ground plots had to be rectified to the uniform coordinate system. In order to obtain the above-mentioned goal, all

images have been assigned by using R to ETRS 1989 Transverse Mercator Coordinate System. Field sample plots have been allocated to the same spatial reference by using arcGIS. A DEM has been used to remove the variation of the spectral values that were caused by the changes in the slope and aspect of the terrain normal. Before running k-NN prediction, it was necessary to decide which image features were used, how many neighbours were used (k) and finally, how the neighbours were weighted (g), where finally decided to use following values: k = 5 and g = 2.

The k-NN Estimation Method

The growing stock can be estimated based on the assumption that each pixel to be estimated has a timber volume similar to the reference pixel, which is derived from field sample plots. Using the k-NN technique, the forest stands volume has been calculated as the weighted mean timber volume of the reference pixel of the k-NN in the feature space. The nonparametric k-NN algorithm as a classification method uses the distance metric, which is defined in the feature space of the satellite image. The k-nearest field sample plot pixels, with respect to distance metric, were sought for each pixel in terms of forestry land satellite image area. Reference sample plots were established based on the degree of similarity between digital number values derived from satellite image bands and a reference field sample plots. In order to perform it, the Euclidian distance equation has been used:

$$d_{t,r} = \sqrt{\sum_{i=1}^{m} (x_{i,t} - x_{i,r})^2} \qquad (1)$$

where $d_{t,r}$ was designated as the distance between target field sample plots and reference sample plots, whereas $x_{i,t}$ and $x_{i,r}$ were designated as the digital number values of each band, and m stands for the number of satellite bands.

Moreover, a maximum geographical distance in horizontal and vertical directions has been used in order to avoid selecting similar plots from surface, in which the response of image variables to field variables is not equal. The main reason for doing so was the reference variables deriving from ground plots, and the fact that the image variables may vary between

various vegetation zones as well as image sub-areas. Therefore, the feasible set of nearest neighbour for pixel p is as follows:

$$\left\{ p_i \mid d_{p,p_i}^{(x,y)} \leq d_{\max}^{(x,y)},\ d_{p,p_i}^z \leq d_{\max}^z,\ R(p_i) = R(p) \right\} \quad (2)$$

where $d_{p,p_i}^{(x,y)}$ is designated as the geographical horizontal distance from pixel p to pixel p_i, d^z states for the distance in the vertical direction, $d_{\max}^{(x,y)}$ and $d_{\max}^z$ are their maximum values and $R(p)$ is designated as the indicator function of given land classes in this case forestry land and other land area.

The plot weights are sum of satellite image pixel weights over the forestry land mask, which have been computed for each ground plot by using following equations:

$$c_{i,h,u} = \sum_{p \in u_h} w_{i,p} \quad (3)$$

$$c_{i,h,u} = c_{i,h,u}^r \frac{A_{h,u}}{A_{h,u}^r} \quad (4)$$

where $w_{i,p}$ is designated as the sums of ground plot weights to pixel and h states for the map stratum over the pixel which belongs to the unit u.

An example of the stratum is digital map data that divides AOI into forestry land and other land, i is denoted as the weight of particular ground plot, and r states for reduced weight sums. $A_{h,u}^r$ estimates the area of the forestry land on map stratum, and $A_{h,u}$ estimates the area of the forestry land on map stratum not covered by the cloud mask. Through the employment of digital map data, land classes located outside forestry land are not taken into consideration, and only variables within forestry land area are predicted by the weighted averages of the k-NN. The pixel weights $w_{pi,p}$, and estimates y_p were computed by using the following equations:

$$w_{pi,p} = \frac{1}{d_{pi,p}^t} \Big/ \sum_{j=1}^{k} \frac{1}{d_{pj,p}^t} \quad (5)$$

$$y_p = \sum_{i=1}^{k} w_{pi,p} \times y_{pi} \quad (6)$$

where i is designated as the field sample plot, p is an arbitrary pixel, p_j stands for pixel corresponding to the field sample plot, j and d are designated as the distance metric defined in the feature space, k is the number of field sample plots and y_{pi} is the observed timber volume of the reference pixel p_i and $t \geq 0$.

Genetic Algorithm application

The main goal of the genetic algorithm employment is to minimise the timber volume prediction errors of the multi-source inventory. Discussed algorithm was chosen because of the complexity of the optimisation problem. The explanatory variable are sought in such a way that the prediction of the timber volume is optimised; however, it must be pointed out that discrepancy with the pixel size from raster image (0.09 ha) and field sample plot (0.05 ha) are interpreted as a measurement error (Halme and Tomppo 2001). The optimisation problem is solved with the help of distance metric, which provides the lowest value for a linear combination of the root-mean-square error (RMSE) and pixel-level bias. First stage generates the initial population with random weight vectors by computing their fitness value through using the following equation:

$$f(\omega, \gamma, e) = \sum_{j=1}^{n_e} \gamma_j \sigma_j(\omega) + \sum_{j=1}^{n_e} \gamma_{j+n_e} e_j(\omega) \quad (7)$$

where $\gamma > 0$ are the coefficients for standard errors σ_j and biases e_j for forest variable j, whereas ω is denoted as the explanatory variable weight vector to be estimated.

In the second stage, a medipopulation known as intermediate group of weight vectors amongst two populations is created, whilst the fitness values are compared. In the third stage, a new population is formed, where vectors carry out uniform crossover in order to produce two offspring. This process repeats until the population consists of sufficient number of weight vectors in the population as well as the number of vectors in medipopulation. In the last stage, the weight vectors in the new population are mutated. The mutated vector changes the original vector as a member of the population if its fitness is better than the original one. The set of manageable weight vectors W fulfils the following condition:

$$W = \begin{cases} \omega \in R_n^+, \ 0 \le \omega_j \le uppe_j, \\ j = 1, \ \dots, \ n, \ \sum_{j=1}^{n} \omega_j = 1 \end{cases} \qquad (8)$$

where $uppe_j$ means the upper bound for the weight variable j and ω is denoted as the explanatory variable that refers to the timber volume derived from field sample plots; however, it must be pointed out that the objective as a function of ω is not continuous.

Statistical errors assessment

In order to obtain the level of errors, the pixel-level RMSE and the pixel level average bias using leave-one-out cross-validation (LOOCV) have been computed. Generally, the RMSE represents the sample standard deviation of the differences between observed values, which were data derived from field sample plots and predicted timber volume by k-NN technique. The bias states difference between true value, which originates from ground plots, and value of the parameter being estimated, which was predicted timber volume. The goodness of the estimates was assessed by means of the pixel-level RMSE and bias from following equations, respectively:

$$RMSE = \sqrt{\frac{\sum_{i \in F} (m_i - m_{iv})^2}{n_F}} \qquad (9)$$

$$bias = \frac{\sum_{i \in F} (m_{iv} - m_i)}{n_F} \qquad (10)$$

where m_i stands for the value of the estimated variable, which in this case was timber volume, m_{iv} its estimate on the ground plot i and n_F is the number of the field sample plots.

It must be pointed out that the RMSE method does not accommodate either dependence, and generally error estimation relies mainly on the ground plots data. Statistical error assessment has been done by using R Statistical Software.

Results

The strength of the relationship between timber volume computed from field sample plots and individual Landsat bands were diversified. Regarding to Landsat bands, the strongest relationship between pixel density

and ground plots timber volume obtained band 7 (short-wave infrared 2), which was equal to −0.262. The weakest relationship between independent variable, which was timber volume derived from ground plots, and dependent variable (band 4 − red) was equal to −0.014. The Pearson correlation coefficient between ground plots timber volume and Landsat bands in all cases was negative and equal for band 1 (coastal aerosol) to −0.138, for band 2 (blue) to −0.217, for band 3 (green) to −0.235, for band 5 (near infrared) to −0.152 and for band 6 (short-wave infrared 1) to −0.236. Relatively strong negative relationship, −0.423, was noticed between explanatory variable, which was timber volume derived from ground plots, and explained variable, which was height above the sea level derived from DEM. Positive relationship, 0.204, was found between timber volume derived from field sample plots and slope image which was computed from digital elevation model.

LOOCV without feature selection gave for k-NN timber volume estimation RMSE equal to 192.208, whereas overall bias reached 11.468. RMSE percentage was equal to 51.51, and bias percentage reached 3.07.

Standard deviation obtained from timber volume derived from field sample plots was amounted to 217.98.
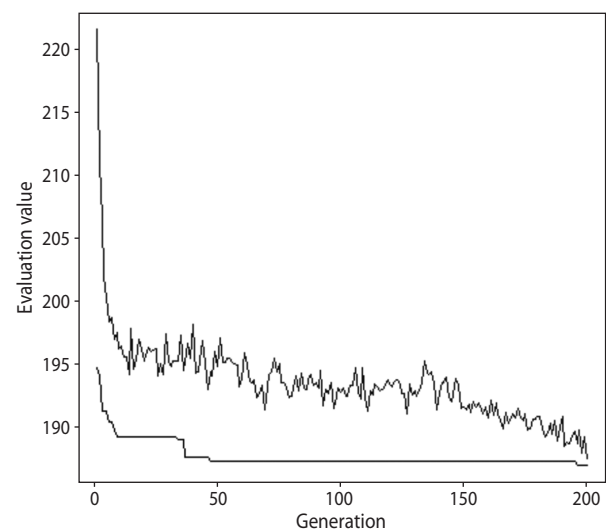

Figure 1. Genetic algorithm output between generation and evaluation value

LOOCV with feature selection used genetic algorithm and its graph output is presented in Figure 1. The graph describes the best score (lower line) and average

score (upper line) for k-NN timber volume estimation over 200 generations (Fig. 1).

RMSE for the k-NN timber volume estimation in this case was equal to 182.576, whereas overall bias reached 0.066. RMSE percentage was equal to 48.92, and bias percentage reached 0.02.
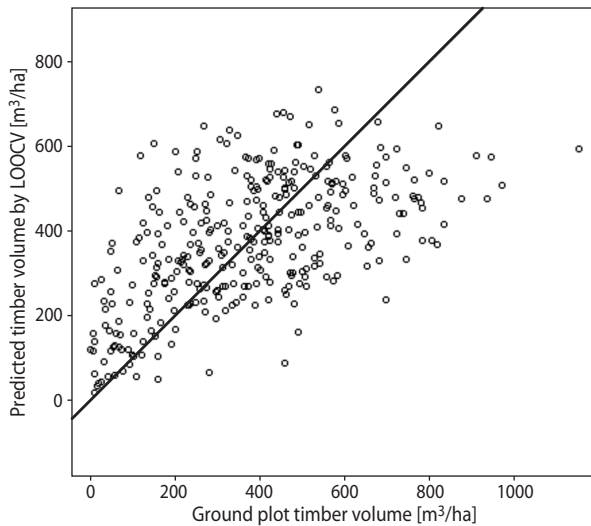

Figure 2. Scatter plot correlation between timber volume derived from field sample plots and predicted timber volume by LOOCV
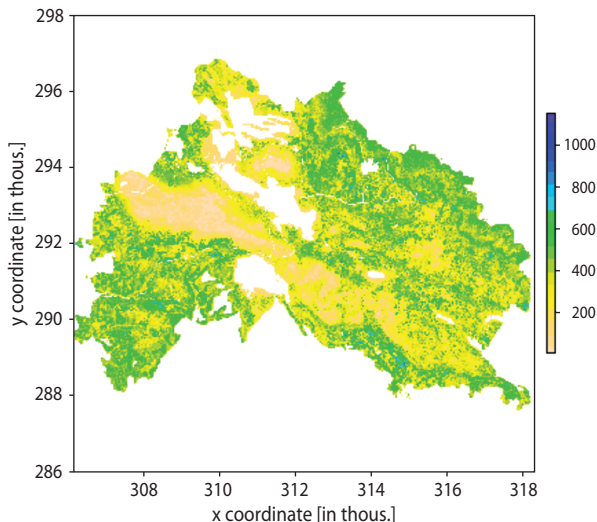

Figure 3. Timber volume map output of growing stock in cubic metres in the study area

Pearson correlation between timber volume derived from field sample plots and independent variable, which

was timber volume estimated by LOOCV method, was equal to 0.556 (Fig. 2).

Total predicted timber volume in the study area reached 2,132,997 m³, which means that average timber volume on the forested area was equal to 366.6 m³/ha. Maximum predicted timber volume reached 1149.5 m³/ha, whereas minimum predicted timber volume was equal to 0.9 m³/ha. Figure 3 presents the final growing stock map of the examined area.

DISCUSSION

MS-NFI is mainly based on k-NN technique, which is generally applicable to all satellite image material where estimation precision depends mostly on the spatial, spectral and radiometric resolutions. In terms of k-NN technique, choosing the appropriate value of k was one of the most decisive issues because this variable always involves a trade-off. Some authors (Nilsson 1997; Tokala et al. 1996) suggested in their publications that the value of k equals to 10–15 is the most suitable for mapping or forest planning purposes. Others indicated that smaller value of k is a better solution (Franco-Lopez at al. 2001; Tomppo 1996; Moeur and Stage 1995). It has been proved that the bias increases and, at the same moment, RMSE decreases when the value of k rises (Franco-Lopez et al. 2001), and therefore, in terms of practical inventories, a compromise is necessary. The most important parameters for the reduction of the estimation error at the pixel level were the value of k, geographic coordinates of the field sample plots and stratification of those plots at the pixel level regarding satellite image, digital map data and DEM. However, it should be noted that the appropriate choice of k depends on several parameters. The first parameter is the number of field sample plots in the study area. The second is the size of those ground plots compared to the pixel level. The third parameter is weighting of the spectral distance in the estimation, and the fourth parameter is density of the training data in spectral space (Katila and Tomppo 2001). Moreover, it is necessary to keep in mind the fact that feature weight selection is an optimisation problem, and generally, it is not possible to find the best solution because the results of different runs give various outputs.

Obtained coefficients concerning Pearson correlation between independent and dependent variables are

quite low. This might be caused by the heavily stocked forest stands – mean timber volume derived from field sample plots was equal to 373 m³/ha. Other might be the result of inaccuracy of field coordinates or by the time difference between satellite image and field data acquisition. The author suspects that incorporation of NDVI might improve discussed relationship between explanatory and explained variables. However, in some MS-NFI trials, the relationship between timber volume and spectral reflectance at the pixel level also indicates a weak or negligible relationship (Trotter at al. 1997). Thus, it should be expected that not only the author struggled with such a problem. The LOOCV method was used to obtain the average RMSE and bias of estimates for pixels. In terms of genetic algorithm graph, it can be noticed that the space between the best score (lower line) and average score (upper line) is relatively large. Much more reliable results and smaller RMSE with overall bias could be obtained if given space was smaller. Nevertheless, received genetic algorithm output is at a satisfactory level in terms of forest management and may have several applications in forest inventory and forest research. Furthermore, it should be pointed out that discussed algorithm has improved considerably overall bias and slightly over RMSE. In terms of Person correlation between field sample plots timber volume and k-NN predicted timber volume, the given coefficient could be bigger. However, obtained result is also at a satisfactory level. It must be emphasised that the presented method is under continuous development, and in the near feature, it can be expected that used algorithms will be optimised better, which may improve the reliability of the results of the forest inventory. All detailed calculations can be found in author's master thesis (Pachana 2015), which can be found in libraries of the Warsaw University of Life Sciences and Eberswalde University for Sustainable Development.

The most serious limitation of the application of MS-NFI method is associated with the availability of satellite images. It was quite difficult to find such images which would be suitable for presented calculations. Good weather without clouds and quality disturbances were the most important. In May 2003, Landsat 7 scan line corrector (ETM+) has been damaged, and after this period, most images could not be used. It should also be pointed out that MS-NFI method is relatively fast and cheap in relation to methods involving light detection

and ranging (LIDAR) measurements. The first advantage of the MS-NFI is the fact that the satellite image can be obtained free of charge and one image covers an area equal to 170 km north-south by 185 km east-west. Such possibility was created by Landsat program, which provides satellite images of the whole Earth. Moreover, USGS supplies free access to DEM data. The overall expenditures associated with discussed method are related to data collection from field sample plots and creation of digital map data. In terms of LIDAR measurements, the overall costs are much higher because it is necessary to charter a plane provided with essential equipment, which is relatively expensive. Furthermore, the acquisition and processing of LIDAR data also entail additional costs and are more time consuming, whilst obtained results regarding to MS-NFI are quite similar. Results from the previous research project based on LIDAR measurements at the same area of interest are as follows: the average timber volume per hectare reached 369.8 m³ and total timber volume for study area was equal to 2,107,600 m³ (Miścicki et al. 2012). However, mean timber volume obtained from LIDAR measurements was a bit higher than that in MS-NFI method, but total volume was lower than the author received. Therefore, it can be assumed that the size of the study area was slightly different or forest stands belonging to private owners have been pulled into calculations. It must be pointed out that private forests represented relatively small share of the overall stands in the examined area. Furthermore, it can be noticed that obtained timber volume maps slightly differ, especially with regard to the places with higher growing stock.

It goes without saying that remote sensing nowadays is a very promising and forward looking technology. Under the budget constraints, the balance between sufficient quantity of field sample plots should be established in order to provide a good co-variation between the remote sensing data and the key variables. The spatial autocorrelation in the forest variables and in the remote sensing data should be taken into consideration in this optimisation process (Wang et al. 2001.) However, it should be pointed out that remote sensing has various factors that reduce the accuracy, and the user should be aware of their occurrence. One of them is geometric accuracy of field sample plots and satellite data. If those two variables does not fit to each other, the results will be wrong. Another issue is radiometric

quality of the satellite image, including atmosphere, because without good quality, the outputs would also be inaccurate. Temporal proximity between ground measurements and satellite image is also a significant issue, and the user should strive to acquire data in similar time periods. Topographic effect from shadow also determines an important issue and that is the reason why DEM should be used in order to enhance reliability of the final results.

MS-NFI method can be applied in various forestry fields but, above all, in modern forest management planning. The main aim of discussed method is obtaining forest inventory results for units smaller than regions, for example, forest district. The size of Stołowe Mountains National Park is a good example, as it is possible to receive large-area timber volume estimation for units equal to typical forest precinct. Currently, in Poland, there are inventory methods that can use the results of MS-NFI in a direct way (Miścicki 2000; Miścicki 2009). The induction of the requirement of using remote sensing data to the Polish Forest Management Instruction is a condition for the application of the presented method. Thanks to the semi-automatic and objective inventory method of certain elements, MS-NFI can support the development of conservation plans in National Parks as well as NATURA 2000 and Corine Land Cover areas. This allows for more accurate characterisation of stands features such as spatial distribution of vegetation as well as estimation of forest biomass. Particularly valuable may be the use of presented solutions in areas inaccessible or dangerous for humans. Another area where MS-NFI could have practical application is fire protection support of the forest areas. This could be achieved by the possibility of determining the overall amount of burning material, forest biomass as well as stands structure. The main problem in present forestry is based on the right combination of modern technologies and human work in such a way that it could bring benefits for both good forest condition as well as respect of the human role.

Conclusions

Research results and their comparison with other research outputs allowed to state that k-NN technique as well as genetic algorithm used in MS-NFI are very promising and powerful tools with regard to the timber volume estimation. Furthermore, k-NN technique is a feasible and an efficient method for processing TM images in order to obtain timber volume estimation. What is more, k-NN is a versatile technique with potential of combining different sources of information and flexibility in incorporation of ancillary data. It must be pointed out that Pearson correlation coefficient between timber volume derived from field sample plots and Landsat bands as well as slope image indicates negligible or weak linear relationship. Nevertheless, field sample plots timber volume indicates strong positive linear relationship with predicted timber volume derived from k-NN technique. Important issue is that genetic algorithm improved results reliability with respect to RMSE and overall bias; however, positive bias derived from k-NN technique in this case leads to timber volume overestimation. Another important aspect of obtaining reliable results is having enough field sample plots in order to cover all variations of the tree dimensions as well as stand density for each canopy type. It must be pointed out that given timber volume results are similar to previous LIDAR measurements, which were done in the same study area, and results derived from k-NN technique may be expected to satisfy national forest inventory and standard analysis.

Acknowledgements

References

Borusiewicz B. 2009. Operat ochrony ekosystemów leśnych Parku Narodowego Gór Stołowych. Taksus SI, Warszawa–Gorzów Wielkopolski.

Franco-Lopez H., Alan R.E., Bauer M.E. 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-Nearest Neighbors method. *Remote Sensing of Environment*, 77, 251–274.

Przemko Pachana

Halme M., Tomppo E. 2001. Improving the accuracy of multi-source forest inventory estimates to reducing plot location error – a multi-criteria approach. *Remote Sensing of Environment*, 78 (3), 321–327.

Katila M., Heikkinen J., Tomppo E. 2000. Calibration of small-area estimates for map errors in multi-source forest inventory. *Canadian Journal of Forest Research*, 30, 1329–1339.

Katila M., Tomppo E. 2001. Selecting estimation parameters for the Finnish multisource National Forest Inventory. *Remote Sensing of Environment*, 76 (1), 16–32.

Katila M., Tomppo E. 2002. Stratification by ancillary data in multisource forest inventories employing k-nearest-neighbor estimation. *Canadian Journal of Forest Research*, 32 (9), 1548–1561.

Koukal T., Suppon F., Schneider W. 2005. The impact of relative radiometric calibration on the accuracy of k-NN predictions of forest attributes. *ForestSat Report*, 8, 17–21.

Maselli F., Chirici G., Bottati L., Corona P., Marchetti M. 2005. Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. *International Journal of Remote Sensing*, 26, 3781–3796.

McRoberts E., Nelson M.D., Wendt D.G. 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique. *Remote Sensing of Environment*, 82, 457–468.

Miścicki S. 2000. Kombinowana dwufazowa inwentaryzacja lasów nizinnych z wykorzystaniem zdjęć lotniczych i stałych kontrolnych powierzchni próbnych. Fundacja „Rozwój SGGW", Warszawa.

Miścicki S. 2009. Pomiar zapasu grubizny z wykorzystaniem zdjęć lotniczych. *Sylwan*, 153 (6), 373–385.

Miścicki S., Będkowski K., Stereńczak K., Zasada M. 2012. Opracowanie metody pomiaru zasobów leśnych z wykorzystaniem lotniczego skaningu laserowego na przykładzie terenu górskiego objętego ochroną. Projekt badawczy finansowany przez Ministerstwo Nauki i Szkolnictwa Wyższego N N309 111937.

Moeur M., Stage A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science*, 41, 337–359.

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

Pachana P. 2015. Forest stands volume estimation by using Finnish Multi-Source National Forest Inventory. Master thesis, Eberswalde University for Sustainable Development, Faculty of Forest and Environment, Warsaw University of Life Sciences, Faculty of Forestry, 47–64.

Pakkala T., Hanski I., Tomppo E. 2002. Disturbance dynamics in boreal forest: Defining the ecological basis of restoration and management of biodiversity. *Silva Fennica*, 36 (1), 279–288.

Reese H., Nilsson M., Sandstrom P., Olsson H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture*, 37 (1), 37–55.

Tokala T., Pitkänen J., Partinen S., Muinonen E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing*, 17, 2333–2351.

Tomppo E. 1988. Standwise forest variate estimation by means of satellite images. University of Helsinki, Department of Forest Mensuration and Management, Research Notes 21, 103–111.

Tomppo E. 1990. Satellite image-based National Forest Inventory of Finland. *The Photogrammetric Journal of Finland*, 12 (1), 115–120.

Tomppo E. 1996. Multi-source National Forest Inventory of Finland – new thrusts in forest inventory. Proceedings of the IUFRO XX World Congress, 6–12 August 1995, Tampere, Finland, Volume 1, 27–41.

Tomppo E., Korhonen K.T., Heikkinen J., Yli-Kojola H. 2001. Multi-source inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. *Silva Fennica*, 35 (3), 309–328.

Trotter C.M., Dymond J.R., Goulding C.J. 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. *International Journal of Remote Sensing*, 18, 2209–2223.

Wang G., Gertner G., Xiao X., Wente S., Anderson A. B. 2001. Appropriate plot size and spatial resolution for mapping multiple vegetation types. *Photogrammetric Engineering and Remote Sensing*, 67 (5), 575–584.