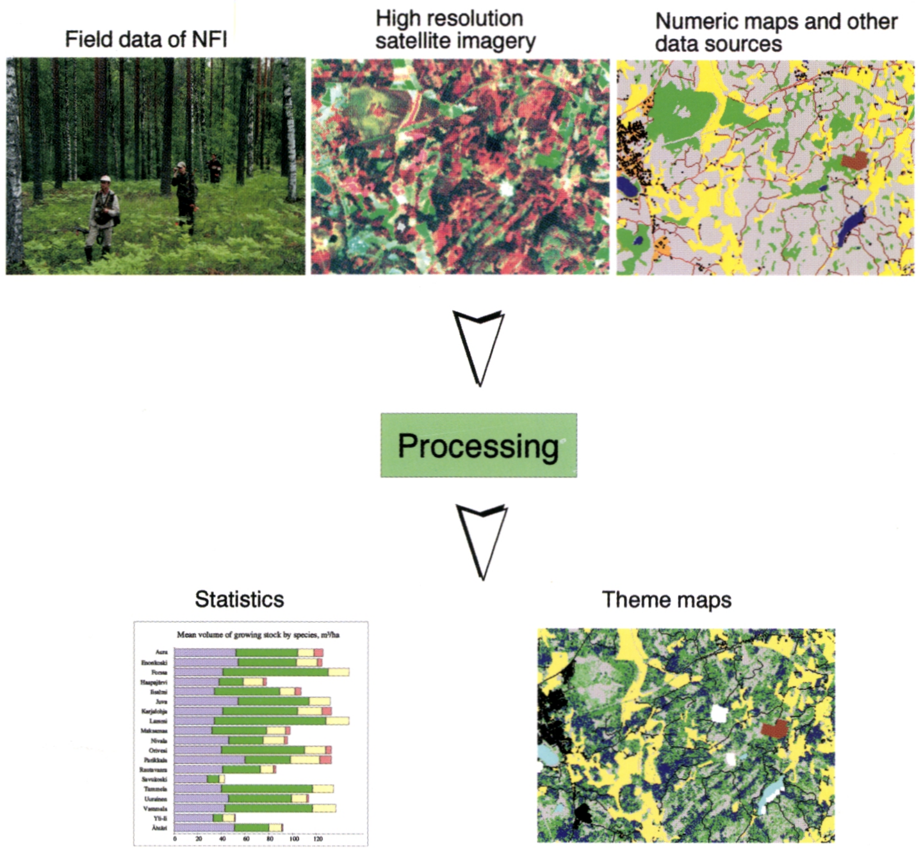
METSÄNTUTKIMUSLAITOKSEN TIEDONANTOJA 910, 2004
THE FINNISH FOREST RESEARCH INSTITUTE, RESEARCH PAPERS 910, 2004



Controlling the estimation errors in the Finnish multisource National Forest Inventory

Matti Katila

VANTAAN TUTKIMUSKESKUS–VANTAA RESEARCH CENTRE

METLA

METSÄNTUTKIMUSLAITOKSEN TIEDONANTOJA 910, 2004
THE FINNISH FOREST RESEARCH INSTITUTE, RESEARCH PAPERS 910, 2004

Controlling the estimation errors in the Finnish multisource National Forest Inventory

Matti Katila

Academic dissertation

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public examination in Auditorium 2 (sali 2), Viikki Infocentre, Viikinkaari 11, on 6 February 2004, at 12 o'clock noon.

VANTAAN TUTKIMUSKESKUS – VANTAA RESEARCH CENTRE

Katila, Matti. 2004. Controlling the estimation errors in the Finnish multisource National Forest Inventory. Metsäntutkimuslaitoksen tiedonantoja 910. Finnish Forest Research Institute, Research Papers 910. ISBN 951-40-1904-0, ISSN 0358-4283. 36 p. + original papers.

Publisher: Finnish Forest Research Institute

Approved: Jari Hynynen, Research Director, January 5, 2004

Author's address:
Matti Katila, Finnish Forest Research Institute,
Unioninkatu 40 A, FIN - 00170 Helsinki.
Tel: +358 10 211 2159. Fax: +358 10 211 2104,
e-mail: matti.katila@metla.fi

Academic dissertation at the Department of Forest Resource Management, University of Helsinki

Supervised by: Prof. Erkki Tomppo
Finnish Forest Research Institute
National Forest Inventory

Ph.D. Juha Heikkinen
Finnish Forest Research Institute
National Forest Inventory

Pre-examiners: Prof. Michael Köhl
Dresden University of Technology
Chair of Forest Biometrics and Computer Sciences

Ronald McRoberts
Mathematical Statistician
North Central Research Station
USDA Forest Service

Opponent: Prof. Risto Päivinen
Director
European Forest Institute

Abstract

This study concentrates on the assessment of the error in the estimates of Finnish multisource National Forest Inventory (MS-NFI) and its minimisation, as well as for the k–nearest neighbour method (k–NN). The MS-NFI utilises optical area satellite images, mainly Landsat TM and ETM+, and digital maps, in addition to field plot data, to produce geo-referenced information, thematic maps and small-area statistics. The non-parametric k–NN estimation method is used in the estimation of forest variables for single pixels and to define weights of field plots to a particular computation unit, e.g. a municipality. First, the estimation parameters that are optimal for the objectives of MS-NFI were achieved by examining the prediction error at the pixel level. Secondly, potential variables, covariates or other exogenous variables, what might explain the residual variation in the k–NN estimates were studied. Finally, two methods were presented aimed at reducing the effect of map errors on MS-NFI small-area estimates.

The selection of the estimation parameters was examined for four study areas that covered a greater part of the variation found in the Finnish forests. The error estimates were obtained by leave-one-out cross-validation. The most important parameters for minimising the estimation error of the total volume and volume by tree species at pixel level were the value of k, the geographical horizontal reference area (HRA) radius used to select the training data and the stratification of the field plot pixels, and training data employing the site class map. With the sampling intensity in the 8th and 9th Finnish National Forest Inventory, a geographical HRA with a radius of 40–50 km was found to be optimal for the total volume estimates and for volumes by tree species on the mineral land map stratum. For the peatland stratum, a wider reference area, 60–90 km, was required.

The main sources of error in the Finnish MS-NFI are considered to be the representativeness of the field sample with respect to the estimation problem, the low dynamic range of spectral channel values on forestry land (FRYL) on high resolution optical satellite data, the small size of the NFI field plots compared to the pixel size in image data and the locational errors in the image and field plot data. The first principal component (PC1) of the Landsat TM or ETM+ channel values of the field plot pixel was strongly related to the residual variation in the volume and basal area estimates. The residual variances of field plot volume were regressed against PC1 and the model was used to remove the trend component of PC1 from

the residuals, but the random error component still remained high in the residuals.

A calibration method was introduced to reduce the map errors on MS-NFI small-area estimates. The method was based on large-area estimates of map errors; i.e. the confusion matrix between land use classes of the field sample plots and corresponding map information. A method to compute the calibrated field plot weights was also presented. These weights were in turn used to calculate the small-area estimates. In the second method, the k-NN estimation was carried out separately within each map strata employing all the field plots from all the land use classes within each stratum.

Comparisons were made between the aggregates of MS-NFI small-area estimates from the two methods and field inventory estimates at the region level in order to determine the total amount of correction, and for the subregions (groups of municipalities) to detect the possible bias in the small-area estimates. Although quite different in nature, both methods corrected the bias in the FRYL area estimates. The FRYL estimates of the calibrated MS-NFI are consistent with post-stratified estimates at the region level. When compared to the field inventory based estimates of tree species volumes for subgroups of municipalities (1738–4238 km^2), the stratified MS-NFI performed better than the original MS-NFI and calibrated MS-NFI. Some of the estimates from the two latter methods differed by more than two standard errors from the field inventory estimates in the subregions of the test data.

The parameter selection methods and the small-area estimation map error correction methods, together with the field inventory estimates and their standard errors, provide a method for reducing the estimation error and a reference of the accuracy of the MS-NFI results. However, if there is a significant systematic error in the small-area estimates of a certain subregion, it may not be possible to remove the error by varying the estimation parameters. Other methods or auxiliar data is needed to do this.

Keywords: multisource forest inventory, k–nearest neighbours, cross-validation, Landsat TM and ETM+, stratification, training data selection, prediction error, statistical calibration

Preface

After working several years with the operative multisource National Forest Inventory, I got the opportunity to concentrate on studying the estimation problems in the multisource method, after admission to the graduate school "Forests in Geographical Information Systems" in the University of Helsinki, in 1998. This work was mainly funded by the Finnish Foresters Foundation and the Finnish Forest Research Institute. I am grateful for their financial support.

I am also grateful for being able to work in the NFI-team led by my supervisor and co-author, Prof. Erkki Tomppo. His personal example of commitment to research work and experienced advice has been invaluable. I am also particularly indebted to my second supervisor and co-author, Dr. Juha Heikkinen for his guidance in the statistical problems, as well as in all practical problems but, most of all, for inspiring discussions of research problems. I would like to thank my pre-examiners, Prof. Michael Köhl and Dr. Ronald McRoberts for their valuable comments.

I would like to acknowledge Prof. Jouko Laasasenaho and Dr. Markus Holopainen for their support and management of the graduate school. I have enjoyed their company and that of the fellow students in the graduate school during these years. I also wish to thank Emeritus Prof. Simo Poso and Prof. Annika Kangas from the Department of Forest Resource Management for their support and valuable comments to my studies.

I have enjoyed discussions with Dr. Helena Henttonen, my first employer in scientific work, concerning statistical problems in forest inventories. Mr. Kai Mäkisara gave me advice on remote sensing and computer science matters, and provided valuable comments on the articles. I also wish to thank Dr. Jari Varjo for encouraging me to apply to the GIS school membership and support as a member of the executive board of the graduate school. With Mr. Jouni Peräsaari, with whom I shared the room during this work, I have had useful discussions on problems in the operative multisource inventory. Mr. Arto Ahola and Mr. Antti Ihalainen have helped me in the application of NFI field data. I am grateful to Dr. Ashley Selby for editing the English language both in compilation and in the articles and Ms. Anna-Kaisu Korhonen for her technical assistance. I wish to extend my gratitude for the whole NFI-team for their support and good working spirit. I also wish to thank the numerous persons not mentioned so far who have contributed to the completion of the work.

Finally, I wish to express my gratitude to my father Hannu, and my mother Eeva for supporting my work.

Helsinki, January 2004 Matti Katila

Contents

List of separate studies

This dissertation includes the following separate studies, which are referred to by roman numerals in the text as follows:

I	Katila M., and Tomppo E. 2001. Selecting estimation parameters for the Finnish multisource National Forest Inventory. Remote Sens. Environ. 76: 16–32.
II	Katila M. 2004. Error variatons at the pixel level in the k–nearest neighbour estimates of the Finnish multisource National Forest Inventory. Manuscript.
III	Katila M., Heikkinen J., and Tomppo E. 2000. Calibration of small-area estimates for map errors in multisource forest inventory. Can. J. For. Res. 30: 1329–1339.
IV	Katila, M., and Tomppo, E. 2002 Stratification by ancillary data in multisource forest inventories employing k–nearest neighbour estimation. Can. J. For. Res. 32:1548–1561.

The articles I, III and IV are reprinted with permission.

Fields of responsibility

In substudy I, Katila designed most of the tests, carried out the programming and drew the conclusions. Tomppo proposed some of the tests concerning the variation in the training data. He also participated in the writing of the manuscript and drawing the conclusions. In substudy IV, Katila designed the stratification in the MS-NFI method, carried out the analysis and wrote the article together with Tomppo. Tomppo carried out the necessary changes in the k–NN estimation program. In substudy III, Katila planned the realisation of calibration in the multisource method, carried out the analysis, drew the conclusions and wrote the manuscript together with Heikkinen and Tomppo. Heikkinen formulated the calibration method to be applied in the multisource inventory and wrote the statistical background of the estimation of standard errors (Appendix A). Tomppo carried out the necessary changes in the k–NN estimation program.

1. Introduction

1.1. The objectives of national forest inventories

There are three main types of forest inventories: the operational, the management and the national forest inventories (Cunia 1978). The objective of national forest inventories is to produce statistically unbiased, reliable forest resource information for large areas for strategic planning, primarily by decision makers. Estimates of both current values and rates of changes of forest resources are required (Cunia 1978). Periodic national forest inventories can provide information on trends in the state of forests (Lund 1993). The estimates are required, e.g. of the forest resources, growing stock, growth, health of forests and, increasingly, of the biodiversity in the forests. The national forest inventory methods should be statistically valid, cost-efficient and flexible (Cunia 1978).

In recent years, there has been a growing interest in obtaining national forest inventory results for smaller areas than had previously been possible based on field samples only, e.g. for municipalities and even for single forest stands, for forest planning, timber procurement and biodiversity assessment purposes (Tomppo 1987, 1991, Schreuder et al. 1993, Kangas 1996, Tokola & Heikkilä 1997, Nilsson 1997, Tomppo et al. 1998, Franco-Lopez et al. 2001). The remote sensing data from airborne and spaceborne sensors has been the key to a more efficient use of forest inventory data. Some of the advantages of remote sensing data are that they offer a synoptic view of the study area, the data can be obtained rapidly for large areas and they can be processed digitally (Schreuder et al. 1993). Traditionally, the remote sensing data has been used as a part of the sampling design, in order to decrease the cost of field work rather than to try to obtain results for significantly smaller areas than normally used in strategic forest inventories (Loetsch & Haller 1973). The classification based on remote sensing data has been used in stratified sampling (Tomppo et al. 2001), multistage-sampling (Schreuder et al. 1993) and multiphase-sampling (Poso 1972, Schreuder et al. 1995). The post-stratification may also provide an effective means to decrease the variance in the estimates after the actual sampling (McRoberts et al. 2002). The concept of multisource forest inventory employing remote sensing data and digital map data has been introduced to forest inventories. One prerequisite for a multisource inventory method is that it should be possible to estimate all the variables measured in the field (Kilkki & Päivinen 1987).

1.2. Multisource national forest inventory

Multisource national forest inventories employ various sources of geo-referenced data, in addition to field inventory data, to obtain more reliable estimates or estimates for smaller areas than when employing the pure field plot data only. Holmgren & Thuresson (1998) list the following types of forest inventory applications employing remote sensing data: land cover classification of timber types, estimation of the forest variables for forest management planning purposes, segmentation to determine stand and other boundaries, landscape ecology analysis and large-scale forest inventories. Continuous variables, such as stand volume, volume by tree species, age and mean breast height diameters have been estimated for forest management planning purposes employing optical area remote sensing data and field plot data. Sampling based methods, parametric and non-parametric regression methods and neural networks have been used, occasionally in conjunction with segmentation techniques (Poso et al. 1987, Tomppo 1987, 1991, Tokola et al. 1996, Hagner 1997, Mäkelä & Pekkarinen 2001). In small-area estimations, indirect estimation methods are used and support is obtained from similar computation units by applying methods to link the field plot data and the auxiliary data (Schreuder et al. 1993). Non-parametric regression has been used for small-area estimation in the Scandinavian countries and the United States (Tomppo 1991, Tokola et al. 1996, Nilsson 1997, Gjertsen et al. 2000, Franco-Lopez et al. 2001). The non-parametric regression methods are relatively easy to use and require no assumptions about the shape of the model.

In multisource forest inventories, both airborne and spaceborne imagery from active or passive sensors may be employed, although optical area remote sensing data has mainly been employed. Aerial photography has demonstrated its applicability for both large area and management inventories (Poso 1972, Loetsch & Haller 1973, Schreuder et al. 1993). Airborne laser instrument and radar data applications in the mapping of forests are still at the development stage (Hyyppä et al. 1997, Naesset 2002).

The earth observation satellites provide continuous image data for large areas (Campbell 1996) and the increase in the number of satellites may help to overcome the problem of cloudiness in the image data. The high resolution image data from Landsat and SPOT satellite programs have been used frequently in large-area land-use or land-cover classification, as well as for multisource forest inventories (Campbell 1996, Eisele 1997, Nilsson 1997, Tomppo et al. 1998, Franco-Lopez et al. 2001). The medium resolution satellites have shown potential in estimating volume and biomass, by covering large areas at low cost (Tomppo et al. 2002). The radar satellite imagery (SAR) has yielded less accurate forest parameter estimates than high resolution optical satellite data (Tomppo et al. 1996). The spectral and spatial resolution of the remote sensing data has been enhanced in multisource forest inventories by employing multitemporal or multiple instrument image data (Poso et al. 1999, McRoberts et al. 2002). New, very high resolution satellite data with 1–5 m pixel size is now available, but it is costly and requires new estimation methods due to the scale of the target, i.e. forest stands and trees (Woodcock & Strahler 1987, Hyppänen 1996, Pekkarinen 2002).

Topographic databases, digital elevation models and other map data are readily available in digital format (National Land Survey of Finland 1996). However, the map data may include location errors, it may be out-of-date and the attributes may not correspond to the ones used in the multisource forest inventory. Despite the possible inconsistencies between map data and remote sensing data, the map data can be used to improve an estimation either as ancillary information or together with remote sensing data in the analysis (Wilkinson 1996).

The Finnish multisource National Forest Inventory (MS-NFI) utilises optical area satellite images and digital maps, in addition to field plot data, to produce georeferenced information, thematic maps and small-area statistics. A non-parametric k–nearest neighbour method (k–NN) is used in the estimation of forest variables for single pixels and to define weights of field plots to a particular computation unit, e.g. a municipality (Tomppo 1991). One advantage of the k–NN method is that all the inventory variables can be estimated simultaneously. Field data from surrounding computation units (municipalities), in addition to the unit itself, are utilised when estimating results for the particular unit. It is therefore possible to obtain estimates for smaller areas than would be the case when employing sparse field data only (Kilkki & Päivinen 1987, Tomppo 1991).

1.3. Aim of the study

This study concentrates on the assessment and minimising of the error in the Finnish MS-NFI and the k–NN estimation method. The errors are studied at the pixel level, for small areas, i.e. municipalities and at the region level. First, the different sources of error and their significance in the MS-NFI estimation are studied. The general outlines of small-area estimation and the non-parametric regression methods are discussed and the application of these methods in the MS-NFI is introduced.

In the k–NN estimation, the overall error is minimised by tuning the estimation parameters. Leave-one-out cross-validation, a resampling technique, is used to guide the parameter selection at the pixel level. These techniques are applied to choose the parameters for the Finnish MS-NFI. The remaining variation in the error is studied and potential explanatory variables are sought to model the prediction error.

Two methods are developed to decrease the error in the small-area estimates caused by the forestry land (FRYL) area delineation based on erroneous map data. FRYL consists of forest land, other wooded land and waste land. A statistical calibration method posterior to the k–NN estimation is compared to the k–NN estimation applied by map strata. The MS-NFI small-area estimates are validated by groups of municipalities –subregions– and at the region level against the field inventory based key forest variable estimates and their standard errors.

2. Error sources in multisource national forest inventory

In multisource forest inventories, the number of errors increase with the number of data sources. Explanatory models or standardised rules must be applied at various phases of data production (Freden & Gordon 1983, Tomppo et al. 1997, Burrough & McDonnell 1998), e.g. a definition of land use classes, volume models for sample trees and calibration equations for the satellite imagery exo-atmospheric radiances. Various types of error taxonomies can be used to describe the error structure of the MS-NFI. The error components of a forest inventory are measurement errors, sampling errors and model estimation errors (Cunia 1965). The accuracy of the spatial data can be grouped into thematic, positional and temporal accuracy (Burrough & McDonnell 1998) or thematic and non-thematic errors (Foody 2002). The measurement errors in remote sensing procedures can be divided into errors in the measurement of field data, errors in the measurement of remote sensing data, and the misregistration in space or time between field variables and remote sensing variables (Curran & Hay 1986). The main sources of error in the Finnish MS-NFI are considered to be the representativeness of the field sample with respect to the estimation problem, the low dynamic range of spectral channel values on FRYL on high resolution optical satellite data, the small size of the NFI field plots compared to the pixel size in image data and the locational errors in the image and field plot data (II; Halme & Tomppo 2001). In the Table 1, several sources of error in the MS-NFI data are presented. They are grouped according to spatial data and forest inventory error types. Some estimates of error magnitudes are given, based on the literature and practical experiences in the Finnish MS-NFI.

Table 1. Sources of error in the data for Finnish MS-NFI employing Landsat Thematic Mapper image data

Data source	Error source	Thematic: Sampling	Thematic: Measurement	Thematic: Model	Positional	Temporal
Field plot	Intensity of sample and size of field plot	Does not cover the variation in the field and in the image	Field plot size (max. 492 m^2) < Instant Field of view of instrument (IFOV) (900 m^2) < Effective IFOV		Location error, RMSE 20 m (Halme & Tomppo 2001)	Field plot measurement and image acquisition date
	Tally trees		diameter at breast height measurement	Generalising sample tree variables for tally trees		
	Sample trees		Variables measured for volume models	Volume equations		
RS instrument	Spatial		Sensor sampling function causing spatial bias within pixels (Bastin et al. 2000)		Landsat 5 TM interband location error 0.2-0.5 pixels	
	Spectral					
	Radiometric		Signal to noise ratio	Precalibration equations		
	Viewing		Varying irradiance due to latitude Varying irradiance due to instrument viewing angle			
	Atmosphere		Scattering			
	Target		Varying irradiance due to topography	Reflection model for topographic correction	Ground altitude variation	Varying irradiance due to seasonal effects (multitemporal images)
	Processing				Image system correction 90 % of errors less than 15 m (Landsat 5 TM) (Freden & Gordon 1983), repeated image lines (Bastin et al. 2000), geo-coordinate rectification model RMSE 15–20 m (Landsat 5 TM) (Tomppo 1996)	
Topographic map data	(Burrough & McDonnell 1998)		Correspondence of map attributes to field plot data	Map conversions, generalisation	Accuracy and precision	Field work date
Digital elevation model	(Burrough & McDonnell 1998)		Elevation curve density in the topographic map	Triangulation method accuracy		

3. Small-area estimation and k–nearest neighbour estimation in multisource national forest inventory

3.1. Small-area estimation

Small-area estimation refers to the calculation of statistics for a small subpopulation (domain) within a large geographical area. Sample sizes are often too small to provide reliable direct estimators for a small area (Rao 1998). Small-area estimates gain support from related areas that are nearby or similar according to auxiliary information (Schreuder et al. 1993). The indirect estimation methods are grouped into estimators based on implicit models and model-based estimators (Rao 1998). The former group contains a synthetic estimator, for which it is assumed that the small areas have the same characteristics as the large areas (Gonzalez 1973). A reliable direct estimator for a large area is used to derive an estimator for a small area (Rao 1998). In the model based methods, either non-parametric or parametric methods are applied to the auxiliary information in order to derive the small-area estimates. Because the small-area estimators are, at least partially, model-based, the estimates obtained are usually biased. However, the biased estimator can still be useful if the mean square error (MSE) of the estimator is smaller than that of the unbiased estimator (Kangas 1996).

Kangas (1996) employed several parametric and non-parametric models in a small-area estimation of municipality level volume estimates using NFI field plot data and their coordinates as auxiliary data. The mixed model estimator was found to be the most reliable of the tested models. In general, models that can be corrected for their observed residuals were recommended: mixed models, the Mandallaz estimator and kriging estimator (Kangas 1996). The area interpretation of weights for field plots used in a small-area estimation for a particular computation unit is useful, e.g. for management planning systems. To obtain this interpretation, all the weights must be positive, the weights must be same for all the target variables and add up to the total area of the calculation unit (Tomppo 1996, Lappi 2001). The weighting approach retains the natural covariation between the field plot variables within each field plot.

In the multisource inventories, non-parametric regression methods have been widely used to estimate the forest variables by associating the field plots directly to the pixels of satellite image data in order to produce thematic maps (Kilkki & Päivinen 1987, Tomppo 1991, Nilsson 1997, Franco-Lopez et al. 2001). Area interpretation is used at least in the reference sample plot method (Kilkki & Päivinen 1987) and Finnish MS-NFI (Tomppo 1991). Lappi (2001) argues that the chosen nearest neighbour field plots may not add up to statistically unbiased or statistically optimal estimates for the region to be estimated. He presented a small-area calibration estimator that minimises the sum of distances between prior and posterior weights of field plots for a distance function while respecting the calibration equation based on spectral values of satellite image. A spatial variogram model was applied for calculating the variances of the calibration estimator.

The bias in the Finnish MS-NFI small-area estimators has been assessed by applying the standard error estimates of the field inventory estimates at the region and subregion level (III), because an explicit error variance estimate is not available. Some small-area estimation methods have estimators for variances. The resampling methods are useful in the estimation of the error for small areas, but unlike in the kriging methods, it is difficult to take into account the possible autocorrelations in the data (Davison & Hinkley 1997).

3.2. k–nearest neighbour estimation method

Nonparametric regression methods are a collection of techniques for fitting a curve when there is little a priori knowledge about the shape of the true function, and the form of the function is not restricted. These methods are applied in exploratory analysis and, increasingly, as stand-alone techniques (Altman 1992, Linton & Härdle 1998). Nonparametric regression methods can be considered to belong to the group of generalised additive models (Hastie & Tibshirani 1997). The general formula for nonparametric regression for a simple bivariate dataset $(X_i, Y_i)_{i=1}^n$ is

$$Y_i = m(X_i) + \epsilon_i, \qquad i = 1, \ldots, n, \tag{1}$$

where ϵ_i is a random error independent over observations, $\mathrm{E}(\epsilon_i|X_i = x) = 0$ and $\mathrm{Var}(\epsilon_i|X_i = x) = \sigma^2(x)$. $m(\cdot)$ is the regression function of Y on X and m is estimated at the group of observations covering some subset $\mathcal{X}$ in support of X. It is a linear smoother of the form $\sum_{i=1}^n W_{ni}(x)Y_i$ for the weights $W_{ni}(x)_{i=1}^n$ depending only on $X_1, \ldots, X_n$ (Linton & Härdle 1998). The kernel and the k–nearest-neighbour estimators are among the most common smoothers in forestry applications.

The kernel estimate is a weighted average of the response variable in a fixed neighbourhood, bandwidth h, of x; the Nadaraya-Watson kernel estimate is

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} K_h(x - X_i)Y_i}{\sum_{i=1}^{n} K_h(x - X_i)}, \tag{2}$$

where $K(\cdot)$ is any kernel function. The k–NN estimate is a weighted average of the response variables in a varying neighbourhood, defined by those X that are among the k–NNs of a point x

$$\hat{m}_k(x) = \frac{\sum_{i \in \mathcal{N}(x)} Y_i}{k}, \tag{3}$$

where $\mathcal{N}(x)$ is the set of indices of the k–NNs of x. Eq. 3 is comparable to a kernel smoother applying a uniform kernel and a variable bandwidth h (Linton & Härdle 1998).

The NN algorithms have been extensively used in the statistical pattern recognition since the paper by Fix & Hodges (1951) in which they presented the simple nearest neighbour classifier. The pattern recognition system typically consists of a feature extraction and classification phase. Dasarathy (1991) reviews several studies concerning the classifier risks for finite and infinite samples, the asymptotic performance of the classifiers, selecting the training data, choice of k and metrics. The nearest neighbour distances are also used in geostatistics (Bailey & Gatrell 1995). Apart from the multisource inventories, the k–NN method and kernel methods have been used in other fields of forest inventory, such as basal area diameter distribution estimation (Haara et al. 1997, Maltamo & Kangas 1998), generalising sample tree data (Korhonen & Kangas 1997) and generalising detailed stand characteristics from stand databases employing less accurate stand information (Moeur & Stage 1995, Malinen 2003).

The choice of k affects the shape of the regression function; when k increases a smoother fit is obtained with a smaller variance but larger local bias for $\hat{m}_k(x)$ with given x and a fixed sample size (Altman 1992). The mean squared error (MSE) is a commonly applied optimality criterion for error minimisation. The quadratic loss by MSE can be studied at a single point x or globally (Linton & Härdle 1998), which may alter the selected smoothing parameter k.

The question may arise, how to select k as the sample size n increases? In pattern recognition, the k–NN classifier has the asymptotic property that when a sequence of k_n satisfies $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$, the classification error approaches the optimal rate of Bayes decision rule for discrete variables (Stone 1977, Keller et al. 1985). However, in practical problems with moderate n, the optimal selection depends largely on the distributions of the variables (X, Y) (Kulkarni et al. 1998).

The k–NN estimates are potentially biased if the true function has substantial curvature (Altman 1992); e.g. the convex relationship between satellite digital numbers (DN) and field plot volume should yield a positive bias in the estimates (Nilsson 1997). The weighting of the neighbours can be used to decrease the bias (Altman 1992).

Resampling techniques, the most popular of them being cross-validation, are frequently applied to the error quantification and parameter selection for classification and estimation problems. Bootstrap methods can be used to estimate the generalisation error and also confidence limits. Efron & Tibshirani (1997) introduced the .632 bootstrap method and improved .632+ bootstrap method for classification problems. These are smoothed versions of cross-validation, partially correcting the bias in the bootstrap variance estimates.

McRoberts et al. (2002) pointed out several weaknesses in the k–NN estimator compared to parametric linear regression: the small k value may result in RMSE values larger than the standard deviation of the observations, and unrelated predictor variables included in the subset of covariates may increase the MSE. The latter case is related to the 'curse of dimensionality'; the rate of convergence for optimal solutions to non-parametric regression is slower in multidimensional cases (Linton & Härdle 1998). In the k–NN estimation, the observations from large feature space distances may be negatively correlated, whereas observations separated by large geographic distances are expected to be uncorrelated (Tokola et al. 1996, McRoberts et al. 2002). The k–NN estimates may be biased near the boundaries of the feature space, because the nearest neighbour distances tend to be greater and the neighbours may be concentrated in one direction only. The spatial distribution of the neighbours in the feature space can be taken into account in the estimation. Local adaptation of non-parametric methods models may help to overcome the edge effect problem as well as the bias caused by strong curvature in the true regression function (Malinen 2003).

The standard techniques for bandwidth selection may fail in a situation where the ϵ_i satisfy $\mathrm{E}(\epsilon_i | X_i = x) = 0$ but are autocorrelated. Altman (1990) studied the selection of bandwidth for the kernel estimator employing data with correlated errors. Cross-validation produces parameters favouring undersmoothing in this kind of situations (Altman 1990). A simple way to correct the effect of autocorrelation in cross-validation is to leave out more than one observation. Altman (1990) suggested either adjusting of the selection criteria or the transformation of residuals. The correlation function should be estimated from the data. However, when the form of the function is not known, the wrong choice of smoothing parameter can induce false serial correlation in the residuals (Opsomer et al. 2001).

3.3. Parameter selection in the MS-NFI k–NN estimation (I)

In the k–NN estimation, the overall error (or other selected criterion) is minimised by tuning the estimation parameters. The selected parameters are the features of interest and their weighting; the distance metric and the smoothing parameter, value of k (Malinen 2003). The MS-NFI also has parameters related to the selection of training data: stratification of the image and field plots on the basis of digital map data; and the geographical reference area from which the nearest neighbours are selected (Tomppo 1996, Tokola 2000).

The aim in (I) is to examine the selection of the estimation parameters employing the error estimates obtained from leave-one-out cross-validation. There were two objectives in the selection of parameters: to minimise the MSE of the key variable estimates and at the same time to retain some of the variation of the original field plot data in the spatial variation of the estimates. The statistical significance of the global bias in the k–NN estimates was also examined in (I). Only one set of parameters per satellite image is preferred to maintain the covariation between the field plot variables in the estimates, consequently a weighting (Tomppo & Halme 2004) or other compromise is required in the operative MS-NFI between the set of parameters obtained for different variables.

The original features of the Landsat TM spectral channel values and Euclidean distance measure were used. The weighting of the Euclidean distance had only a slight effect on the global MSE in (I), (c.f. Tokola et al. 1996). A mild topographic correction was carried out for the DN values of satellite image spectral channels using a modification of the Lambertian surface reflectance assumption employing digital elevation model. Outside of northern Finland, the topographic correction had only local significance.

The two somewhat contradictory objectives –minimising the MSE and retaining variation– have led to heuristic rules or subjective selection of k in MS-NFI applications employing Landsat TM or ETM+ image data. Several values of k have been applied: one (Franco-Lopez et al. 2001), 5–10 (Tomppo 1996), 10–15 (Tokola et al. 1996, Nilsson 1997), a minimum relative decrease RMSE k in (I) and an 'objective criteria' (minimum MSE) k_{opt} (McRoberts et al. 2002). In (I), the objectives defined earlier were met under the condition of minimum decrease of 0.5 % between k and $k+1$ sought from a window ranging from $k+1$ to $k+5$. This criterion was needed when different geographical reference areas were used to select the training data. It yielded k values 7–11 for the total volume estimates.

Landsat images cover geographically large areas that may contain edafic and climatic variation both horizontally and vertically. The atmospheric conditions and the radiometric properties of the image data may also vary within the image (Helder et al. 1992, Tomppo et al. 1998). The MS-NFI estimates will be biased for a forest area if there is locational dependency in the spectral values of pixels within the training data (Kilkki & Päivinen 1987). Kilkki & Päivinen (1987) proposed the use of the same training data (locationally uncorrelated) covering the particular surveyed forest area. On the other hand, the training data should be large enough to cover the true range and variation in the inventory area. A fixed size moving geographical horizontal (and vertical) reference area windows (HRA and VRA) have been used in the Finnish MS-NFI (Tomppo 1996). Because the locational dependencies are difficult to model explicitly, the global unbiasedness is checked using the cross-validation method.

The RMSE of the total volume and volume by tree species were studied against the geographical HRA radii. The mineral and peatland strata were analysed separately because there is high moisture content and moisture variation in the peatland soils compared to mineral soils. A near minimum MSE for volume estimates was obtained for mineral land already with a 20 km radius and for peatland with a 30 km radius, or employing 150–300 field plots. The maximum radius was sought by estimations based on field plots outside different geographical HRA. Significantly biased estimates were obtained for spruce and pine volume in some subregions that employed field plots from 40–60 km and larger radii. On mineral stratum, the 40–50 km geographical HRA radius yielded, on average, 400–600 field plots to the training data and did not increase the RMSE or decreased the bias in some cases. Nilsson (1997) in a simulation study recommended the same number of field plots for the estimation of total volume.

The area of peatlands is smaller than for mineral soils and their proportion varies across the country; generally larger geographical HRA radii, 60–90 km, are required to obtain a sufficient number of field plots. However, if the average number of field plots in the peatland stratum falls below 300, an estimation in two strata may not be justified. This map-based stratification is not very accurate and there are also differences within the peatland forests (Tomppo 1996). However, it was demonstrated in (I) that the stratification significantly decreased the global bias of the volume estimates within both strata.

Tokola (2000) found a 20 km geographical HRA radius to be optimal for total volume and pine and a 30 km radius for spruce and deciduous volume estimates in a study with NFI data in Eastern Finland applying cross-validation for error estimation. However, the decrease in the degree of determination was slow and the study material enabled radii only up to 40 km. Lappi (2001) in a small-area estimation study that used a calibration estimator and NFI field plots, concluded that 500 field plots outside the county to which the timber volume was to be estimated was reasonable in addition to the field plots of the county itself. To an average size county in the particular study area this would yield an approximately 35 km geographical HRA radius fixed to the centre of the county, assuming circular counties. However, the field plots outside the county obtained less weight in the estimation.

The parameters obtained are generally suitable for the MS-NFI, but a significant global bias in the results may still remain. Local bias may occur in the small-area estimates, especially in the edges of satellite image data or inventory area, when trend-like large-scale changes occur in the forest. The NFI sample is too small for reliable error estimation in small areas. The bias in the key field plot variables can be studied in the parameter selection phase or posterior to the k–NN estimation by comparing the MS-NFI estimates in the subregions (groups of municipalities) to the NFI field inventory estimates.

3.4. Error variations at the pixel level in the k–NN estimates of the MS-NFI (II)

There are several sources of error in the multisource forest inventories because they employ measurement data and models of different natures and scales. These errors contribute to the uncertainty in the k–NN estimates. At the pixel level, the prediction errors measured with relative RMSE are usually high, e.g. 50–80 % for field plot volume (I; Tokola et al. 1996). These error estimates are obtained by cross-validation.

The aim in (II) has been to study the variation in the error (residuals of the k–NN estimation by cross-validation) and to see whether there is a functional dependency between observable covariates and the prediction error. The potential explanatory variables for which the values could be obtained for every pixel were tested: i.e. estimated values of forest variables, variables of the selected nearest neighbour field plots and the spectral channel or digital map data values of pixels. The field plots in the training data were studied as an independent sample, ignoring the possible spatial autocorrelation between the field plots within the same cluster. The focus was on pixel-level prediction error of field plot volume and weighted mean of basal area (BA) observations in the k–NN estimation. The possible cumulation of systematic error in small areas was beyond the scope of the study.

The effect of locational error, which is quite significant in the MS-NFI training data, was minimised by employing a procedure to reassign the satellite image information to the field plot data (Halme & Tomppo 2001), or by restricting the number of mixed pixel field plots in the training data. The weighted mean of BA observations in and near the field plot was used instead of pure field plot BA to decrease the sampling error in the dependent variable. The use of weighted BA decreased the random variation (coefficient of variation) in the training data, as well as the MSE in the cross-validation. These results suggest that the optimum field plot size for MS-NFI purposes is larger than that currently applied when high resolution optical satellite data is used.

The standard deviation of the k neighbours' field plot variable was found to be a good measure of uncertainty. The estimated volume and BA correlated with the standard deviation and can be potentially employed in the analyses of uncertainty.

The residuals were studied against the spatial neighbourhood spectral variables, numerical map data (3×3 window) values and variables describing the spatial distribution, direction and clustering of neighbours in the Euclidean feature space. The first principal component of the field plot pixels, the spectral brightness feature (Horler & Ahern 1986), strongly correlated with the volume and BA estimates, and with their residuals from the k–NN estimation. Concerning the spatial neighbourhood, the bias in the estimates increased close to the non-FRYL map mask. This result supports the use of map data to stratify the MS-NFI in (IV). At the edges of the feature space, there should be more error in the k–NN estimates, but the variables describing the spatial distribution of the k neighbours did not correlate with the volume or BA residuals. The distances in DN for the majority of field plot pixels in the feature space are quite small compared to the possible magnitude of error in the Landsat TM data (Curran & Hay 1986).

The effect of the first principal component was removed from the residuals by using a model of field plot volume residual variances. The remaining variation was weakly correlated with the other potential explanatory variables. The random error component remained considerable in the k–NN residuals. At single field plot level, the cause of the error seemed to be case sensitive: mislocation of the field plot, the radiation from the surrounding land use classes or stands, the deviation of the target field plot from the surrounding forest and extreme field plot variable values.

3.5. Correction of map errors in the MS-NFI small-area estimates (III,IV)

The delineation of the inventory area is one of basic steps in planning and executing a forest inventory. The forest area estimate can be based on the sample and the remote sensing and map data can be employed as auxiliary data, e.g. in stratification (Loetsch & Haller 1973). The error component of the estimate of the area of FRYL is included in the total error of the estimate. In the Finnish NFI, the land area is assumed to be known, and the estimates, both for mean and total values, are based on ratio estimators of field sample plots (Tomppo et al. 1997). The standard errors are estimated using local quadratic forms (Matérn 1960). In the MS-NFI, the FRYL area has been delineated based on the numerical map data and in some cases from satellite image data (Tomppo 1991). More precisely, other land use has been estimated from the map data and the rest has been considered to be FRYL consisting of the forest land, other wooded land and waste land. The problem with the current MS-NFI map data is that it is not necessarily up-to-date, there are locational errors and it does not correspond exactly to the NFI land use classes. The aim in (III) and (IV) has been to reduce the map error in the MS-NFI small-area estimates: to obtain better FRYL area estimates and to correct the effect of map error in the forest resource estimates.

The error probabilities from the cross-tabulation (confusion) matrix of a classification can be used to correct or calibrate for misclassification bias in (remote sensing based) statistical estimates of class proportions (Hay 1988, Czaplewski & Catts 1992). The confusion matrix must be based on a statistical sampling scheme (Card 1982). In (III), a calibration method is introduced to reduce the map errors in MS-NFI small-area estimates. The method is based on the confusion matrix between land use classes of the field sample plots and corresponding map information, estimated from a large region. If the map strata can be expected to be reasonably homogeneous with respect to the map errors and land use class distribution, the proportions estimated for large region can be used for small areas (synthetic estimation) (Gonzalez 1973). In the calibration literature, the method is identified as "inverse calibration for classification error" (Brown 1982), introduced by Tenenbein (1972). In (III), the aggregates of the estimated land use class areas over the large region agree with unbiased post-stratification estimators (Holt & Smith 1979).

In (III), a method is found to calibrate the field plot weights $c_{i,U}$ for computation unit U in such a way that the sum of the calibrated weights over all training data plots is equal to the calibrated FRYL area estimates when applying the confusion matrix and the above method. The calibration of the weights is not straightforward because there are only FRYL field plots in the training data and there is a lack of correspondence between the NFI land use classes and the map strata. In addition, the calibrated MS-NFI may produce negative weights $c_{i,U}$ for some field plots.

In (IV), the k–NN estimation was employed by map strata. All the field plots within each map stratum, irrespective of the field measurement based land use class, were used for estimating the areas of land use classes and forest variables of the particular stratum. The applied strata were formed so as to be as homogeneous as possible with respect to the NFI based land use classes. However, the number of strata was restricted by the fact that there should be a sufficient number of field plots for the k–NN estimation (IV). The aim of the method was to obtain simultaneously the FRYL area estimate and accurate forest variable estimates within each stratum. A compromise was made in the parameter selection between the high overall accuracy of FRYL classification and minimising the MSE of the key forest variables. The stratified MS-NFI resembles the field inventory estimation in the sense that all the field plots within a stratum are retained in the training data. The final estimates are obtained by combining the stratum-wise estimates.

In (III) and (IV), the stratified and calibrated MS-NFI reduced the error in the FRYL area estimates caused by errors in the map data. Comparisons were made between the aggregates of MS-NFI small-area estimates and field inventory estimates at the region level in order to determine the total amount of correction, and at the subregions (groups of municipalities), to detect the possible bias in the small-area estimates. At the region level, the calibrated FRYL area estimates were by construction, equal to the post-stratified FRYL area estimates, and the post-stratification efficiently reduced the standard error of the estimate in land use classes that were homogeneous with the map strata (III). For the stratified MS-NFI, FRYL area correction remained between the original MS-NFI and the calibrated estimates. The calibration typically increased the volume estimates at both the region and subregion levels. The original MS-NFI estimates were calibrated upwards or downwards more or less systematically. The stratified MS-NFI small-area estimates, especially for volume and volume by tree species, varied more compared to the original MS-NFI estimates. The calibrated and stratified MS-NFI estimates of FRYL and total volume did not differ significantly from the field inventory estimates in subregions of size ranging from 1728 to 4238 km^2. However, only the stratified MS-NFI estimates of tree species volumes were within two standard errors of the field inventory estimates in the subregions of the test data. If the original MS-NFI estimates are clearly biased in the subregions, the calibration method alone can not correct the bias.

In the calibration method, the confusion matrices were calculated for large regions, where several thousands of field plots were available. The assumption of constant misclassification probabilities within the strata may not have held. The confusion matrices could be formed for subregions: according to Czaplewski & Catts (1992) improvement in the estimation precision of the classes starts to diminish after 500–1000 sample plots in a simple random or systematic sample. However, in (III) the smallest strata had less than 50 field plots.

Formation of the strata is more simple in the stratified MS-NFI, but the estimation parameters must be sought for all the strata applying cross-validation. The FRYL area estimates for each stratum were not very sensitive to the values of k or geographical HRA in (IV). The field plot weights w_{i,p_h} to pixel p_h in stratum h, i.e. the fuzzy membership values of field plot i, retain the variation in the training data in the estimates. The classification accuracy for FRYL and non-FRYL was not very high in (IV); the number of field plots within minor strata may be too small for efficient classification.

4. Discussion

In (I), the most important parameters for minimising the estimation error of the total volume and volume by tree species at pixel level were the value of k, the geographical HRA radius to select the training data and the stratification of the field plot pixels, and training data employing the site class map. With the parameter selection criteria employed, the parameters obtained were quite similar in the four different study areas that represented different geographical areas of Finland. This indicates a consistency in the quality of Landsat TM image data and in the NFI field plot data. The selection of k was based on the the condition of minimum decrease of 0.5 % between k and $k+\ 1$ on a smoothed prediction error curve in (I). According to McRoberts et al. (2002), the threshold percentage should be taken from the minimum RMSE. In general, if there is more than one criterion for selecting the estimation parameters, e.g. minimising the MSE and retaining some of the original variation in the field plot data in the estimates, it would be more objective to state and apply them in an analytical way. The use of a small value of k may be appealing because it retains the original variation of the field plot data in the produced map data (Franco-Lopez et al. 2001). However, a consequence may be that k–NN yields a MSE larger than the variance in the observations (McRoberts et al. 2002). Secondly, there is less variation in the forest variables for units the size of a Landsat TM pixel (30×30 m^2) than in the NFI field plots, c.f. Nyyssönen et al. (1967).

In (I), the geographical HRA radii for mineral land and peatland strata were determined using the following criteria: to minimise the MSE of the key variables, to exclude from the training data field plots that would introduce bias into the estimates (maximum HRA radius) as well as to obtain a sufficient number of field plots on average in the training data (minimum HRA radius). Tokola (2000) found a smaller HRA radius to be optimal when the criterion was to minimise the MSE of volume and volume by tree species from the cross-validation estimates. However, Nilsson (1997) recommended that the same number of field plots should be employed in the training data as were found to be suitable in (I) on mineral stratum. In northern Finland, there is more variation in the altitude and, according to experiences in the operative MS-NFI, the use of geographical VRA will decrease the bias in the vertical subsets of the training data (Tomppo et al. 1998).

Stratifying the image and field plots for mineral strata and peatland strata significantly decreased the bias of the volume estimates within those strata in (I). In general, stratifying the low radiometric resolution satellite data employing auxiliary data that reduces the within strata variation, e.g. a forest site quality map (Tokola & Heikkilä 1997) or stand characteristics data (Nilsson 1997, Tomppo et al. 1999) will reduce the bias within strata and possibly the global MSE in the k–NN estimation. The k–NN estimates of forest stand border pixels have a larger bias than those inside the stand and a separate estimation of stand boundaries would decrease this error (Tokola & Kilpeläinen 1999). The bias in the estimates also increases close to non-FRYL map strata in (II). In (IV), The MS-NFI by strata was employed. The relatively large amount of training data required limits the number of strata to be formed. Combining remote sensing data and map data will propagate different types of error in the output data (Wilkinson 1996). The stratified remote sensing classification may produce artificial boundaries on the output thematic maps (Hutchinson 1982).

In (I and II), the cross-validation has been applied assuming independent sampling, despite the fact that the key forest variables between neighbouring field plots within clusters are spatially correlated. E.g. the volume for forest and other wooded land had a correlation coefficient greater than 0.3 up to a distance of approximately 500 m within the same cluster in Central and Northern Finland in the 7th NFI (Tomppo et al. 2001). Spatial autocorrelation also occurs in the satellite image spectral channel values. This derives from both the sensor spatial properties and the spatial structure of the scene (Collins & Woodcock 1999). However, in the cross-validation it has not been detected in practice that the nearest neighbours would be more often from the same cluster as the target field plot. Nevertheless, the spatial autocorrelation range from the left-out pixel in cross-validation should be taken into account either by modifying the cross-validation (Altman 1990) or simply by the 'leave-some-out' method (Linton & Härdle 1998).

It is inevitable that the prediction error at the pixel level will be considerable in an MS-NFI that employs high resolution satellite data. The size of the field plot is small compared to the instant field of view of the satellite, the amount of mixed pixels is large and the image spectral channel values contain little variation for well-stocked stands (Ripple et al. 1991, Ardö 1992). However, reducing the main sources of error in the MS-NFI, e.g. in the field plot data, should decrease the prediction error in the k–NN estimates. Reducing the field plot locational error in the training data not only decreases the RMSE of mean volume estimates obtained from the cross-validation, but also retains more of the correct variation in the estimates (Halme & Tomppo 2001). It also corrects the typical shrinkage towards the mean in the k–NN estimates rather more than when a small value of k is used. The sampling error in the training data is decreased by the use of weighted mean of BA observations from a larger area than a field plot (II).

These results lead to the larger question of the optimal field sampling design for MS-NFI purposes. This will include the questions concerning the size of the field plot, the distance between field plots, the representativeness of the sample. When the field sample is used in a remote sensing application, an optimal spatial resolution of the remote sensing data may be selected for the estimation (Hyppänen 1996) or the resolution –and the sensor– may be fixed. Under budget constraints, a balance should be found between the need for a large enough field plot size to provide a good covariation between the remote sensing data and the key variables, and the need for the training data to cover the variation of field variables within the satellite image cover (I). The spatial autocorrelation in the forest variables and in the remote sensing data should be taken into account in this optimisation process, cf. Wang et al. (2001).

Further refinement of the estimation parameters could increase the accuracy of the forest variable estimates. The predictive power of the feature space variables employed can be summarised by applying canonical correlation analysis (Moeur & Stage 1995) or weighting the features based on optimisation rules (Tomppo & Halme 2004). This is useful when only one set of parameters is used for all the forest variables. The local adaptation of the k–NN method could be used, based on the selected nearest neighbours or on the spectral features. The larger k–NN estimates also had a larger residual variation and variation in the selected nearest neighbours in (II) and it might be possible to decrease the prediction error by applying a stronger smoothing for the pixels where high volume estimates will be produced. On the other hand, the spatial distribution of the k neighbours varies at the edges of the feature space and the Euclidean distances in DN are small between the field plot pixels of high stand volume, whereas in open land and in young forests the distances can be quite high.

The confusion matrices used for the calibration in (III) were estimated for entire forestry centres. If the error probabilities in the confusion matrix vary significantly within such large regions, the calibration could be split into subregions. *A priori* information of the map accuracies, efficient stratification to subregions and the evaluation of standard errors of the misclassification probabilities, c.f. (Card 1982), could be used to determine the optimal size and distribution of the subregions for calibration. In general, the stratified MS-NFI was a more simple method than calibration and provided, on average, more accurate estimates of the volume by tree species for small areas.

The field inventory estimates and their standard errors for large regions and subregions (groups of municipalities) are useful in assessing the systematic error of the MS-NFI estimates within a satellite image or some subarea of it (III; Tomppo & Katila 1992). The errors for field inventory estimates are large for areas less than 150 000 ha of FRYL, and other methods could be tested to evaluate the accuracy of the MS-NFI results, e.g. post-stratified field inventory estimates or resampling methods at the municipality level. There is both map error and forest variable estimation error in the aggregates of MS-NFI small-area estimates and this makes comparison with the field inventory estimates more difficult than in the cross-validation at pixel level, where only FRYL field plot pixels are employed. The parameter selection methods studied in (I) and the small-area estimation map error correction methods in (III and IV), together with the field inventory estimates, provide a method to reduce the estimation error and a reference of the accuracy of the MS-NFI results. However, if there is a significant systematic error in the small-area estimates of a certain subregion, it may not be possible to remove the error by varying the parameters studied in (I). In practice, the small-area estimates are dependent upon where the small area is located with respect to the employed satellite image and the training data. The satellite images and the large regions covered by the field inventory data form a mosaic of 'estimation images' that are analysed separately. Consequently, neighbouring pixels and small areas may employ training data from different geographical reference areas. This may cause bias in the results. It has been found necessary to take the tree species composition of the reference area into greater account , i.e. large scale trend-like changes of forest variables (Tomppo & Halme 2004). This indicates that the correlation between covariates and the volumes by tree species may not be strong enough to define the field plot weights $c_{i,U}$ for the small areas, and the use of averages of variables from a window defined by large scale trends around a municipality, decreases the error in the small-area estimates. The bias in the small-area estimator could be therefore corrected, e.g. by applying a combination of k–NN estimator and a direct sample estimator, a composite estimator, weighted by some criteria (Schreuder et al. 1993).

The parameter selection in the cross-validation is based on the global MSE and bias criteria. The systematic error in the aggregates of small-area estimates at the region and subregion levels are assessed by applying field inventory estimates. The aim in the MS-NFI is to obtain unbiased estimates for the small areas as well. The question is open as to, how much the optimal parameters for small areas or subregions would differ from the global optimum.

A spatial presentation of the estimation of uncertainty would be useful for the data analyst. Building an error estimation method based on sources of error is a complex problem (Bastin et al. 2000). The measures of uncertainty studied in (II) may be far from the true prediction of error and more information of the target pixels, especially mixed pixels, are needed. The finer resolution PAN images could help to assess the representativeness of the field plots and to decrease the estimation error. Also, the fact that pixel-level estimation errors can be spatially autocorrelated must be taken into account in the error estimation method (Congalton 1988, Flack 1995). Wallerman (2003) in a study employing Landsat TM and an intensive field sample, found the spatial dependence of the residuals from a spatial regression model to be lower than the residuals from ordinary least squares regression, but only with field plot data sampled by distances of less than 300 m.

Although a reliable method for estimating pixel-by-pixel error could be produced, such a method would not be suitable for deriving the error estimates for larger computation units such as forest stands and municipalities. The error estimates for larger areas cannot be obtained directly by combining the error estimates for single pixels due to spatial autocorrelation both in the satellite image and field data and, in the case of cross-validation error estimates, due to locational errors in the field plot data. The error variance of the MS-NFI for small areas could be estimated employing models describing the second order properties of the MS-NFI error estimates for pixels, obtained from cross-validation (Lappi 2001). However, the field plot volume prediction error of the MS-NFI estimates depends not only on distance between pixels but, e.g. on the true volume. In addition, the k–NN prediction errors may not be treated as the residuals of a trend surface of a spatial model. The several sources of error in the MS-NFI, both in the field plot data and the remote sensing data, can reduce the reliability of the spatial modelling of errors.

References

Altman, N. S., 1990. Kernel smoothing of data with correlated errors. J. Am. Stat. Assoc. 85:749–758.

—, 1992. An introduction to kernel and nearest-neighbour nonparametric regression. Am. Stat. 46(3):175–185.

Ardö, J., 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. Int. J. Remote Sens. 13(9):1779–1786.

Bailey, T. C. & Gatrell, A., 1995. Interactive Spatial Data Analysis. Longman Scientific & Technical, Harlow.

Bastin, L., Edwards, M. & Fisher, P., 2000. Tracking the positional uncertainty in 'ground truth'. In: Heuvelink, G. B. M. & Lemmens, M. J. P. M. (eds.), Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 12–14 July 2000, Amsterdam, The Netherlands. De Rode Hoed, Amsterdam, pp. 39–42.

Brown, P. J., 1982. Multivariate calibration. J. R. Stat. Soc. Ser. B 44(3):287–321.

Burrough, P. A. & McDonnell, R. A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, UK.

Campbell, J. B., 1996. Introduction to remote sensing. The Guilford Press, New York, 2 edition.

Card, D. H., 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogrammetric Engineering and Remote Sensing 48(3):431–439.

Collins, J. B. & Woodcock, C. E., 1999. Geostatistical estimation of resolution-dependent variance in remotely sensed images. Photogrammetric Engineering and Remote Sensing 65(1):41–50.

Congalton, R. G., 1988. Using spatial autocorrelation analysis to explore errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54:587–592.

Cunia, T., 1965. Some theory on reliability of volume estimates in a forest inventory sample. For. Sci. 11:115–128.

—, 1978. On the objectives and methodology of national forest inventories. In: Forest Resource Inventory. Joint Meeting of IUFRO groups S4.02 and S4.04. 18-26 June 1978. Institutul de Cercetari si Amenageri Silvice, XI-XXIX.,Bucharest.

Curran, P. J. & Hay, A. M., 1986. The importance of measurement error for certain procedures in remote sensing at optical wavelenghts. Photogrammetric Engineering and Remote Sensing 52(2):229–241.

Czaplewski, R. L. & Catts, G. P., 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sens. Environ. 39(1):29–43.

Dasarathy, B. V., 1991. Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques. IEEE Computer Society Press.

Davison, A. C. & Hinkley, D. V., 1997. Bootstrap methods and their application. Cambridge University Press.

Efron, B. & Tibshirani, R., 1997. Improvements on cross-validation – the .632+ bootstrap method. J. Am. Stat. Assoc. 92(438):548–560.

Eisele, F.-A., 1997. Practical use of the multi-source forest inventory system under Central European conditions. In: Managing the resources of the world's forests, Lectures given at the 1997 Marcus Wallenberg Prize Symposium, 14 October 1997, Stockholm, Sweden. The Marcus Wallenberg Foundation, Falun, Sweden, pp. 71–86.

Fix, E. & Hodges, J. L., 1951. Discriminatory analysis – nonparametric discrimination: Consistency properties. Report 4, USAF School of Aviation Medicine, Randolf Field, Texas. Technical Report Project 21-49-004.

Flack, J. C., 1995. On the Interpretation of Remotely Sensed Data Using Guided Techniques for Land Cover Analysis. Ph.D. thesis, Curtin University of Technology, School of Computing.

Foody, G. M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80:185–201.

Franco-Lopez, H., Ek, A. R. & Bauer, M. E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote Sens. Environ. 77:251–274.

Freden, S. C. & Gordon, F., Jr., 1983. Landsat satellites. In: Colwell, R. N. (ed.), Manual of remote Sensing. Vol I., Theory, instruments and techniques, American Society of Photogrammetry, The Sheridan Press, Falls Church, Virginia, U.S.A., pp. 517–570. Second edition.

Gjertsen, A. K., Tomter, S. & Tomppo, E., 2000. Combined use of NFI sample plots and Landsat TM data to provide forest information on municipality level. In: Zawila-Niedzwiecki, T. & Brach, M. (eds.), Remote sensing and forest monitoring, Proceedings of IUFRO conference, 1–3 Jun. 1999, Rogow, Poland. Office for Official Publications of the European Communities, Luxembourg, pp. 167–174.

Gonzalez, M. E., 1973. Use and evaluation of synthetic estimators. In: Proceedings of the Social Statistics Section. American Statistical Association, Washington, pp. 33–36.

Haara, A., Maltamo, M. & Tokola, T., 1997. The k-nearest-neighbour method for estimating basal-area diameter distribution. Scand. J. For. Res. 12(2):200–208.

Hagner, O., 1997. Textur i flygbilder för skattning av beståndsegenskaper. Working Report 29, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics, Umeå, Sweden. (in Swedish).

Halme, M. & Tomppo, E., 2001. Improving the accuracy of multisource forest inventory estimates by reducing plot location error – a multi-criteria approach. Remote Sens. Environ. 78:321–327.

Hastie, T. & Tibshirani, R., 1997. Generalized additive models. In: Kotz, S., Read, C. B. & Banks, D. L. (eds.), Encyclopedia of Statistical Sciences, update vol. 1, Wiley, New York, pp. 261–269.

Hay, A. M., 1988. The derivation of global estimates from a confusion matrix. Int. J. Remote Sens. 9(8):1395–1398.

Helder, D. L., Quirk, B. K. & Hood, J. J., 1992. A technique for the reduction of banding in Landsat Thematic Mapper images. Photogrammetric Engineering and Remote Sensing 58(10):1425–1431.

Holmgren, P. & Thuresson, T., 1998. Satellite remote sensing for forestry. Scand. J. For. Res. 13:90–110.

Holt, D. & Smith, T. M. F., 1979. Post stratification. J. R. Stat. Soc. Ser. A 142(1):33–46.

Horler, D. N. H. & Ahern, F. J., 1986. Forestry information content of thematic mapper data. Int. J. Remote Sens. 7(3):405–428.

Hutchinson, C. F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. Photogrammetric Engineering and Remote Sensing 48(1):123–130.

Hyppänen, H., 1996. Spatial autocorrelation and optimal spatial resolution of optical remote sensing data. Int. J. Remote Sens. 17(17):3441–3452.

Hyyppä, J., Pulliainen, J., Saatsi, A. & Hallikainen, M., 1997. Radar-derived forest inventory by compartments. IEEE Transactions on Geoscience and Remotes Sensing 35:392–404.

Kangas, A., 1996. Small-area estimates using model-biased methods. Can. J. For. Res. 26:758–766.

Keller, J. M., Gray, M. R. & Givens Jr., J. E., 1985. A fuzzy k-nearest neighbor algorithm. IEEE Transactions on Systems, Man, and Cybernetics 15(4):580–585.

Kilkki, P. & Päivinen, R., 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. In: Remote sensing-aided forest inventory. Proceedings of seminars organised by SNS, Hyytiälä, Finland, Dec. 10-12, 1986. University of Helsinki, Department of Forest Mensuration and Management,Research Notes No 19, Helsinki, Finland, pp. 209–215.

Korhonen, K. & Kangas, A., 1997. Application of nearest-neighbour regression for generalizing sample tree information. Scand. J. For. Res. 12(1):97–101.

Kulkarni, S., Lugosi, G. & Venkatesh, S., 1998. Learning pattern classification—a survey. IEEE Trans. Inf. Theory 44:2178–2206. Reprinted in S. Verdú, S.W. McLaughlin (editors.), Information Theory: 50 Years of Discovery, IEEE Press, New York, 1999.

Lappi, J., 2001. Forest inventory of small areas combining the calibration estimator and a spatial model. Can. J. For. Res. 31:1551–1560.

Linton, O. & Härdle, W., 1998. Nonparametric regression. In: Kotz, S., Read, C. B. & Banks, D. L. (eds.), Encyclopedia of Statistical Sciences, update vol. 2, Wiley, New York, pp. 470–485.

Loetsch, F. & Haller, K. E., 1973. Statistics of forest inventory and information from aerial photographes. BLV Verlagsgesellschaf, München.

Lund, H. G., 1993. Sampling designs for national forest inventories. In: Nyyssönen, A., Poso, S. & Rautala, J. (eds.), Proceedings of Ilvessalo symposium on national forest inventories,Finland, Aug. 17-21, 1992. The Finnish Forest Research Institute, Research Notes No 444, Helsinki, Finland, pp. 16–24.

Malinen, J., 2003. Locally adaptable non-parametric methods for estimating stand characteristics for wood procurement. Silva Fennica 37(1):109–120.

Maltamo, M. & Kangas, A., 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. Can. J. For. Res. 28(8):1107–1115.

Matérn, B., 1960. Spatial variation. Medd. fr. St. Skogsf. Inst. 49(5). Also appeared as number 36 of Lecture Notes in Statistics. Springer-Verlag, New York, 1986.

McRoberts, R. 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 Sens. Environ. 82:457–468.

Moeur, M. & Stage, A. R., 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. For. Sci. 41(2):337–359.

Mäkelä, H. & Pekkarinen, A., 2001. Estimation of timber volume at the sample plot level by means of image segmentation and Landsat TM imagery. Remote Sens. Environ. 77:66–75.

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

National Land Survey of Finland, 1996. Maastotietojen kohdemalli, Feature catalog of the topographic database. Maanmittauslaitoksen julkaisuja 71, National Land Survey of Finland. (in Finnish).

Nilsson, M., 1997. Estimation of Forest Variables Using Satellite Image Data and Airborne Lidar. Ph.D. thesis, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics. Acta Universitatis Agriculturae Sueciae. Silvestria 17.

Nyyssönen, A., Kilkki, P. & Mikkola, E., 1967. On the precision of some methods of forest inventory. Acta Forestalia Fennica 81(4):1–58.

Opsomer, J., Wang, Y. & Yang, Y., 2001. Nonparametric regression with correlated errors. Statistical Science 16:134–153.

Pekkarinen, A., 2002. Image segment-based spectral features in the estimation of timber volume. Remote Sens. Environ. 82(2-3):349–359.

Poso, S., 1972. A method for combining photo and field samples in forest inventory. Commun. Inst. For. Fenn. 76(1):1–133.

Poso, S., Paananen, R. & Similä, M., 1987. Forest inventory by compartments using satellite imagery. Silva Fennica 21(1):69–94.

Poso, S., Wang, G. & Tuominen, S., 1999. Weighting alternative estimates when using multi-source auxiliary data for forest inventory. Silva Fennica 33(1):41–50.

Rao, J. N. K., 1998. Small area estimation. In: Kotz, S., Read, C. B. & Banks, D. L. (eds.), Encyclopedia of Statistical Sciences, update vol. 2, Wiley, New York, pp. 621–628.

Ripple, W. J., Wang, S., Isaacson, D. & Paine, D., 1991. A preliminary comparison of Thematic Mapper and SPOT-1 HRV multispectral data for estimating coniferous forest volume. Int. J. Remote Sens. 12(9):1971–1977.

Schreuder, H. T., Gregoire, T. G. & Wood, G. B., 1993. Sampling methods for multiresource forest inventory. Wiley, New York, U.S.A.

Schreuder, H. T., LaBau, V. J. & Hazard, J. W., 1995. The Alaska four-phase forest inventory sampling design using remote sensing and ground sampling. Photogrammetric Engineering and Remote Sensing 61(3):291–297.

Stone, C. J., 1977. Consistent nonparametric regression. The Annals of Statistics 5:595–645.

Tenenbein, A., 1972. A double sampling scheme for estimating from misclassified multinomial data with applications to sampling inspection. Technometrics 14(1):187–202.

Tokola, T., 2000. The influence of field sample data location on growing stock volume estimation in Landsat TM-based forest inventory in Eastern Finland. Remote Sens. Environ. 74(3):422–431.

Tokola, T. & Heikkilä, J., 1997. Improving satellite image based forest inventory by using a priori site quality information. Silva Fennica 31(1):67–78.

Tokola, T. & Kilpeläinen, P., 1999. The forest stand margin area in the interpretation of growing stock using Landsat TM imagery. Can. J. For. Res. 29:303–309.

Tokola, T., Pitkänen, J., Partinen, S. & Muinonen, E., 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. Int. J. Remote Sens. 17(12):2333–2351.

Tomppo, E., 1987. An application of a segmentation method to the forest stand delineation and estimation of stand variates from satellite images. In: Proceedings of the 5th Scandinavian Conference on Image Analysis. Stockholm, pp. 253–260.

—, 1991. Satellite image-based national forest inventory of Finland. International Archives of Photogrammetry and Remote Sensing 28:419–424. Proceedings of the Symposium on Global and Environmental Monitoring, Techniques and Impacts, 17–21 Sept. 1990, Victoria, British Columbia, Canada.

—, 1996. Multi-source National Forest Inventory of Finland. In: Päivinen, R., Vanclay, J. & Miina, S. (eds.), New Thrusts in Forest Inventory. Proceedings of the subject group S4.02-00 'Forest Resource Inventory and Monitoring' and subject group S4.12-00 'Remote Sensing Technology', vol. 1, IUFRO XX World Congress, 6–12 Aug. 1995, Tampere, Finland. European Forest Institute, Joensuu, Finland, pp. 27–41.

Tomppo, E., Goulding, C. & Katila, M., 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scand. J. For. Res. 14:182–192.

Tomppo, E. & Halme, M., 2004. Using large area forest variables as ancillary information and weighting of variables in k-nn estimation - a genetic algorithm approach. Manuscript.

Tomppo, E., Henttonen, H. & Tuomainen, T., 2001. Valtakunnan metsien 8. inventoinnin menetelmä ja tulokset metsäkeskuksittain Pohjois-Suomessa 1992–94 sekä tulokset Etelä-Suomessa 1986–92 ja koko maassa 1986–94. Folia For. 1B/2001:99–248. (in Finnish).

Tomppo, E. & Katila, M., 1992. Satellite image-based national forest inventory of Finland - first large area results. In: Folving, S., Ertner, G. & Svendsen, T. B. (eds.), European Collaborative Programme workshop on Remote Sensing for Forestry Applications, Copenhagen 13-15 Nov., 1991. pp. 63–67.

Tomppo, E., Katila, M., Moilanen, J., Mäkelä, H. & Peräsaari, J., 1998. Kunnittaiset metsävaratiedot 1990–94. Folia For. 4B/1998:619–839. (in Finnish).

Tomppo, E., Mikkelä, P., Veijanen, A., Mäkisara, K., Henttonen, H., Katila, M., Pulliainen, J., Hallikainen, M. & Hyyppä, J., 1996. Application of ERS-1 SAR data in large area forest inventory. In: Guyenne, T.-D. (ed.), Proceedings of the Second ERS Applications Workshop, London, UK, 6-8 December 1995. Noordwijk. pp. 103–108.

Tomppo, E., Nilsson, M., Rosengren, M., Aalto, P. & Kennedy, P., 2002. Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. Remote Sens. Environ. 82:156–171.

Tomppo, E., Varjo, J., Korhonen, K., Ahola, A., Ihalainen, A., Heikkinen, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S. & Tuomainen, T., 1997. Country report for Finland. In: Study on European Forestry Information and Communication Systems. Reports on forestry inventory and survey systems, Office for Official Publications of the European Communities, Luxembourg, volume 1, pp. 145–226.

Wallerman, J., 2003. Remote sensing aided spatial prediction of forest stem volume. Ph.D. thesis, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics. Acta Universitatis Agriculturae Sueciae. Silvestria 271.

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.

Wilkinson, G. G., 1996. A review of current issues in the integration of gis and remote sensing data. International Journal of Geographical Information Science 10(1):85–101.

Woodcock, C. E. & Strahler, A. H., 1987. The factor of scale in remote sensing. Remote Sens. Environ. 21:311–332.

I

ELSEVIER

Remote Sensing of Environment 76 (2001) 16–32

Remote Sensing of Environment

www.elsevier.com/locate/rse

Selecting estimation parameters for the Finnish multisource National Forest Inventory

Matti Katila*, Erkki Tomppo[1]

Finnish Forest Research Institute, Unioninkatu 40 A, FIN-00170 Helsinki, Finland

Received 12 January 2000; accepted 12 September 2000

Abstract

The paper examines the selection of parameters for the nonparametric k-NN estimation method that is used in the Finnish multisource National Forest Inventory (MS-NFI). The MS-NFI utilises NFI field plot data, optical area satellite images and digital maps and produces forest variable estimates from the single pixel level up to the national level. The most important parameters to be selected are: the distance metric, the number of the nearest neighbours, k, parameters related to the digital elevation model, stratification of the image data, as well as the width of the moving geographical horizontal and vertical reference areas (HRAs and VRAs). The root mean square errors (RMSEs) and significance of biases at pixel level were evaluated in order to find optimal parameters. A leave-one-out cross-validation method was applied. The emphasis is placed on the search for moving geographical HRAs and VRAs, as well as in the stratification of the field plots and the satellite images on the basis of auxiliary data. Stratification reduces the bias of the estimates significantly within each strata. With the current sampling intensity of the Finnish national forest inventory, a geographical HRA with a radius of 40–50 km was found optimal for the total volume estimates and for volumes by tree species in the mineral land map stratum. On the average, there was a sufficient number of field plots to cover the variation of forest variables within the image area to be analysed. The inclusion of field plot data beyond this area introduced bias to the estimates. For the peatland strata, a wider reference area, 60–90 km, was needed. A VRA, together with topographic correction of the digital values of images, reduced the standard error of the volume estimates in Northern Finland.

Keywords: Nonparametric estimation; Satellite images; Multisource forest inventory; Stratification; Cross-validation; Training data selection

1. Introduction

The trend in large area inventories is towards geographically accurately located information and small area estimates. Under Finnish conditions this means municipality and forest holding level estimates (Eisele, 1997; Franco-Lopez et al., 2000; Gjertsen et al., 2000; Nilsson, 1997; Tokola & Heikkilä, 1997; Tomppo, 1991; Tomppo et al., 1999a, 1999b).

The use of satellite images in forest inventories has been studied since the beginning of 1970s. The focus has been on the estimation of basic variables, such as volumes by tree species, basal area, age and mean breast height diameter of stand (Hagner, 1997; Tokola et al., 1996; Tomppo, 1987, 1991). Parametric and nonparametric regression, as well as neural networks, together with segmentation techniques, have been used. Forest inventories involve high numbers of variables measured in the field, typically between 100 and 400 variables concerning, e.g., site, volume and increment of growing stock, forest damages and forest biodiversity. Estimates for all of these are usually necessary.

The Finnish multisource National Forest Inventory (MS-NFI) has utilised optical area satellite images and digital maps, in addition to field plot data, since 1990. A nonparametric k-nearest neighbour method (k-NN) deviates from the usually applied methods and has made it possible to estimate all inventory variables at the same time (Tomppo, 1991). Field data from surrounding units (municipalities), in addition to the unit itself, are utilised when estimating results for one unit; the method is known as synthetic estimation in statistical literature (e.g., Rao, 1998). This

* Corresponding author. Tel.: +358-9-857-05-312; fax: +358-9-625-308.

E-mail address: matti.katila@metla.fi (M. Katila).

[1] Tel.: +358-9-857-05-340; fax: +358-9-625-308. *E-mail address*: erkki.tomppo@metla.fi (E. Tomppo).

0034-4257/00/$ – see front matter
PII: S0034-4257(00)00188-7

makes it possible to obtain estimates for smaller areas than would be possible with sparse field data only (Franco-Lopez et al., 2000; Nilsson, 1997; Tomppo, 1996). The method produces georeferenced information, thematic maps and small area statistics. The field plot data should, however, cover the variation of field variables within the satellite image cover. Consequently, a large number of observations is required.

The k-NN algorithm searches the feature space for the k nearest pixels, whose field data vectors are known, applying a distance measure, d, defined in the feature space. Field data from the k nearest pixels is transferred to the unknown pixel. The method has been widely studied in pattern recognition (Cover & Hart, 1967; Keller et al., 1985) and statistics (Linton & Härdle, 1998). Altman (1992) showed that the k-NN estimator may give biased estimates as the value of k increases, but that the bias can be reduced with weighted averages of the k neighbours. The error rate asymptotically approaches the optimal rate of the Bayes decision rule for discrete variables when both the k and n (number of observations) tend to infinity in such a way that $k/n \rightarrow 0$ (Keller et al., 1985).

A set of parameters is chosen for the k-NN method in the operative MS-NFI. Examples are: (1) the image features; (2) the distance measure; (3) the value of k, i.e., the number of the nearest neighbours; (4) parameters related to the possible use of digital elevation model; (5) stratification of the image and field plots to mineral land and peatland on the basis of a digital site class map, produced by the National Land Survey (NLS); and (6) the geographical reference area from which the nearest field plots are selected. The geographical reference area is crucial for the estimation procedure and is selected separately for each pixel in the Finnish MS-NFI (Tomppo, 1996).

Franco-Lopez et al. (2000) and Nilsson (1997) studied different distance metrics using the k-NN method. Several studies have been conducted for selecting the optimal value of k (Franco-Lopez et al., 2000; Nilsson, 1997; Tokola et al., 1996; Tomppo, 1996; Tomppo et al., 1998b). It is affected by the layout and the size of the field plots, size of the pixel and the variation of the field variables. Compromises are often needed due to the fact that retaining the variation of field variables in the estimates may presume a low value of k, while minimising of pixel level root mean square errors (RMSEs) presumes higher value of k (cf. Franco-Lopez et al., 2000). Stratification of the study area and field plots has been studied on the basis of supplementary data such as site quality maps and old forest management planning data, e.g., by Tokola and Heikkilä (1997), Tomppo et al. (1999b) and Tomppo et al. (1998a, 1998b).

The selection of the geographical reference area on a large scale has not been systematically studied, partly due to lack of large scale test data. The objective of this paper is to fill this gap. The paper addresses the selection of the reference area, both in horizontal and vertical directions (horizontal and vertical reference area (HRA and VRA)). Another goal is to study the stratification of the field plots based on supplementary data: in this case, the digital peatland map. The selection of the parameters, especially the value of k, must also be addressed in order to complete the reference area selection and stratification in an optimal way.

There are several reasons for the use of pixel-dependent geographical HRA and VRA from the possible nearest field plots to the pixel to be analysed. A large forest area, covered, e.g., by one Landsat 5 or 7 Thematic Mapper (TM) satellite image (with a size of 183×172 km), may involve a gradual change in vegetation structure. In Finland, the vegetation zone may change, e.g., from South Boreal to Middle Boreal. This often implies that the average structure of the growing stock, as well as other forest variables, also change. The proximity of large lakes or sea, as well as elevation variations, affect the average structure of the growing stock and other vegetation composition as well. The relationship between growing stock and image features may vary because of these changes. Too wide an HRA, i.e., too large a value for the geographical maximum distance, may lead to biased estimates. On the other hand, when field plot layout is sparse, a minimum distance is needed to include all the local variation of the forest variables in the field plots.

The high moisture content and large moisture variation make the reflectance of peatland forests very different from that of mineral soil forests, even with a similar structure of the growing stock (Tomppo, 1987). A stratification of the image area and the field plots have been made according to the digital site class map in the operative MS-NFI (Tomppo, 1996; Tomppo et al., 1998b). The proportion of peatlands of the land area varies with inventory areas. The hypothesis presented here is that different geographical HRAs are needed for peatland strata and mineral soil strata.

Developing an analytical method for deriving the standard error of the estimates of a forest area is a challenging task due to the spatial dependencies of the forest variables and the image data itself. A satisfactory analytical solution is still under development. However, statistically reliable error estimates of forest variables from the pure field inventory data can be used to assess the MS-NFI results (Katila et al., 2000; Tomppo & Katila, 1992). For these purposes, a large enough part of the image must be analysed.

The leave-one-out cross-validation method is applied in this paper to estimate the average biases and RMSE of predictions at the single pixel level for different combination of k-NN estimation parameters: particularly VRA, HRA and strata. The parameters are chosen in such a way that the RMSEs of the estimates are minimised and the biases of the estimates are simultaneously kept within twice the standard error from the value 0. The procedure is also applied to control errors by strata defined by field variables. Errors by volume classes are important, especially in map production (Franco-Lopez et al., 2000).

2. Materials

Four areas in Finland were chosen for this study. These were located approximately between longitudes 21°40′E, 30°25′E and latitudes 59°40′N, 68°10′N (Fig. 1). The land areas varied between 13 878 and 38 220 km^2 (Table 1). The test data contains field measurements from the 8th and 9th NFI, and digital map data and satellite images as applied in the MS-NFI. The study areas were chosen in such a way that the image acquisition and the field inventory were from the same year, that the image quality was good and that the image area contained as many field plots as possible. The structure and average volume (m^3/ha) of the growing stock vary within study areas and especially between study areas. The four study areas cover the greater part of the variation in land use classes, soil properties, tree species variation and climatic in Finland. The Western Finland study area: (1) contains large peatland areas, the Central Finland study area; (2) is rich in fertile mineral soils and the southwestern image; and (3) has a relatively high nonforestry land proportion. The Northern Finland study (4) area is from the north boreal vegetation zone, in which Scots pine (*Pinus sylvestris* L.) dominates, and has moderately high elevation variation. The forests of the study areas are either pine or Norway spruce (*Picea abies* (L.) Karst.) dominated with birch (*Betula* spp.) and other deciduous species as a mixture.

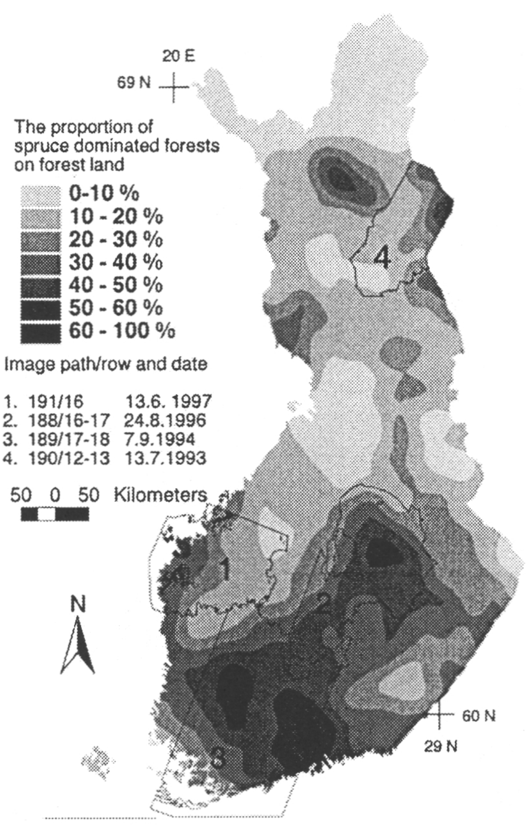


Fig. 1. The coverage of the four Landsat 5 TM satellite images plotted over the map of the proportion of spruce-dominated forests on forest land based on the 8th NFI field data.

The field sample of the NFIs were measured from systematically located clusters of sample plots. The sample plots (14–18 per cluster) were located along a rectangular or L-shape tract at 200- to 300-m intervals, depending on the area. Trees were measured from field plots belonging to forest and other wooded land (FOWL) stands. The tally trees were selected with PPS sampling (sampling with probability proportional to size), applying a basal area factor of 1.5 in the Northern study area and 2 elsewhere. The probability of a tree's inclusion is proportional to its cross-sectional area at a height of 1.3 m; a maximum radius of 12.45 m was used in Northern Finland and 12.52 m in Central and in Western Finland. Unrestricted PPS sampling (sampling with no maximum distance) was applied in the southwestern study area. The inclusion of 'border' trees is carefully checked. The distance of the nearest forest stand boundary from the field plot centre point was recorded in 10-m classes from 0 to 40 m.

The field plots that are within forestry land (FRYL) are selected from the NFI field sample for the following analyses. They are divided into forest land, other wooded land and waste land, according to site productivity (Table 2) (Tomppo et al., 1998a, 1999b; Tomppo et al., 1997). The mean and the standard deviation of the volumes of the field plot measurements for the main tree species, mean basal area and age of the field plot stand are presented in Table 3.

The Landsat 5 TM satellite images employed were rectified to the national grid coordinate system with regression models of first or second-order polynomials calculated from 35 to 70 ground control points. These were identified from topographic maps and satellite images. The model residuals were checked over images to ensure an even distribution of the rectification model errors. The mean square error of the model, together in the direction of rows and columns, varied between 0.6 and 0.7 satellite image pixels. Nearest neighbour resampling was used with a pixel size of 25×25 m^2 for intensities (Tomppo et al., 1998b).

3. Methods

3.1. MS-NFI estimation method

Multisource estimates are computed for FRYL pixels. FRYL is separated from the other land use classes by means of digital map data in the current MS-NFI. Cloud-free FRYL areas of a satellite image are analysed with the FRYL field plots i chosen for the training data set. Incorrectly located field plots and those that contain non-FRYL land use classes

Table 1
Study areas: satellite images of the 8th MS-NFI and 9th MS-NFI, field inventory data

	Field plot data						Satellite image	
Study area	NFI	Land area (km^2)	FRYL (%)	Year	Cluster distance (km)	Plots per cluster	Landsat 5 TM	Date
Southwestern Finland	8+	38220	69.1	1994	7 × 16	16	189/17–18	940709
Central Finland	9	18787	82.0	1996	7 × 7	18/14[a]	188/16–17	960824
Western Finland	9	13878	73.0	1997	7 × 7	18/14[a]	191/16	970613
Northern Finland	8	13687	98.1	1993	10 × 10	15	190/12–13	930713

[a] Every fourth cluster had 14 field plots in the Central and Western Finland study areas.

are excluded from the training data set, the proportion is usually in the range of 2–6%.

The MS-NFI estimates are weighted averages of the field plot variables. The k-NN method is used to calculate the weights (Keller et al., 1985; Tomppo, 1991). Data from the k nearest field plots in the feature space, $i_1(p)$, ..., $i_k(p)$ are utilised in the analysis of each pixel p. More precisely, the field plots are sorted according to Euclidean feature space distance between p_i and p, and the k nearest plots are chosen.

Stratification of the FRYL area and the training data to peatlands and mineral soils according to a numerical map data has usually been applied in such a way that only pixels within the same stratum as the target pixel are accepted as neighbours. The geographical distance to the potential nearest neighbours has been restricted to 40–120 km due to gradual changes of vegetation type. Simultaneous upwards and downwards vertical maximum distances of 50–150 m have also been applied, particularly in North Finland. Cross-validation-based error estimation and the large area subregion estimates from field data have been used to decide upon the suitable geographical reference area while keeping in mind the need for a certain minimum number of field plots (Tomppo, 1996; Tomppo et al., 1998b).

The weight $w_{i,p}$ of the field plot i to the pixel p is defined as

$$w_{i,p} = \frac{1}{d^t_{p_i,p}} \Big/ \sum_{j \in \{i_1(p),\ \ldots, i_k(p)\}} \frac{1}{d^t_{p_{(j)},p}}, \quad \text{if and only if } i \in \{i_1(p), \ldots, i_k(p)\} = 0, \quad \text{otherwise,} \tag{1}$$

where $\{i_1(p), \ldots, i_k(p)\}$ is the set of the field plots whose corresponding pixels are the k nearest ones to the plot p.

A value $t = 1$ was applied for the weighting parameter in this study. A small positive value is given for 0 distances.

The weight $w_{i,p}$ can be interpreted as that share of the pixel p that obtains data from the field data vector of the plot i. For a single pixel p, the estimate of the average of a continuous variable is expressed by Eq. (2)

$$\hat{m}_p = \sum_{i \in \text{FRYL}} w_{i,p} m \tag{2}$$

For more details, see (Tomppo, 1996).

3.2. Feature selection

The original Landsat 5 TM channels 1–5 and 7 are used. Topographic correction for the digital number values of channels has been made on rugged terrain with a modification of the Lambertian surface reflectance assumption. The normalised intensity value I^* is calculated from the observed intensity value I and angle α between sun and the normal of the land surface (Eq. (3)).

$$I^* = I/\cos^r \alpha \tag{3}$$

The exponent $r < 1$ has been added to the denominator, because the Lambertian reflecting surface assumption is not necessarily true for a varying forest area. The value $r = 1$ usually leads to overcorrection (Tomppo, 1992).

3.3. Results validation

The choice of the classification parameters was tested with a leave-one-out cross-validation method: a single field plot p_i belonging to the ground truth data set is classified with the other plots (Linton & Härdle, 1998). Other possible

Table 2
The land use class distribution of the FRYL field plots over the study areas, the minimum and maximum elevation above the sea level and the proportion of field plots in peatland and mineral soil strata of the site class map

Study area	Elevation range (m)	Forest land (%)	Other wooded land (%)	Waste land (%)	FRYL (no. of plots)	Peatland stratum (%)	Mineral soil stratum (%)
Southwestern Finland	0–213	95.1	3.5	1.4	3546	11.7	88.3
Central Finland	79–301	96.9	1.7	1.4	6220	19.4	80.6
Western Finland	0–223	89.2	5.9	5.0	4661	30.1	69.9
Northern Finland	149–549	67.4	20.1	12.5	2013	29.1	70.9

Table 3
The mean and the standard deviation of the volume of the growing stock of field plots by tree species, basal area (BA) of the field plot stand (three measurement points) and the field plot stand age on FOWL of study areas

	Southwestern Finland		Central Finland		Western Finland		Northern Finland	
Variable	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
Volume (m^3/ha)	131.8	103.0	114.6	99.9	89.0	82.4	35.4	40.9
Volume of pine (m^3/ha)	41.9	57.4	37.8	54.8	44.3	53.7	21.9	35.3
Volume of spruce (m^3/ha)	68.3	89.3	53.9	84.3	27.8	57.4	8.7	19.4
Volume of deciduous species (m^3/ha)	21.6	40.0	22.8	38.1	17.0	33.9	4.7	11.0
BA (m^2)	16.9	10.4	15.7	10.5	13.2	9.9	5.6	6.3
Age (years)	56.8	37.1	51.7	36.1	58.1	39.6	68.7	83.8

methods are the hold-out estimator (i.e., "data splitting"), jack-knifing and bootstrapping.

The RMSE has been used as a measure of reliability of the continuous variables (Eq. (4)).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}, \tag{4}$$

where y_i, $i=1, \ldots, n$ are the values of variables in the training data set and $\hat{y}_i$ is the estimated value. Other criteria are bias (Eq. (5)) and the standard error of bias (Eq. (6)).

$$\bar{e} = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \tag{5}$$

$$s(\bar{e}) = \frac{s(e)}{\sqrt{n}}, \tag{6}$$

where $s(e)$ is the standard deviation of errors $\hat{y}_i - y_i$ and also the variance component of the RMSE, which does not include the possible bias.

The quantity $s(\bar{e})$ can be used for testing whether the bias deviates significantly from zero. Deviations greater than $2s(\bar{e})$ from the field plot based estimate of mean are here considered to be statistically significant.

The cross-validation errors are studied within strata of variables or by location, i.e., as soil class and subareas of the study area in order to obtain an idea of the possible bias in subclasses.

The options considered in the field plot data selection and stratification in the error analysis are: (1) maximum geographical distance (vertical and horizontal) from the pixel under analysis to the potential nearest neighbours and (2) stratification of field plots and image area based on auxiliary data (digital site class map).

4. Results

4.1. Selection of the number of k nearest neighbours

A practical rule for the selection of k was developed in the following tests. The RMSE normally decreases as k increases until a minimum RMSE is reached (Fig. 2). The minimum may not be reached before $k=30$, but the decrease levels off between 10 and 15 for mineral soil stratum and slightly earlier for peatland stratum, where there are fewer observations (Table 2).

Four different ways of selecting the value of k were tested for geographical HRAs of 10–200-km radii: (1) the minimum RMSE between 1 and 30; (2) the minimum RMSE controlled by the significance of the bias; (3) a choice under the condition of minimum decrease of 0.5% in RMSE between k and $k+1$; and (4) a fixed $k=10$ (Fig. 3).

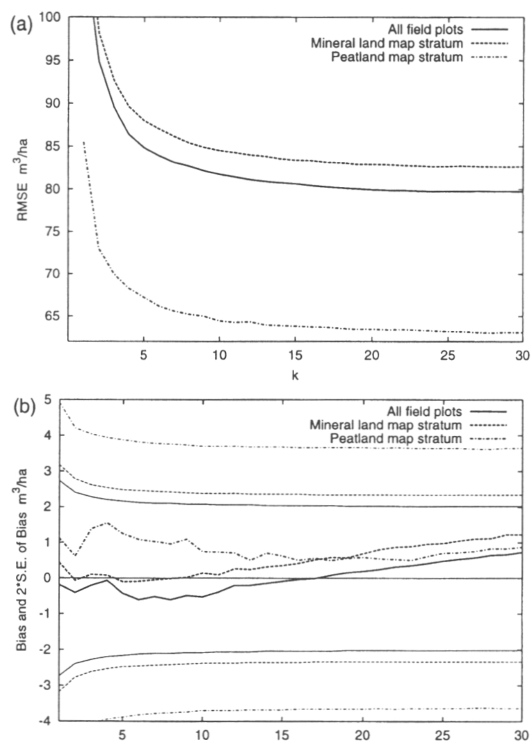


Fig. 2. RMSE (a), bias and double S.E. of bias (b) of total volume estimates against number of k for 40 (mineral soil stratum and all plots) and 70 km (peatland stratum) geographical HRA, stratification and no stratification, Central Finland study area.

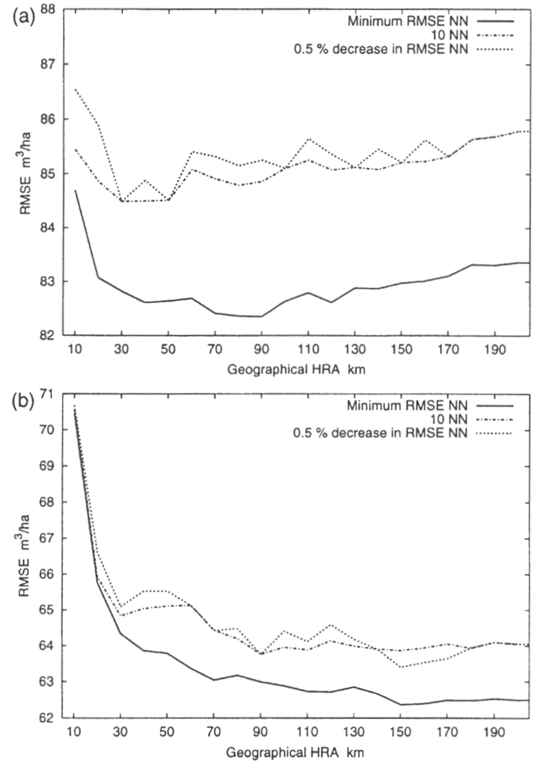


Fig. 3. RMSE of total volume estimates for differently chosen k values against geographical HRAs of 10–200 km; (a) mineral soil stratum and (b) peatland stratum, Central Finland study area.

The biases of the minimum RMSE option were nonsignificant for all HRAs. The values of k were equal to options (1) and (2). The values of k obtained with the 0.5% decrease in the RMSE rule were near to 10. Note, however, that the test is for the total volume only.

Options (1) and (2) produced the smallest RMSE, but since it was desirable to retain some of the variation of the original field plot data in the pixel level estimates (cf. Moeur & Stage, 1995), alternatives (3) and (4) were employed in the following calculations.

4.2. Stratification of the field data to peatlands and mineral soils

The stratification of the field plot data according to (1) site class map and (2) field plot main site class was tested in the cross-validation. The RMSE and the bias of the total volume estimates from the cross-validation were studied separately for each stratum and for the whole field sample plot data of the study area. The value of k was selected with the condition of minimum decrease of 0.5% in RMSE between k and $k+1$ in this analysis.

The advantage of stratification becomes clear when the average biases of the volume estimates are compared within both site class map strata, classified by all the sample plots and by only the sample plots within each stratum. The former estimates from the cross-validation are significantly biased (Fig. 4). The bias of the total volume estimate changed from -2.7 to -6.4 m^3/ha on the mineral soil stratum and from -1.5 to 6.0 m^3/ha on the peatland stratum when the map based stratification was left out in Western Finland study area with a 50-km radius of HRA. The site class map based stratification decreased the global RMSE value only by approximately 0.5 m^3/ha, with 40 km and larger radius of geographical HRA. The use of precise site class information for the stratification, based on the field plot data, did not improve the accuracy of the estimates. The stratification is applied overall in the following tests.

4.3. Horizontal and vertical geographical reference area

Different sizes of pixel-dependent HRA and VRA were studied; i.e., the variation available in the corresponding

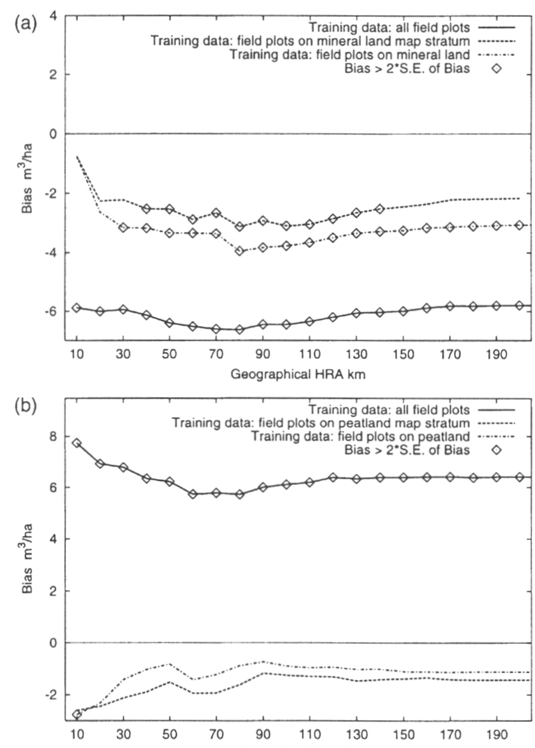


Fig. 4. Bias and the significance of bias of total volume estimates for the Western Finland study area. Target field plot data stratified to mineral soil (a) and peatland (b) according to the site class map; training data: all field plots, particular stratum field plots stratified according to map data and field plot data.

training data and the RMSE and bias of the pixel level estimates of total volume and volumes by tree species. Geographical reference areas were sought, which would yield the minimum RMSE and unbiased estimates and would have a sufficient amount of field plots in the training data.

4.3.1. Minimum size of the reference area

The applied k-NN estimation method utilises a pixel-dependent geographical HRA. The spatial variation of field variables affects the selection of the area: The larger the geographical area of the training data, the better it covers the true variation of the values of the field variables. This can be seen from the distribution of the standard deviations of volume computed from the surrounding training data of each field plot in the Central Finland area. When the HRA increases, the standard deviations concentrate around the one computed from the whole field data (volume 103.6 m^3/ha) (Fig. 5).

A larger geographic HRA is necessary for more rare combinations of field plot variables, as can be seen in Fig. 6. For example, in the training data for the Western Finland FRYL area on mineral soil stratum, an average HRA of 40-km radius is required to obtain 10 field plots from spruce-dominated forests of volume 250–300 m^3/ha and age >80 years. For the particular image, the 40-km HRA would seem to be the minimum for obtaining a sufficient amount of nearest neighbour candidates.

4.3.2. Maximum size of the reference area

The contribution to the volume estimates of field plots from different geographical distances was studied in order to better reveal their value in the estimation of volume. A cross-validation test, complementary to the above tests for selection of the training data, was made. Only the field plots beyond a certain radius were used.

The RMSE and the bias of tree species' volume estimates were calculated from the field plot data outside geographical HRAs of 0–200 km (cf. Tokola, 1998). The area of South-western Finland was chosen due to its large area and variation in the spruce dominance of the forests. The RMSE and biases of the estimates were studied for the whole study area and for three smaller subareas (Fig. 7): (1) spruce-dominated (335 target plots) and pine-dominated areas (2) (515 target plots) and (3) (687 target plots). The number of field plots selected for the training data was kept constant for each sample plot (or cluster). Only mineral soil stratum field plots were applied with $k=10$.

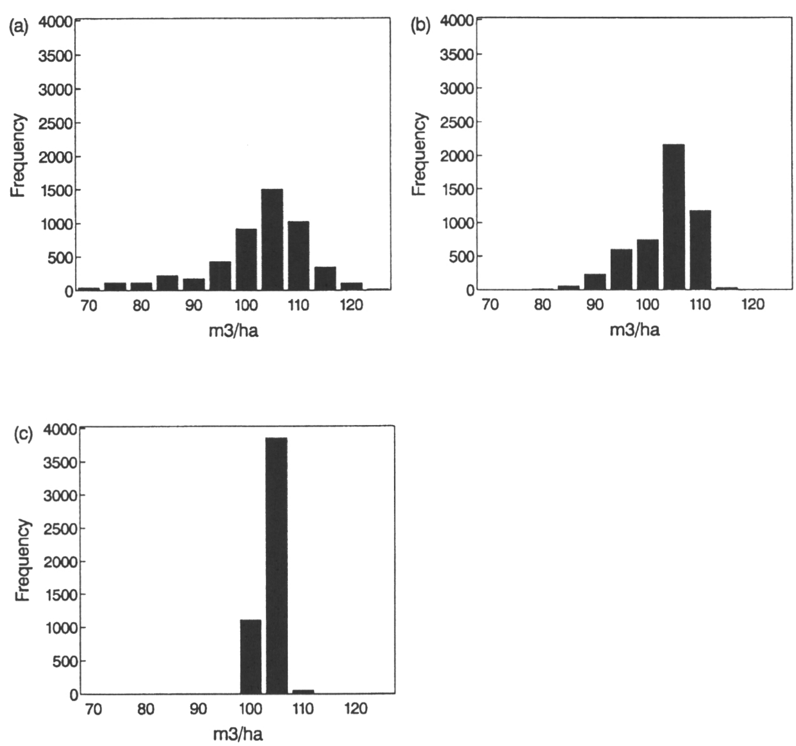


Fig. 5. The distribution of standard deviation of volume in the training data for the target field plots of the Central Finland study area; (a) 20 km, (b) 40 km and (c) 100 km geographical HRAs.

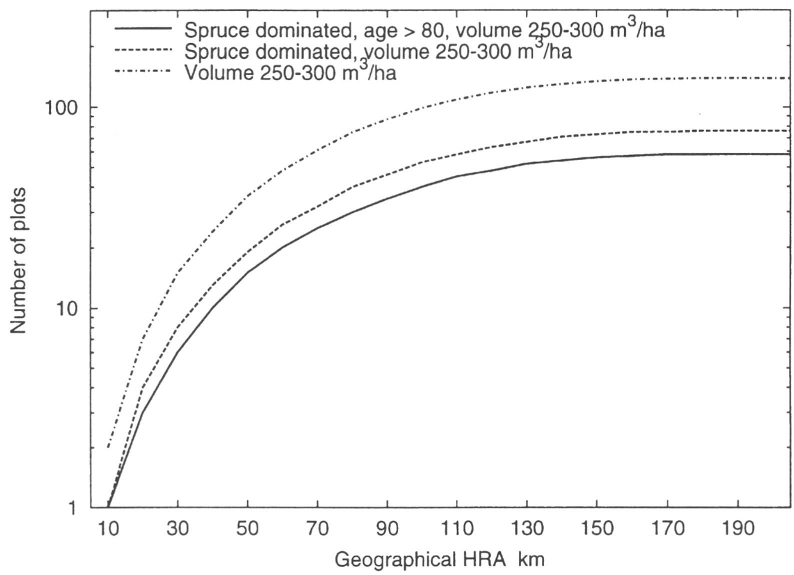


Fig. 6. The average number of 250–300 m^3/ha field plots and subsets of sample plots (logarithmic scale) in the training data in Western Finland study area when different HRA are applied.

The global relative RMSE, i.e., the percentage of the RMSE of the mean volume of the field plot data, of the pine and spruce volume estimates increased by only a few percent and the biases were not significant as the HRA increased. Naturally, there was more variation in the RMSE and in the bias of the estimates in the subareas (Fig. 8). The relative RMSE of subareas (1) and (3) increased slightly both for spruce and pine volume estimates as the distance increased from 0 to 100 km. The biases $\bar{e}$ for subareas showed a clear increase with the remoteness of the training data. The bias of the spruce volume estimate became significant for sample plots beyond 40 km geographical distance for the spruce-dominated subarea (1) ($\bar{e}=-11.4$, $s(\bar{e})=4.8$) and for subarea (3) ($\bar{e}=6.2$, $s(\bar{e})=2.6$). The Northwest subarea (2) did not produce biased estimates until the sample plots were further than 80 km; at this

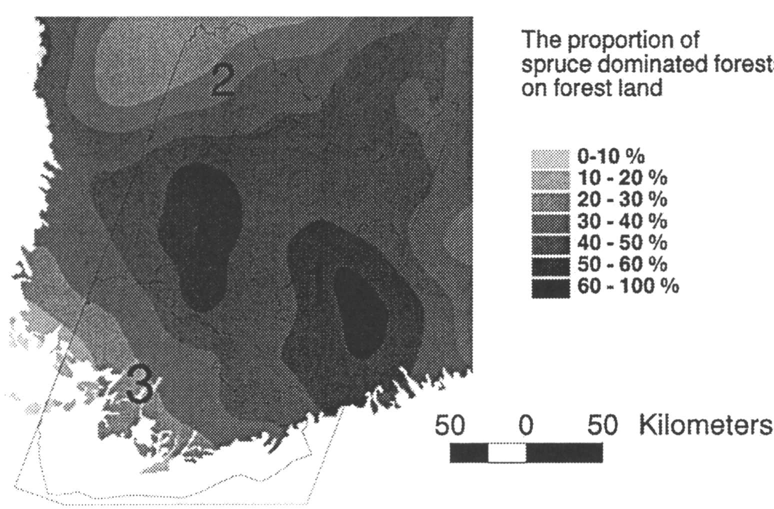


Fig. 7. The coverage of the Southwestern Finland study area and three subareas (1), (2) and (3) plotted over the map of the proportion of spruce dominated forests on forest land based on the 8th NFI field data.

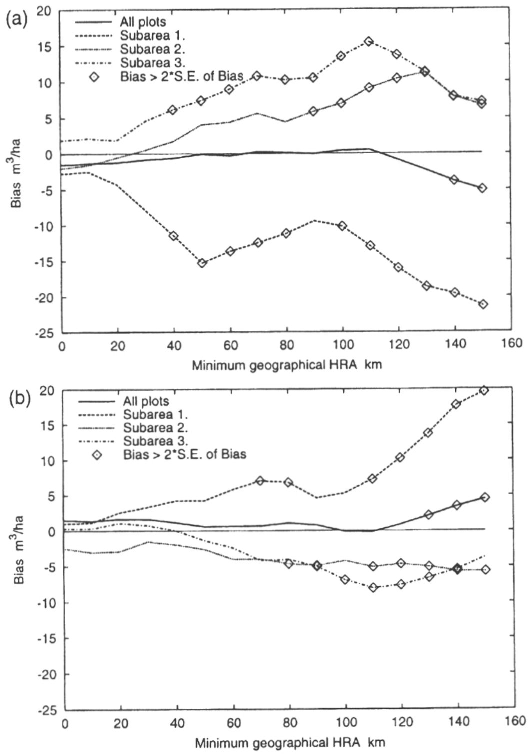


Fig. 8. Bias and the significance of bias of spruce (a) and pine (b) volume estimates for Southwestern Finland area. A total of 350 sample plot training data selected beyond various geographical distances.

distance, the pine volume estimate became significantly biased ($\bar{e} = -4.7$, $s(\bar{e}) = 2.2$). There seems to be a second minimum in the error estimates for both pine and spruce when applying the sample plots beyond a distance of 80 km. For subarea (1), there is a second wide spruce-dominated area, and for subareas (2) and (3), the training data further than from the neighbouring spruce-dominated areas is applied.

When the reference area was defined to contain all the field plots within the HRA, only the estimate of spruce volume became significantly biased; from distances 90 and 130 km upwards for subareas (1) and (2), respectively.

4.3.3. Principle of reference area selection

The number of sample plots in the training data selected with a constant geographical HRA radius varies considerably within the image area; the proportion of FRYL varies between inventory areas and within images. Near the image boundaries, there is lower number of field plots available. An alternative choice for selecting the training data was tested. Instead of a constant geographical HRA radius, a constant number of field plots HRA was employed, cf. minimum number of plots criterion (Tokola, 1998). The other MS-NFI parameters were: (1) k selected with a condition of minimum decrease of 0.5% in RMSE between k and $k+1$ and (2) stratification of field plot data according to the site class map.

In practice, the geographical HRAs were calculated for each cluster in such a way that the required constant number of field plots was approximately achieved. For example, for 300 mineral soil field plots, the geographical distance was on the average 27 km and varied between 21 and 53 km, and for 150 peatland stratum field plots, the distance was on average, 44 km and varied between 20 and 100 km within the Central Finland area.

There were only slight differences between the RMSE values of the two methods of training data selection (Fig. 9). There were no noticeable differences between the RMSE of the volume estimates for the subgroups of volume classes of the training data.

4.3.4. HRA and VRA for total volume estimates

The relative RMSE of the total volume was tested for various HRAs between 10 and 200 km. The comparison was made for all the images and also separately for both site class strata. All the other parameters, except the pixel-

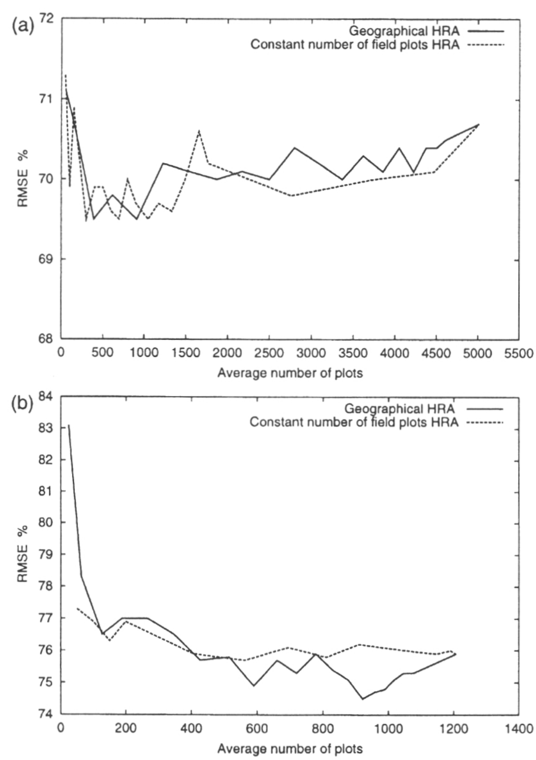


Fig. 9. Relative RMSE of total volume estimates on (a) mineral soil and (b) peatland stratum for the Central Finland study area. Training data selected with a geographical HRA and with a constant number of field plots HRA.

dependent HRA, were fixed: (1) k selected with a condition of minimum decrease of 0.5% in RMSE between k and $k+1$ and (2) stratification of field plot data according to the site class map.

With respect to the mineral soil, the RMSE of the estimates did not change much beyond geographical distances of 30–50 km (Fig. 10); concerning the number of field plots available for classification, there was only a slight decrease after 200–300 field plots. Sample plots beyond a geographical distance over 100 km or 1000 plots gave a slight increase in the RMSE for the Central Finland and Western Finland areas. The biases correspondingly had a decreasing, though nonsignificant, trend.

For the peatland stratum, the relative RMSE of the Northern Finland area was >100% and was not included in Fig. 11. The number of field plots for the peatland stratum was low for the Southwestern Finland area, only 414 (Fig. 11(b)), which may be insufficient for the estimation of other MS-NFI variables. The decrease in the RMSE of the estimates levels off in the peatland-dominated Western Finland area with a 60-km HRA. For the other two study areas, the decrease in RMSE continues over a 100-km range. The proportion of peatlands varies with the study area: the RMSE graphs are more alike when plotted against the average number of plots available. The decrease continues after 200 plots, but quite slowly. It seems that different geographical HRAs are needed for mineral and peatland strata due to different proportion of the strata. RMSEs close to the minimum level of RMSE for the volume estimate were obtained using 200–300 plots for both mineral and peatland strata, except for the Southwestern Finland area where only 50 plots were required (small proportion of peatlands).

The minimum number of plots required in training data, rather than a certain geographical HRA data, was tested by taking subsamples from the training data with $k=10$. Two less intensive sampling designs were tested: 53% (nine plots per cluster) and 18% (three plots per cluster) of the original sample for the mineral soil stratum of Western Finland area. Larger geographic HRA radii were needed for the 18% subsample (three plot clusters) (Fig. 12(a)). The minimum number of field plots required was approximately 100–200 plots for all the different field samples. The smallest subsamples seemed to benefit most from the increasing number of sample plots (from remote geographical distances) (Fig. 12(b)). This could be due to the poor variance reduction power of the sparse (18%) sample.

The altitude above sea level varies from 150 to 550 m in the Northern Finland study area (Table 2), and the VRA area was tested in addition to the HRA. The elevation variation of the terrain changes the irradiance properties of the vegetation and a simple modified Lambertian cosine correction with an exponent was used. The parameters from the operative MS-NFI were tested: a ±100 m inclusion

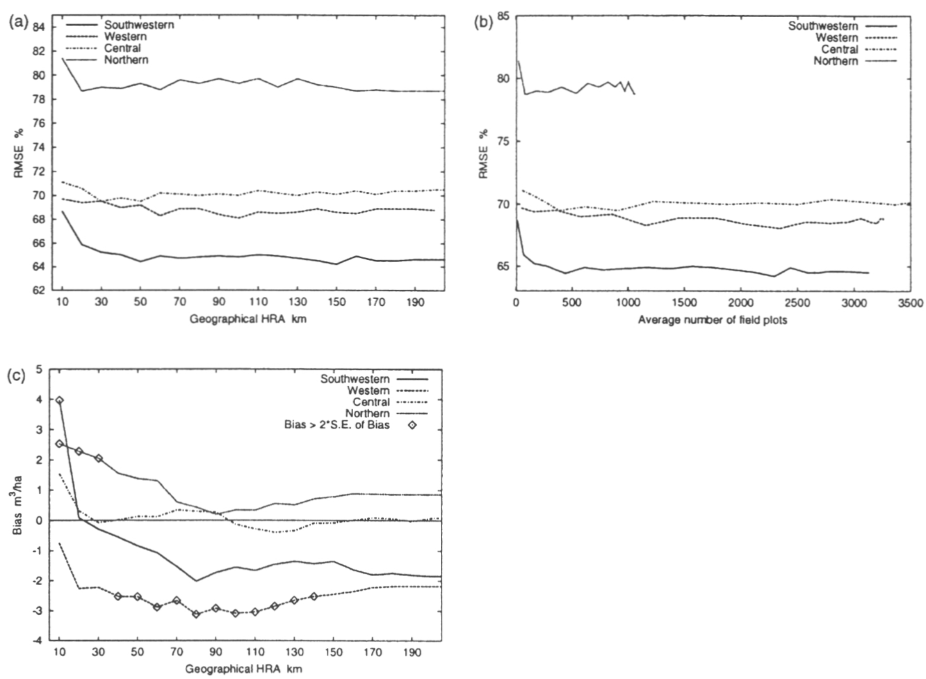


Fig. 10. Relative RMSE % ((a) and (b)), bias and significance of bias (m^3/ha) (c) of total volume estimates on the mineral soil stratum.

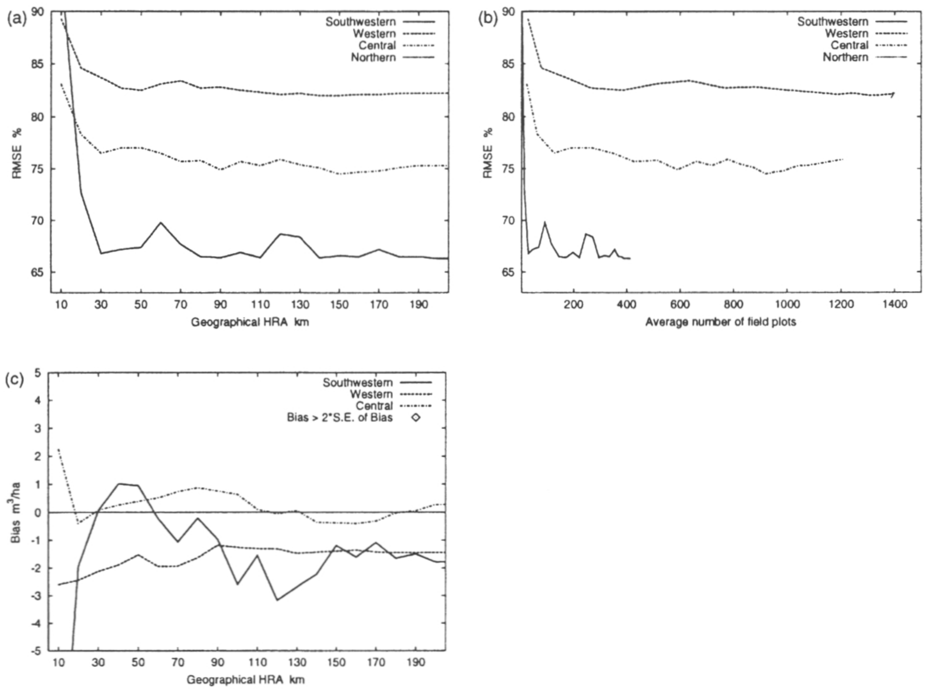


Fig. 11. Relative RMSE % ((a) and (b)), bias and significance of bias (m^3/ha) (c) of total volume estimates on the peatland stratum.

range and the normalised intensity values I^* with the exponents 0.4 and 0.8. The value of k was selected by applying a condition of minimum decrease of 0.5% in the RMSE of the estimate between k and $k+1$. Both the 100-m VRA limitation and the cosine correction with $p=0.4$ gave the largest decrease (5%) in the RMSE of the total volume estimates on the mineral soil stratum at 50-km radius of HRA. These parameters also increased the bias, but not significantly (Fig. 13).

4.3.5. HRA for estimates of volume by tree species

The dependence of the RMSE of the volume estimates were studied by tree species with the different geographical HRAs. The RMSE of the volume estimates for the three main tree species groups were tested against geographical HRAs of 10–200 km in the Central Finland (Fig. 14) and peatland-dominated Western Finland areas (Fig. 15). The other MS-NFI parameters were: (1) k selected with a condition of minimum decrease of 0.5 in RMSE between k and $k+1$ and (2) stratification of the field plot data according to the site class map.

The RMSE of the volume estimates for the two main tree species (pine and spruce) decreased to 20–30-km radii of HRA on the mineral soil stratum but did not decrease much after these distances. The relative RMSE of the deciduous tree species' volume estimates had a slowly decreasing trend, but the explanatory power R^{*2} (Eq. (7)) of the k-NN estimates was close to zero (Table 4).

On the peatland stratum, there was a greater difference in the culmination of the RMSE decrease against the HRA: 30–50-km radii for both Central Finland and Western Finland areas (Figs. 14(c) and 15(c)). Spruce volume estimates were biased in the peatland stratum in Western Finland (Fig. 15(d)); the volume of spruce varies substantially in the particular stratum (Table 4). Note that the 40-km geographical inclusion distance, on the average, gives 190 and 270 sample plots, respectively, for the training data on peatland stratum for Central Finland and Western Finland areas. The inclusion of all the field plots in the training data causes extra variation in the estimates due to numerous mixed pixels. The same graphs produced with the sample plots selected with the distance to the stand boundary >20 m had sharper changes in the RMSE.

4.3.6. The precision of the volume estimates at the pixel level

The cross-validation results of the tree species' volume estimates were compared with the sample plot statistics. The MS-NFI parameters were chosen for each tree species, stratum and test area: (1) k selected with a condition of minimum decrease of 0.5% in RMSE between k and $k+1$; (2) stratification of field plot data according to the site class

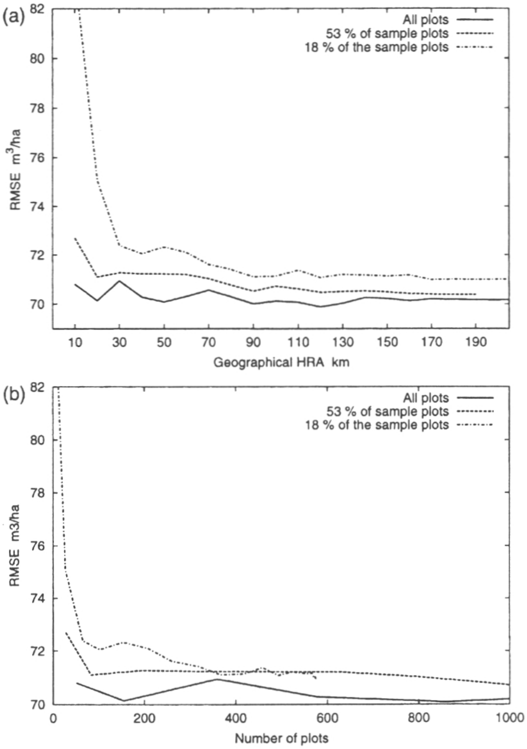


Fig. 12. RMSE (m^3/ha) on the mineral soil stratum for Western Finland study area, different densities of the original field plots.

map; and (3) the HRA and VRA parameters producing the minimum RMSE of the total volume estimate were applied for each stratum.

The mean and the standard deviation of the field plot data, and the absolute and relative RMSE, bias and the standard error of bias of the estimates, were calculated (Table 4). The RMSE of the MS-NFI estimates were compared with the standard deviation $s(m)$ of the field plot data variables. A $R*^2$ coefficient was computed to compare the predictivity, the amount of variation reduced by the "model", of different variables and methods (Eq. (7)) (cf. Tokola et al., 1996):

$$R^{*2} = 1 - \frac{\mathrm{MSE}}{s(m)^2} \tag{7}$$

As mentioned above, the volume estimates of tree species with highest volumes have the smallest relative RMSE in each strata. In the Western Finland area, two of the tree species' volume estimates are significantly biased for mineral soils, i.e., larger than twice the standard error of bias $s(\bar{e})$. This indicates that obtaining unbiased estimates for all the tree species' volume estimates is not an easy task.

The $R*^2$ coefficient varied between 0.16 and 0.42 for the total volume estimates and 0.06 and 0.46 for the dominant tree species estimates of the different stratum. Estimates for the Northern Finland peatland stratum had a poor explanatory power. There are many treeless mires, for which the variation of moisture is large and this may cause severe misclassifications.

Of all three species, the spruce volume estimates had the highest $R*^2$. The spruce volume also had the highest variation among the tree species' volumes in the field plot data. The variation of the spruce estimates is still significantly reduced, although the relative RMSE is over 100%. The $R*^2$ coefficients for pine are lower, especially when pine is not the tree species with the highest volumes. The deciduous tree species estimates have a low $R*^2$ coefficient. These species occur mostly in mixed forests with coniferous species. The early summer satellite image of the Western Finland area gave the highest $R*^2$ for the deciduous species estimates. For the peatland stratum, differences in $R*^2$ values by tree species were small. The $R*^2$ values for the mineral soil were approximately the same magnitude as those presented by Tokola et al. (1996). However, the

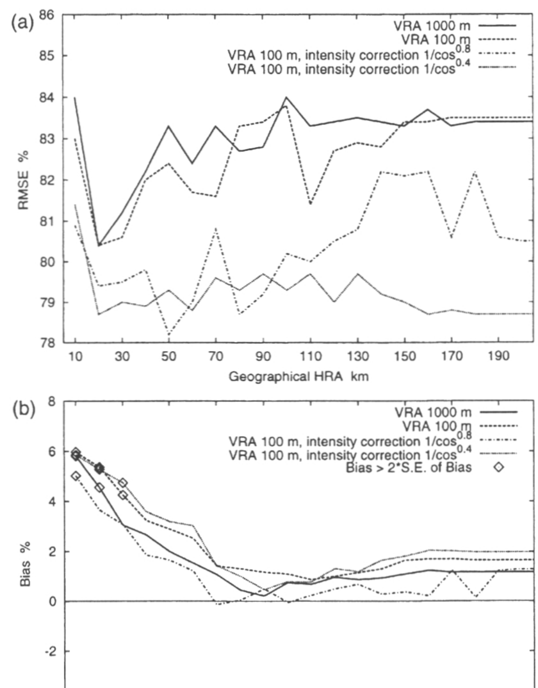


Fig. 13. Relative RMSE, bias and significance of bias for total volume estimates for the Northern Finland study area, VRA and HRA, intensity correction, mineral soil stratum.

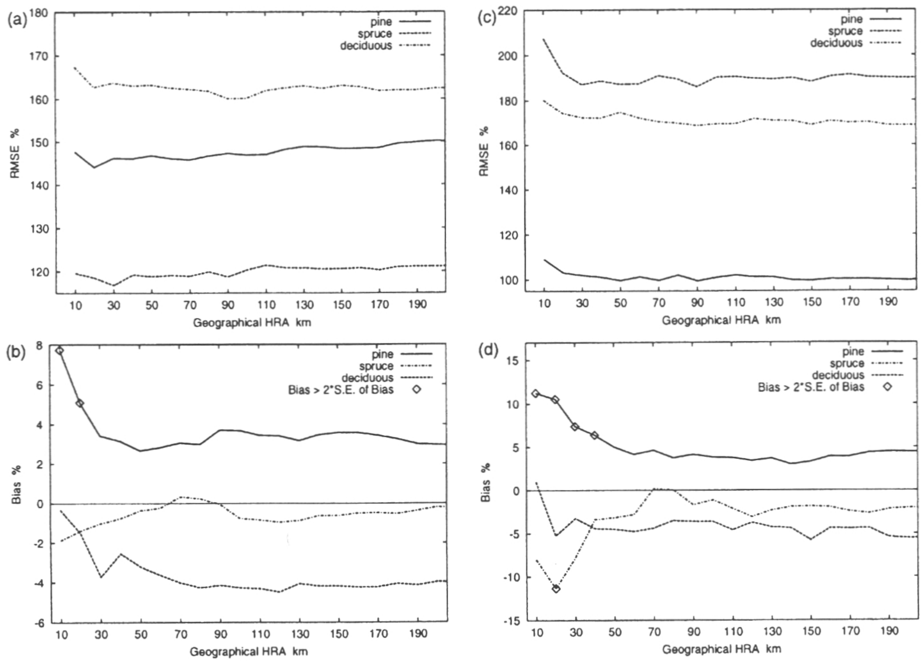


Fig. 14. Relative RMSE, bias and significance of bias for tree species volume estimates on mineral soil stratum ((a) and (b)) and peatland stratum ((c) and (d)) for the Central Finland study area.

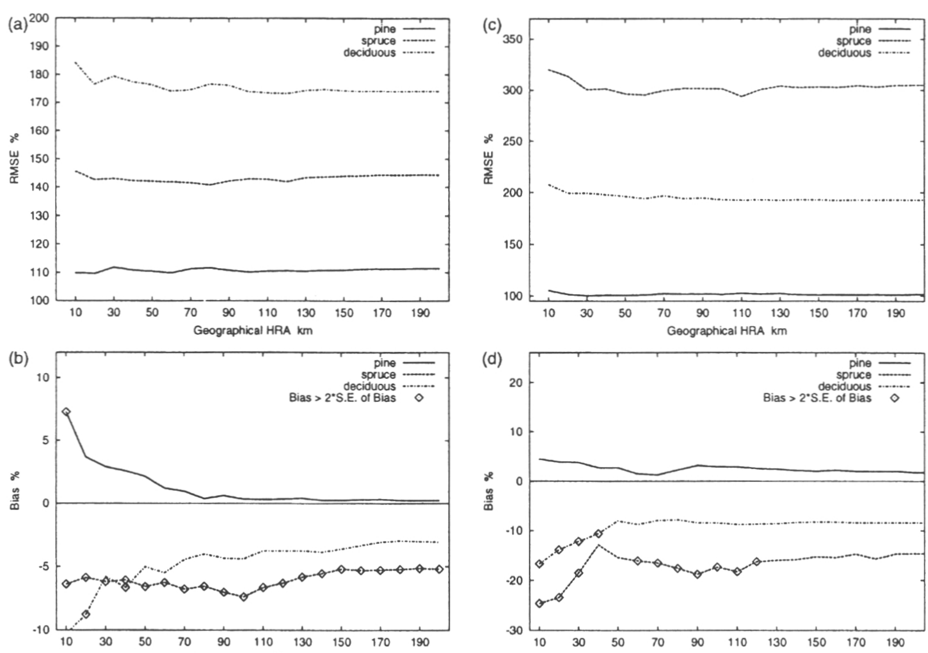


Fig. 15. RMSE, bias and significance of bias for tree species volume estimates on mineral soil stratum ((a) and (b)) and peatland stratum ((c) and (d)) for the Western Finland study area.

Table 4
The absolute and relative RMSE, bias ($\bar{e}$), the standard error of the bias ($s(\bar{e})$) for the volume estimates and variable mean ($\bar{m}$), the standard deviation of the variable ($s(m)$) and R^{*2} coefficient by study areas; stratification and different geographical HRA radius according to the site class map.

Study area	Strata/HRA	Tree species	$\bar{m}$ (m^3/ha)	RMSE (k-NN) (m^3/ha)	RMSE (%)	$\bar{e}$ (m^3/ha)	$s(\bar{e})$ (m^3/ha)	$s(m)$ (m^3/ha)	R^{*2}	No. of plots
Western Finland	mineral 40 km	total	102.4	70.7 (8)	69.0	**− 2.52**	1.24	86.3	0.33	3258
		pine	47.0	52.0 (8)	110.8	1.21	0.91	57.1	0.17	
		spruce	36.5	52.0 (6)	142.3	**− 2.22**	0.91	64.3	0.35	
		deciduous	18.9	33.5 (9)	177.3	**− 1.26**	0.59	36.0	0.13	
	peatland 60 km	total	57.9	48.1 (7)	83.1	− 1.94	1.28	62.4	0.40	1403
		pine	38.1	38.4 (9)	101.0	0.55	1.03	44.4	0.25	
		spruce	7.4	21.9 (5)	295.4	− 1.19	0.58	27.4	0.36	
		deciduous	12.4	24.2 (10)	194.3	− 1.09	0.64	27.8	0.25	
Central Finland	mineral 40 km	total	121.7	84.9 (9)	69.8	0.02	1.20	103.6	0.33	5012
		pine	37.1	54.2 (9)	146.1	1.17	0.77	57.1	0.10	
		spruce	60.7	72.3 (8)	119.2	− 0.46	1.02	88.4	0.33	
		deciduous	23.9	39.0 (9)	163.0	− 0.61	0.55	39.3	0.02	
	peatland 70 km	total	85.1	64.4 (10)	75.7	0.74	1.85	76.0	0.28	1208
		pine	40.7	40.6 (10)	99.6	1.89	1.17	44.2	0.16	
		spruce	25.9	49.5 (8)	190.9	0.04	1.42	56.3	0.23	
		deciduous	18.5	31.5 (9)	170.5	− 0.81	0.91	32.6	0.07	
Southwest Finland	mineral 50 km	total	136.0	87.5 (11)	64.4	− 0.84	1.56	104.2	0.30	3132
		pine	42.3	56.6 (10)	134.0	1.83	1.01	58.8	0.07	
		spruce	71.8	75.5 (10)	105.2	− 1.92	1.35	90.8	0.31	
		deciduous	21.9	38.3 (11)	174.7	− 0.78	0.68	40.1	0.09	
	peatland 90 km	total	100.4	66.6 (9)	66.4	− 0.96	3.28	87.6	0.42	414
		pine	39.5	37.3 (10)	94.6	1.43	1.84	45.3	0.32	
		spruce	41.7	53.3 (8)	127.8	− 1.23	2.62	72.3	0.46	
		deciduous	19.3	37.1 (9)	192.9	− 1.25	1.83	39.3	0.11	
Northern Finland	mineral 50 km	total	43.2	34.0 (9)	78.8	1.31	0.90	43.2	0.38	1428
		pine	27.3	33.2 (8)	121.6	1.15	0.88	39.1	0.28	
		spruce	10.5	18.1 (10)	172.6	0.17	0.48	21.3	0.28	
		deciduous	5.4	11.1 (8)	206.3	− 0.05	0.29	11.6	0.08	
	peatland 80 km	total	16.3	24.2 (8)	148.9	− 0.27	1.00	26.4	0.16	585
		pine	8.8	17.0 (8)	192.2	0.24	0.70	17.5	0.06	
		spruce	4.3	12.4 (6)	292.2	− 0.21	0.51	12.9	0.08	
		deciduous	3.2	8.1 (9)	254.7	− 0.19	0.33	9.1	0.22	

stratification was not applied in that study and the NFI sample data were from smaller areas.

5. Discussion

The selection of the appropriate geographical HRA and VRA and the effect of stratification of the field plots and images on the basis of digital site class map for the Finnish MS-NFI method were studied applying RMSE and biases of volume estimates at pixel level. The leave-one-out cross-validation method was used to obtain average RMSE and biases of estimates for pixels. The main findings with the applied test data were: (1) The stratification of the satellite image and the field plot data with the site class map significantly decreased the bias of the volume estimates. (2) The geographical HRA radius of 40–50 km on mineral soil stratum included a sufficient number of field plots (400–600) for different variable combinations, and to minimise the RMSE of the volume estimates with the current sampling intensity in the NFI. Field plots from larger distances increased the bias of the volume estimates in the image subareas. (3) For the peatland stratum, covering a minor part of the FRYL, larger HRA radii (60–90 km) were needed. (4) The VRA, together with normalised intensity values, decreased the global relative RMSE of the total volume estimate on the mineral soil stratum in Northern Finland.

5.1. Sources of error in the training data

Errors in the field measurements and in the location of the field plots, location errors of the pixels, imaging system errors and atmospheric condition errors cause extra variation in the estimates (Curran & Hay, 1986; Tomppo et al., 1999a). The location errors decrease the precision of pixel level estimates of the cross-validation in two ways: both the pixel to be analysed and the field plot pixel may possess location errors. Even if there was no location error, the size of the NFI field plot measured with PPS sampling is much smaller than the area of a single pixel (625 m^2); the trees of d.b.h. 10 and 20 cm are tallied from area of 39 and 157 m^2, respectively, when a basal area factor of 2 is applied. The smaller the average size of the trees, the smaller is the area of the forest stand covered by the field plot.

5.2. Distance metric and the value of k

The original digital values of channels 1–5 and 7 of Landsat 5 TM were used as image features, although improvement in the precision of the estimates could have been expected with transformations or weighting of the original channels (Franco-Lopez et al., 2000; Tokola et al., 1996; Tomppo et al., 1999a, 1999b). The Euclidean distance measure was used in the feature space. The weighting of spectral distances would be expected to reduce the bias in the estimates (Altman, 1992). Weights between 0 and 2 have been tested in other studies when optical area satellite data and point sampling or concentric circular field plots material have been applied (Nilsson, 1997; Poso et al., 1999; Tokola et al., 1996; Tomppo, 1991). The weighting of distance with $t=1-2$ was found to give a smaller RMSE of volume estimates than nonweighted distances, especially for smaller HRAs (<50 km). The weighting with $t=1$ gave slightly better results on mineral soils and was chosen for this study.

Two objectives have been kept in mind when selecting the value of k: (1) minimising of the RMSE of the estimates of the key variables and (2) retaining the original variation of the field plot data in the spatial variation of the estimates. These objectives conflict to some extent, the RMSEs of the estimates decrease slightly until the value of $k=20-30$, wherefore, e.g., Nilsson (1997) and Tokola et al. (1996) suggested a value $k=10-15$. For example, for mapping or for forest planning purposes, Franco-Lopez et al. (2000) and Moeur and Stage (1995) suggested a much smaller value, even the value $k=1$, which retains the variation of the original data. A compromise is necessary for practical inventories. Weighting of contradicting objectives, e.g., the RMSE and retaining the variation, would be needed to select the value k in an analytical way. The weighting, however, depends on for what purposes the estimates are used. A heuristic rule has therefore been applied in selecting the value of k in the Finnish MS-NFI. The value has usually been between 5 and 10 (Tomppo, 1996).

A moderate value of k can also be argued by the fact that the stratumwise biases may increase when the value of k increases. A value higher than 1 for k, on the other hand, can be argued by the fact that the area of a NFI sample plot is smaller than the area of a pixel. The field plot data involves also theoretically more variation than pixel level data should involve.

Selecting the value of k with minimum decrease of 0.5% in RMSE between k and $k+1$ led to values of $k=5-12$. A constant $k=10$ was also used. The value selection for k was not very sensitive to the number of observations in the training data when $n>100$.

When the global RMSE criterion is used the appropriate choice of k depends on several parameters: (1) number of sample plots in the training data; (2) size of the field plots compared to the pixel size; (3) weighting of the spectral distance in the estimation, a higher distance weight reduces the importance of the last neighbours; and (4) the density of the training data in spectral space.

5.3. Stratification

The stratification, which applied the site class map (peatlands and mineral soils), significantly decreased the bias of the volume estimates (Fig. 4), although the NLS's peatland delineation is different from that of the NFI. Maps often underestimate the area of peatlands and they also contain location errors (Tomppo et al., 1998b). The global RMSE of the combined total volume estimates only improved by 1%. Contrary to expectations, the use of the precise sample plot data for the stratification did not significantly improve the global RMSE of the volume estimates compared to the map based site class stratification. Tokola and Heikkilä (1997) obtained a 5% reduction of the global RMSE in the pixel level estimates of total volume with a stratification based on forest site quality maps and NFI data when using an estimation method similar to the one used here.

The stratification of low radiometric resolution satellite data with the auxiliary data, which is correlated with the estimated variables, will most often reduce the bias of the estimates within the strata. The stratification can help to avoid mismatches in the classification of certain type of forests, e.g., peatlands and mineral soils or old-growth forest stands. However, the minimum number field plots in the training data for each stratum must be maintained (Tomppo et al., 1999a).

5.4. Geographical reference area

In the presence of spatial trends in the forest variables, pixel-dependent geographical HRAs of radius 10–200 km were tested for the selection of training data for the Finnish MS-NFI. The RMSE and the bias of the volume estimates based on the cross-validation were calculated separately for the two strata. Different HRAs were required for the peatland and mineral soil strata due to the different proportion of the strata (Table 2). On the mineral soil map, a suitable HRA was 40–50 km for total volume estimate, although most of the variance reduction was already gained at a distance of 20 km (Fig. 10). On the peatland stratum, the suitable HRA varied more, from 60 to 90 km, and the variance reduction was slower than in the case of mineral soils (Fig. 11). When the average number of sample plots in the training data was studied instead of the HRA, a near minimum RMSE was achieved with 150–300 sample plots. However, increasing the number of sample plots to between 400 and 500 decreases the error of both strata. Increasing the HRA to

100 km and over did not increase the global relative RMSE of the total volume estimate, except for two images on mineral soil stratum.

The RMSE of the volume estimates for spruce and pine reached a local minimum with a slightly smaller HRA radius than that for the total volume. As the RMSEs were not significantly higher, if the HRA was the same as with the total volume, the same radius could be used.

The number of sample plots is more important than the geographical distance when subsets of the original sampling design were selected for the field plot data (Fig. 12). However, when the estimation was made with the training data containing sample plots only beyond a 50–60-km or larger HRA radius, the spruce and pine volume estimates were significantly biased in the subareas of the Southwestern Finland study area (Fig. 8). These radii are comparable with results presented by Tokola (1998).

It can be concluded that in the boreal forests of Finland, a 40–50-km geographic HRA radius for mineral soil is suitable, depending on the intensity of sampling. This also yields a reasonable amount of variation in the sample plot data (400–600 sample plots) for subclasses of variables (Fig. 6) and does not lead to significant biases within the subareas of an image. Nilsson (1997) found the same number of sample plots sufficient for a total volume estimation on the FRYL with simulated forest map and Landsat TM data.

On the average, the peatlands account for 26.6% of the FRYL in southern and 40.6% in northern Finland (Finnish Forest Research Institute, 1999), that is to say, less than the mineral soils. Thus, a geographical HRA radius of 60 km for peatland-dominated areas and 90-km radius for areas with low peatland proportion is recommended for the Finnish MS-NFI. However, if the average amount of sample plots in the training data falls below 300, stratification may not be an appropriate method.

The VRA, together with normalised intensity values, decreased the global relative RMSE of the total volume estimate on the mineral soil stratum in Northern Finland (Fig. 13). Changes in altitude have a clear impact on the vegetation in Northern Finland (Seppälä & Rastas, 1980). The VRA distance of ± 100 m alone did not, however, affect the global RMSE of volume estimates, but from earlier experience, it is known that it decreases the bias of the estimates in vertical subareas.

The RMSEs of the volume estimates were high at the pixel level, but seem to be of the same magnitude for the same strata for the different study areas and satellite images (Table 4). The worst estimates were obtained for the mixed pixels, i.e., those near stand boundaries. The explanatory powers as measured by R^{*2} increased to over 0.5 for the total volume estimates when the sample plots near the stand boundaries were omitted from the cross-validation (cf. Tokola & Kilpeläinen, 1999).

The estimates for the surveyed area will be biased, if there is a locational dependency in the spectral values of pixels within the HRA (Kilkki & Päivinen, 1987). Kilkki and Päivinen (1987) proposed the use of the same (locationally uncorrelated) training data for each pixel of the surveyed forest area. In the Finnish MS-NFI, a fixed size moving HRA is applied and artificial boundaries are avoided (Tomppo, 1991). The locational dependencies for FRYL within the HRA satellite images are quite difficult to model explicitly due to the complexity of imaging systems, atmospheric attenuation and target reflectance properties. In the operative inventory, the global and local unbiasedness of the estimates were checked using the cross-validation method and large area forest statistics prior to the classification. A knowledge of the range and shape of vegetation cover changes (Fig. 1) has been used to define the appropriate form and size of the reference areas (rectangular or circle) (Tomppo, 1996).

In this context of local unbiasedness, obtaining a variable number of field plots in the training data for each pixel, due to image boundaries and proportion of FRYL, seems to have only minor effects on the precision of the MS-NFI results within the particular image. There were no significant differences in the global RMSE and bias of the total volume estimates between the selection of training data with a constant number of sample plots HRA or using geographic HRA (Fig. 9). The study of the biases by subareas also failed to reveal significant differences between the two training data selection methods.

These results do not cover the problem which is present particularly when trend-like large-scale changes occur in forests. The small area estimates are highly dependent upon how the area is located with respect to the applied satellite image. The satellite images obtained for the inventory area of a certain year form an image 'mosaic.' Since each satellite image will be analysed separately, neighbouring pixels, or small areas, may employ training data from a different geographical reference area depending on how the area is located with respect to the applied satellite image.

Other possible ways to define the geographic HRAs could be a combination of VRA and HRA — an ellipsoid, or the mean effective temperature sum of thermal season and vegetation zones as a surface use of 7 h. However, because the forests in Finland are not in a natural state, the pure edafic and climatic factors may only partly explain the location dependent variation in the forests. In addition, the silvicultural regimes vary between forest owner groups — private, state and companies.

New and enhanced map data, e.g., soil and bedrock maps, could be studied for stratification purposes in the future, since there will be more digital map data available. The Finnish MS-NFI is proceeding in its 9th cycle and the independent 8th MS-NFI estimation result could also be tested for stratification purposes. More directly, the successive MS-NFI image cover intensity values could be used as multitemporal features.

Acknowledgments

We thank Dr. Juha Heikkinen, Dr. Helena Henttonen, Lic. Kai Mäkisara and two anonymous referees for their valuable comments. The English language was edited by Dr. Ashley Selby. The study was partly funded by the Graduate School Program "Forests in GIS", University of Helsinki.

References

Altman, N. S. (1992). An introduction to kernel and nearest-neighbour nonparametric regression. *American Statistician*, *46*, 175–185.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*, 21–27.

Curran, P. J., & Hay, A. M. (1986). The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, *52*, 229–241.

Eisele, F. A. (1997). Practical use of the multi-source forest inventory system under central european conditions. *Managing the resources of the world's forests, Lectures given at the 1997 Marcus Wallenberg Prize Symposium, 14 October 1997* (pp. 71–86) Stockholm, Sweden, The Marcus Wallenberg Foundation, Falun, Sweden.

Finnish Forest Research Institute. (1999). *Finnish statistical yearbook of forestry.* Helsinki, Finland: Finnish Forest Research Institute.

Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2000). Estimation and mapping of forest stand density, volume and cover type using the *k*-nearest neighbors method. *Remote Sensing Environment*, (in press).

Gjertsen, A. K., Tomter, S., & Tomppo, E. (2000). Combined use of NFI sample plots and Landsat TM data to provide forest information on municipality level. In: T. Zawila-Niedzwiecki, & M. Brach (Eds.), *Remote sensing and forest monitoring, Proceedings of IUFRO conference, 1–3 Jun. 1999* (pp. 167–174) Rogow, Poland, Office for Official Publications of the European Communities, Luxembourg.

Hagner, O. (1997). Textur i flygbilder för skattning av beståndsegenskaper. Working Report 29. Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics, Umeå, Sweden (in Swedish).

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

Keller, J. M., Gray, M. R., & Givens, J. E. Jr. (1985). A fuzzy *k*-nearest neighbor algorithm. *IEEE Transaction on Systems, Man and Cybernetics*, *15*, 580–585.

Kilkki, P., & Päivinen, R. (1987). Reference sample plots to combine field measurements and satellite data in forest inventory. In: *Remote sensing-aided forest inventory. Proceedings of seminars organised by SNS, Hyytiälä, Finland, Dec. 10–12, 1986* (pp. 209–215) University of Helsinki, Department of Forest Mensuration and Management, Research Notes No. 19, Helsinki, Finland.

Linton, O., & Härdle, W. (1998). Nonparametric regression. In: S. Kotz, C. B. Read, & D. L. Banks (Eds.), *Encyclopedia of statistical sciences, update vol. 2* (pp. 470–485). New York: Wiley.

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.

Poso, S., Wang, G., & Tuominen, S. (1999). Weighting alternative estimates when using multi-source auxiliary data for forest inventory. *Silva Fennica*, *33*, 41–50.

Rao, J. N. K. (1998). Small area estimation. In: S. Kotz, C. B. Read, & D. L. Banks (Eds.), *Encyclopedia of statistical sciences, update vol. 2* (pp. 621–628). New York: Wiley.

Seppälä, M., & Rastas, J. (1980). Vegetation map of the northernmost Finland with special reference to subarctic forest limits and natural hazards. *Fennia*, *158*, 41–61.

Tokola, T. (1998). Critical aspects in establishment of satellite remote sensing based regional forest inventory system. PhD thesis, Faculty of Forestry, University of Joensuu, Research notes 77.

Tokola, T., & Heikkilä, J. (1997). Improving satellite image based forest inventory by using a priori site quality information. *Silva Fennica*, *31*, 67–78.

Tokola, T., & Kilpeläinen, P. (1999). The forest stand margin area in the interpretation of growing stock using Landsat TM imagery. *Canadian Journal of Forest Research*, *29*, 303–309.

Tokola, 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. (1987). An application of a segmentation method to the forest stand delineation and estimation of stand variates from satellite images. In: *Proceedings of the 5th Scandinavian conference on image analysis* (pp. 253–260) Stockholm.

Tomppo, E. (1991). Satellite image-based national forest inventory of Finland. In: *Proceedings of the symposium on global and environmental monitoring, techniques and impacts, 17–21 Sept. 1990.* Victoria, British Columbia, Canada, *International Archives of Photogrammetry and Remote Sensing*, *28*, 419–424.

Tomppo, E. (1992). Satellite image aided forest site fertility estimation for forest income taxation purposes. *Acta Forestalia Fennica*, *229*, 1–70.

Tomppo, E. (1996). Multi-source national forest inventory of Finland. In: R. Vanclay, J. Vanclay, & S. Miina (Eds.), *New thrusts in forest inventory. Proceedings of the subject group S4.02-00 'Forest Resource Inventory and Monitoring' and subject group S4.12-00 'Remote Sensing Technology', vol. 1, IUFRO XX World Congress, 6–12 Aug. 1995* (pp. 27–41) Tampere, Finland, European Forest Institute, Joensuu, Finland.

Tomppo, E., Goulding, C., & Katila, M. (1999a). Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. *Scandinavian Journal of Forest Research*, *14*, 182–192.

Tomppo, E., Henttonen, H., Korhonen, K. T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., Mikkelä, H., Tonteri, T., & Tuomainen, T. (1998a). Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968–97. *Folia Forestalia*, *2B*, 293–374 (in Finnish).

Tomppo, E., Henttonen, H., Korhonen, K. T., Aarnio, A., Ahola, A., Heikkinen, J., & Tuomainen, T. (1999b). Pohjois-Savon metsäkeskuksen alueen metsävarat ja niiden kehitys 1967–96. *Folia Forestalia*, *2B*, 389–462 (in Finnish).

Tomppo, E., & Katila, M. (1992). Multi-source national forest inventory of Finland. *International Archives of Photogrammetry and Remote Sensing*, *29*, 671–674.

Tomppo, E., Katila, M., Moilanen, J., Mäkelä, H., & Peräsaari, J. (1998b). Kunnittaiset metsävaratiedot 1990–94. *Folia Forestalia*, *4B*, 619–839 (in Finnish).

Tomppo, E., Varjo, J., Korhonen, K., Ahola, A., Ihalainen, A., Heikkinen, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S., & Tuomainen, T. (1997). Country report for Finland. In: *Study on European forestry information and communication systems. Reports on forestry inventory and survey systems, vol. 2* (pp. 145–226). Luxembourg: Office for Official Publications of the European Communities.

II

Error variations at the pixel level in the k–nearest neighbour estimates of the Finnish multisource National Forest Inventory

Matti Katila
Finnish Forest Research Institute
Unioninkatu 40 A, FIN-00170 Helsinki, Finland

Abstract

The paper examines the residual variation in the k–nearest neighbour (k–NN) estimates of the Finnish multisource National Forest Inventory (MS-NFI). In the MS-NFI, field plots, satellite images and digital maps are utilised. The prediction errors at single pixel level for field plot volume and weighted mean of basal area (BA) observations were studied applying leave-one-out cross-validation method and potential explanatory covariates were seeked. The standard deviation of the field plot variable from the k neighbours was a good measure of the estimation uncertainty. The first principal component (PC1) of the Landsat TM or ETM+ channel values of the field plot pixel had a strong relation to the volume and BA estimates and to the prediction error. The residual variances of field plot volume were regressed against PC1 and the model was used to remove the trend component of PC1 from the residuals. The random error component still remained high in the residuals.

Keywords: cross-validation, k–nearest neighbours, Landsat TM and ETM+, multisource forest inventory, prediction error

1. Introduction

In recent years, there has been a growing interest in obtaining national forest inventory results for small areas, i.e. for municipalities and forestry holdings (Schreuder et al. 1993, Tomppo 1996, Kangas 1996), and even for forest stands (Tomppo 1987). For this purpose, auxiliary data is required in addition to sparse field measurements. In multisource inventories, remote sensing and numerical map data is combined with the forest inventory data to obtain estimates of forest variables for single pixels (Tomppo 1996, Tokola et al. 1996, Nilsson 1997, Franco-Lopez et al. 2001). Since 1989, the Finnish multisource National Forest Inventory (MS-NFI) has utilised optical area high resolution satellite images, numerical map data and field plot data to produce thematic maps and forestry statistics for municipalities. All the forest variables can be estimated simultaneously by employing a non-parametric k–nearest neighbour (k–NN) method. A large number of field plots is required because the training data should cover the range and variation present in the inventory area (Tomppo 1996, Katila & Tomppo 2001).

Multisource inventory methods involve several sources of error because they combine measurement data and models of different nature and scale. In the MS-NFI, the data at each step is produced by an explanatory model or standardised rule: the land use classes are defined by certain rules, volume models are employed for sample trees, the satellite imagery exo-atmospheric radiances are calibrated to digital numbers using linear models.

There have been various attempts to represent the spatial variation of the classification error. Error maps have been produced by employing extrapolation of errors from the training data set (Steele et al. 1998), magnitude and partitioning of class membership in fuzzy classification (Zhang & Foody 1998) and geostatistical approaches to model the variation in accuracy (DeBruin 2000).

In the k–NN estimation, the overall error is minimised by tuning the estimation parameters. Error quantification methods include resampling techniques such as leave-one-out cross-validation and bootstrap methods (Katila & Tomppo 2001, Franco-Lopez et al. 2001). The numerous error sources increase the uncertainty in the MS-NFI estimates. The prediction errors, described with relative RMSE for mean volume estimates at the field plot level, have been high, 50–80 %, and the proportion of explained variation in the field plot data has been 30–40 % (Tokola et al. 1996, Katila & Tomppo 2001).

The spectral channel values of Landsat TM and ETM+ satellite images contain little variation in the well-stocked stands (Ardö 1992). It might be expected that variation in the estimates increases, as the volume of the target (field plot) increases. If there is a functional dependence between observable covariates and the prediction error, a model can be estimated for the given form of heteroskedasticity, c.f. heteroskedastic linear regression models (Polasek et al. 1998).

The objective of this paper is to study the residual variation in the k–NN estimation and to determine whether there is a functional dependence between the residuals and covariates or other exogenous variables. In addition, some suggestions are made for reducing the random error in the k–NN estimates. This paper is one step in deriving an analytical method for estimating the error of multisource estimates from pixel level to region level. The next phase will be finding suitable models to estimate the k-NN estimation error, taking into account the spatial dependencies of the errors. The explanatory variables should be such that their values can be obtained for every pixel. Potential explanatory variables are target field plot pixel values, the estimated values of forest variables and the variables of the selected k–NN field plot pixels (forest and spectral variables). A simple empirical error estimation model is tested for the MS-NFI data. The leave-one-out cross-validation method is employed for the error prediction at the single pixel level. The behaviour of the prediction error is studied in a realistic setting created by two geographically different study areas in Finland.

2. Material

The two study areas are located between longitudes 21°40′E and 31°36′E and latitudes 61°21′N and 63°50′N (Fig. 1). The test data contains field measurements from the 9th NFI and satellite image data from the same years (Table 1). The Western Finland study area contains large peatland areas and the Eastern Finland study area consists largely of medium fertile mineral soils (Table 2). The forests of the study areas are characterised by Scots pine (*Pinus sylvestris* L.) or Norway spruce (*Picea abies* (L.) Karst.), mixed with birch (*Betula* spp.) and other deciduous species.

The NFI field samples were measured from systematically located clusters of sample plots. The sample plots (10–18 per cluster) were located along a rectangular or L-shape tract at 250 or 300 m intervals, depending on the area. The average

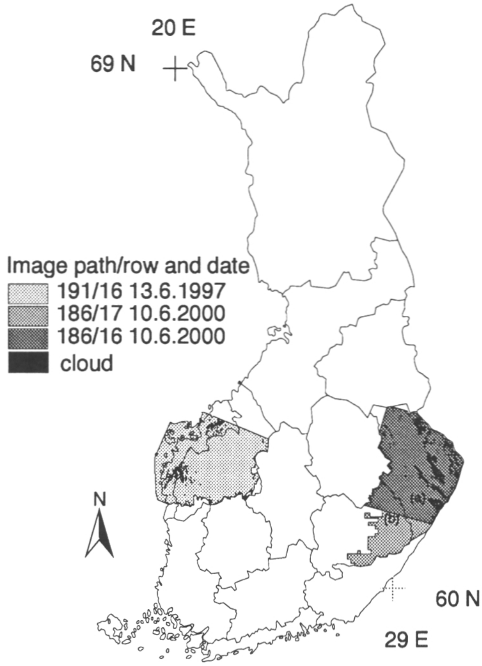


Fig. 1. Location of the study areas and path, row and time of acquisition of the Landsat 5 TM and Landsat 7 ETM+ images employed.

Table 1. Study areas: satellite images of the 9th MS-NFI and field inventory data.

	Field plot data					Satellite image	
study area	land area (km^2)	forestry land (%)	year	cluster distance (km)	plots per cluster	Landsat path/row	date
Western Finland	13920	72.7	1997	7 × 7	18/14 (a)	Ls5 TM 191/16	13.6.1997
Eastern Finland/north	14660	88.8	2000	7 × 7	18/14 (a)	Ls7 ETM+ 186/16–17	10.6.2000
Eastern Finland/south	6670	86.9	2000	6 × 6	14/10 (b)	Ls7 ETM+ 186/16–17	10.6.2000

(a) Every fourth cluster had 14 field plots.
(b) Every fourth cluster had 10 field plots.

field plot location error has been estimated to be 20 m (Halme & Tomppo 2001). This is caused by map error and field plot location applying precision compass and tape. Trees were measured on parts of field plots belonging to forest and other wooded land (FOWL) stands. The tally trees were selected with PPS-sampling (sampling with probability proportional to size), applying a relascopic factor of two. The probability of a tree's inclusion was proportional to its cross-sectional area at a height of 1.3 m; a maximum radius of 12.52 m was used. The distance of the nearest forest stand boundary from the field plot centre point was recorded by 10 m classes from 0 to 40 m and the bases for stand delineation were recorded, e.g. land use class, site class, development class, tree species composition and storey and completed drainage (Tomppo et al. 1998, 2001).

In addition to field plot measurements, three basal area (BA) observations were made on forest land, in the stand to which the field plot is located. The first BA measurement was made from the field plot centre point, if the BA observation was not transected by another stand. The two other BA measurements -or all three- were made at a distance of 20 m to the field plot centre, preferably from two of the four main cardinal directions (Fig. 2). The basal area factor two was applied. If a field plot was cut by a stand or land use class boundary, the entire plot was considered to consist of two or more parts. The BA observations were made on each field plot part (stand) belonging to forest land: more precisely, on the field plot centre part and on the other field plot parts, where there were tallied trees.

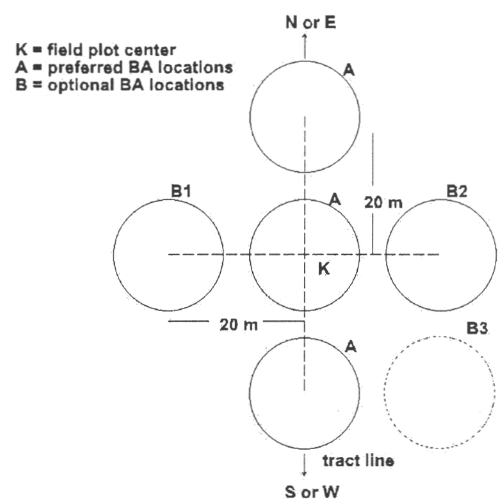


Fig. 2. Location of three basal area measurement points on the field plot stand.

For field plot parts belonging to forest land, a weighted mean of the three BA observations G_{obs} in each field plot part was calculated to better estimate the BA on the area of a pixel ($625m^2$). If there was a BA observation from the field plot centre point, it was given a weight of 0.5, while the other two observations were given a weight 0.25. Otherwise equal weights were employed. For other wooded land, the BA estimated by the crew leader was used. Finally, the BA estimates from each field plot part were combined for the whole field plot, weighted by the proportions of the assessed areas of the field plot parts. It must be noted that the BA observations contain measurement error because the border trees are not checked and observations are biased towards the centre of the stands.

The field plots that were totally on forestry land (FRYL) were selected from the NFI field sample for the following analyses. They were divided into forest land, other wooded land and waste land, according to site productivity (Table 2) (Tomppo et al. 1997). The mean and the standard deviation of the volumes of the field plot measurements and weighted mean of BA observations on forest and other wooded land are presented in Table 3.

Table 2. The land use class distribution of the forestry land field plots and the proportion of field plots on peatland and mineral soil strata according to the site class map, by study areas.

study area	forest land (%)	other wooded land (%)	waste land (%)	forestry land plots	peatland stratum (%)	mineral soil stratum (%)
Western Finland	89.5	5.7	4.8	4829	39.4	60.6
Eastern Finland	95.3	2.4	2.3	7492	28.0	72.0

Table 3. The mean and the standard deviation of the volume of the growing stock and the weighted mean of basal area (BA) observations of the field plot parts on forest and other wooded land by study areas.

variable		Western Finland		Eastern Finland	
		$\bar{y}$	s	$\bar{y}$	s
Volume	(m^3/ha)	93.7	82.0	110.0	94.6
BA	(m^2)	13.8	9.3	15.5	10.0

The Landsat 5 TM and Landsat 7 ETM+ satellite images were rectified to the national grid coordinate system using regression models of the first order polynomials calculated from 53 to 71 ground control points. These were identified from topographic maps and satellite images. The RMSE of the rectification model from the panchromatic image data, together in the direction of rows and columns, was 0.63 and 0.18 satellite image pixels (30×30 m^2) for the Landsat 7 images 186/16 and 186/17 respectively. The RMSE for Western Finland Landsat 5 image was 0.55 pixels. The channels 1–5 and 7 from Landsat 5 and all the eight channels of Landsat 7 ETM+, including the thermal and panchromatic channels, were used in the k–NN estimation. Nearest neighbour resampling was used with a pixel size of 25×25 m^2 for all the channels (Tomppo 1996).

A multi-criteria procedure to reassign the satellite image information to the field plot data was employed in the Western Finland study area (Halme & Tomppo 2001). A weighted function of the correlation coefficients of the selected image and field variables is used as a scaling function in the multicriteria optimisation. This procedure reduces the effect of the locational errors on the training data and decreases the prediction errors, particularly for the total volume estimates.

A topographic correction for the digital number (DN) values of satellite image spectral channels was carried out using a modification of the Lambertian surface reflectance assumption (Tomppo 1996).

3. Methods

3.1. Multisource National Forest Inventory estimation method

In the operative MS-NFI, multisource estimates are computed for FRYL pixels. Cloud-free FRYL areas of a satellite image are analysed using the FRYL field plots i chosen for the training data set. Field plots with uncertainty concerning their location, and those that contain non-FRYL land use classes, are excluded from the training data set; the excluded proportion is usually in the range of 2–6 %.

The MS-NFI estimates are weighted averages of the field plot variables. The k–NN method is used to calculate the weights (Keller et al. 1985, Tomppo 1991). Data from the k nearest field plots, $i_1(p), \ldots, i_k(p)$, in the feature space are utilised in the analysis of each pixel p. The field plots are sorted according to distance $d_{p_i,p}$ between field plot pixel p_i and p in the image feature space, and the k nearest plots are then chosen.

Stratification of the FRYL area and the training data to peatlands and mineral soils, according to a numerical map data, has usually been executed in such a way that only pixels within the same stratum as the target pixel are accepted as neighbours. The horizontal geographical reference area (HRA), i.e. the maximum geographical distance to the potential nearest neighbours, has been restricted to 40 to 90 km due to gradual changes of vegetation type and, is selected by image.

The weight $w_{i,p}$ of the field plot i for estimating the value for the pixel p is defined as

$$w_{i,p} = \begin{cases} \frac{1}{(d_{p_i,p}+c)^t} / \sum_{j \in \{i_1(p),...,i_k(p)\}} \frac{1}{(d_{p_j,p}+c)^t}, & \text{if } i \in \{i_1(p),...,i_k(p)\} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\{i_1(p),...,i_k(p)\}$ is the set of the field plots whose corresponding pixels are the k nearest neighbours to the pixel p. Here, a value $t = 1$ is applied for the weighting parameter (Katila & Tomppo 2001) and an arbitrary constant $c = 1$ is added to the Euclidean distances to smooth the weighting of 0 distances.

The weight $w_{i,p}$ can be interpreted as the share of the pixel p that obtains data from the field data vector of plot i. For a single pixel p, the estimate of the average of a continuous variable is,

$$\hat{y}_p = \sum_{i \in FRYL} w_{i,p} \cdot y_i, \tag{2}$$

where y_i are the values of variables in the training data set.

3.2. Results validation and parameter selection

The choice of estimation parameters, k and geographical HRA radius, was tested using the leave-one-out cross-validation method: a single field plot pixel p_i belonging to the ground truth data set was estimated with the other plots (Linton & Härdle 1998).

The root mean square error has been used as a measure of reliability of the continuous variables (eq. 3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}, \tag{3}$$

The estimates of biases and the standard error of biases have been used as further criteria (Katila & Tomppo 2001). Residuals, $e_i = \hat{y}_i - y_i$, of the main field plot variables were produced for each observation i in the training data from the cross-validation.

3.3. Indicators of estimation uncertainty from field plot variables

The values of total volume and weighted mean of BA observations for the selected k neighbours were studied in each validated field plot. The standard deviation of the k neighbours' field plot values for the variable estimated was used to evaluate the uncertainty of the k–NN estimate of each observation. The greater the deviation between the neighbours, the greater is the expected prediction error , cf. class membership probabilities from statistical classification (Canters 1997).

The average sampling error of the field plot BA G_i estimated from tally trees used to estimate the BA of a larger area (a pixel) was studied using the difference between the field plot centre point BA and the weighted mean of BA from the three observations, $G_{obs} - G_i$. For a single field plot, a high difference indicates that the field plot measurement differs from the average BA of the surrounding forest stand.

3.4. Variables describing the spatial neighbourhood and the Euclidean feature space neighbourhood

The edges in the spatial neighbourhood of the field plot were studied employing differences of field plot pixel and the pixel values in the surrounding 3×3 window, the number of non-FRYL pixels in the 3×3 window (from the numerical map data) and the magnitude of maximum change in pixel values in the 5×5 window of Landsat 7 Pan image of the Eastern study area, as defined by Sobel gradient operator (Gonzales & Woods 1993). The principal component transformation was applied to the field plot pixel spectral channel values of the study areas. The transformation was made on the covariance matrix. In Table (4), the first principal component (PC1) contains 81 % of the variation in the spectral channel values in the Western Finland study area. PC1 is a weighted sum of all bands and Horler & Ahern (1986) call it the spectral brightness-type feature.

The k–NN estimates may be biased at the edges of the spectral feature space because the k–NN method cannot extrapolate beyond the observations in the training data. Suitable variables for describing the spatial distribution, direction and clus-

Table 4. The eigenvectors for the principal components, reassigned training data, Western Finland study area.

Eigenvector	TM channel						Variation explained
	1	2	3	4	5	7	%
1	0.14	0.11	0.18	0.34	0.83	0.36	81.4
2	-0.05	0.00	-0.12	0.93	-0.25	-0.23	15.1
3	0.83	0.29	0.39	0.03	-0.28	0.02	2.2
4	-0.41	0.13	0.48	0.09	-0.37	0.66	0.7
5	-0.31	0.18	0.68	-0.04	0.17	-0.61	0.4
6	-0.16	0.93	-0.34	-0.05	0.01	-0.02	0.1

tering of the k neighbours in the feature space were tested.

One measure of the spatial distribution of the neighbours around the pixel value to be estimated was obtained by dividing the spectral feature space into two half spaces applying a hyperplane that goes through the field plot pixel spectral vector $\mathbf{p}_i$. The difference between the number of nearest neighbours in the the half spaces was calculated. The hyperplane $H_i = \{\mathbf{p}_i \in R^n \mid \langle \mathbf{w}_i, \mathbf{p}_i \rangle \leq a\}$ divides the feature space R^n to open half spaces (Fig. 3). The normal vector $\mathbf{w}_i$ that defines the hyperplane and is perpendicular to it was obtained by subtracting the target field plot pixel spectral vector $\mathbf{p}_i$ from the spectral value vector $\hat{\mathbf{p}}_i$ estimated for the field plot by the k–NN method. In this way, the number of nearest neighbours was expected to be distributed as unevenly as possible into the halfspaces.

The polar coordinates of the k neighbouring spectral values from the target field plot pixel were calculated on a plane formed by the two first principal components of the training data set. The mean of the angles $\theta_{i,j}$ between the adjacent neighbours was used as a measure of the spatial distribution of the neighbours.

Other variables in the feature space of the k neighbours were the distance to the first nearest neighbour, the standard deviation of the nearest neighbour distances and the Euclidean distance $d_{p_i,\hat{p}_i}$ from the true pixel value to the channel values estimated with k–NN to the field plot pixel.

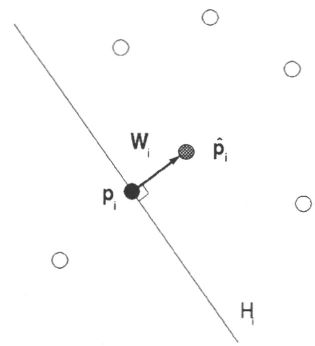


Fig. 3. Hyperplane H_i defined by normal vector $\mathbf{w}_i$ and field plot pixel spectral vector $\mathbf{p}_i$.

3.5. Specific error models

It was soon noticed that the PC1, the spectral brightness value, of the target pixel is a dominating explanatory variable for the residual variation in the k–NN estimates. The pixels with low PC1 values obtained the highest estimates and also the highest absolute residuals in the k–NN estimation. An attempt was made to remove the effect of the spectral brightness value from the k–NN volume estimate residuals obtained from cross-validation. The effect of PC1 on the volume residual variance was modelled by assuming the variance to be a multiplicative function of two components

$$e_i^2 = f_i \delta_i^2 \qquad (4)$$

with the trend component f_i estimated from $\mathrm{E}(e_i^2) = f(PC1)$. Generalized linear models were estimated employing Poisson regression with a logarithmic link function. The $\mathrm{E}(e_i^2)$ follows Poisson distribution and the log transformation is used to adjust for the skewness in the Poisson distribution. The PC1 trend was removed from the field plot volume residuals and the remaining variance component $\delta_i = e_i/\sqrt{f(PC1)}$ was studied against the potential explanatory variables.

4. Results

4.1. The selected parameters for k–NN estimation

The estimation parameters for the k–NN method were selected on the basis of the pixel-level estimates of the variables. The goal was to obtain accurate volume and BA estimations in the two strata. The pixel-dependent geographical HRAs, which were found to be optimal in the earlier study, were used because a sufficient number of field plots should remain in the training data (Katila & Tomppo 2001). k values of 5 for the peatland stratum and 10 for the mineral stratum were chosen for estimating the field plot volume and BA. These values were near the ones obtained in the earlier study (Katila & Tomppo 2001) and were considered sufficiently large because the obtained RMSEs for the field plot volume were, in the Western Finland study area, only 1 % and 6 % larger than the minimum RMSE values for the mineral soil and peatland strata respectively. The significant global bias of the volume estimate in the Eastern study area was not considered a major problem in this study. Most of the results presented hereafter have been estimated with the reassigned training data (Halme & Tomppo 2001) that includes all the FRYL field plots not intersecting stand boundaries. The prediction errors, and the employed estimation parameters for the FRYL field plots on mineral soil and peatland strata, for the Eastern Finland (with minimum distance of 30 m to the stand boundary) and Western Finland study area (using the reassigned training data) are summarised in Table 5.

4.2. Visual inspection of residuals

In the following, the residual pattern figures are presented mostly for the Western Finland study area mineral soil stratum. Figures from other study areas and strata are also presented if the residual patterns are notably different.

There is a negative correlation between field plot volume and most of the reflectance values of the Landsat satellite images. The low dynamic range of Landsat image spectral channel values on FRYL, the large amount of noise, small size of field plot and other type of errors caused considerable variation in the scatter plot of the field plot volume and PC1 of the spectral channel values, although the locational errors in the training data were reduced by reassigning the spectral values to the field plots (Halme & Tomppo 2001) (Fig. 4).

Table 5. The absolute and relative RMSE and bias ($\overline{e}$), the standard error of the bias ($s(\overline{e})$) for the weighted mean of basal area (BA) observations and field plot volume estimates, mean ($\overline{y}$) and the standard deviation ($s(y)$) of the variable and R^{*2} coefficient. Stratification and different geographical horizontal reference area (HRA) radius according to the site class map, forestry land field plots with minimum distance $\geq$ 30 m to the nearest stand boundary (Eastern Finland study area) and forestry land field plots not intersecting stand boundaries, reassigned training data (Western Finland study area).

study area	strata /HRA	variable	$\overline{y}$	RMSE (k-NN)	RMSE	$\overline{e}$	$s(\overline{e})$	$s(y)$	R^{*2}	No. plots
Eastern	mineral	BA (m^2/ha)	16.0	5.9 (10)	36.9 (%)	-0.20	0.18	10.9	0.71	1026
Finland	50 km	volume (m^3/ha)	114.3	67.3 (10)	58.8 (%)	**-4.35**	2.10	101.8	0.56	
	peatland	BA (m^2/ha)	10.9	4.8 (5)	43.7 (%)	0.36	0.24	9.0	0.72	393
	80 km	volume (m^3/ha)	64.4	44.4 (5)	68.9 (%)	1.96	2.24	69.9	0.60	
Western	mineral	BA (m^2/ha)	14.5	5.0 (10)	34.6 (%)	-0.11	0.10	10.2	0.76	2768
Finland	40 km	volume (m^3/ha)	98.8	45.1 (10)	45.8 (%)	-1.41	0.86	87.9	0.74	
	peatland	BA (m^2/ha)	9.5	4.0 (5)	42.3 (%)	-0.18	0.11	8.5	0.78	1235
	60 km	volume (m^3/ha)	54.1	30.1 (5)	55.6 (%)	-1.57	0.85	61.6	0.76	

Note: Significant bias is printed in bold font.

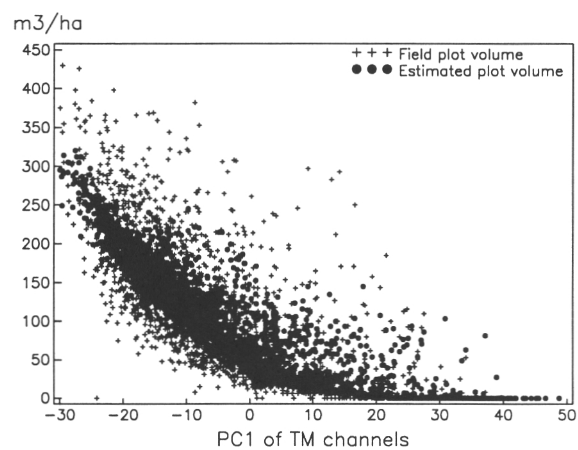


Fig. 4. Field plot volume and estimated volume plotted against the PC1 of channel values for the field plot pixel. Western Finland study area, mineral soil stratum, reassigned training data, k= 10, 40 km geographical horizontal reference area.

The k–NN estimation (k =10) of the field plot volume reduced much of the variation, but averaged the results; high volume estimates are missing (Fig. 4). Employing a small value of k would slightly decrease the shrinkage towards mean, and would globally better preserve the original variation in the field plot data, c.f.

Franco-Lopez et al. (2001). On the other hand, this may yield a RMSE value larger than the standard deviation of the observations (McRoberts et al. 2002). The estimated volume and the variation of the residuals had a relatively strong correlation (Fig. 5).

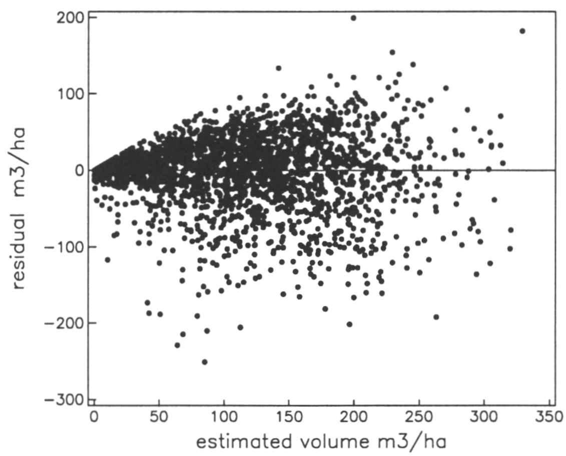


Fig. 5. Residuals e_i of the field plot volume estimate plotted against the estimated field plot volume $\hat{y}_i$. Western Finland study area, mineral soil stratum, reassigned training data, k= 10, 40 km geographical horizontal reference area.

Field plot variables

The standard deviation of the k–neighbours' field plot values correlated with the residuals of the variable to be estimated, field plot volume and weighted mean of BA observations (Figures 6b and 6d). The average of the differences of field plot BA and the weighted mean of BA from the three observations, $G_{obs} - G_i$, in the selected k neighbours correlated weakly with the field plot volume residuals only at the extreme values of neighbours' average $G_{obs} - G_i$ (Fig. 7b).

Spectral variables

The highest residuals for the estimates of field plot volume and weighted mean of BA observations occurred at the low end of the PC1 values (Fig. 6a and 6c). In the spatial neighbourhood of the field plot pixels, the variation between the centre pixel and the surrounding pixels is related to the spectral brightness (PC1) of the pixel value. The highest volume residuals had the lowest variation in the field plot

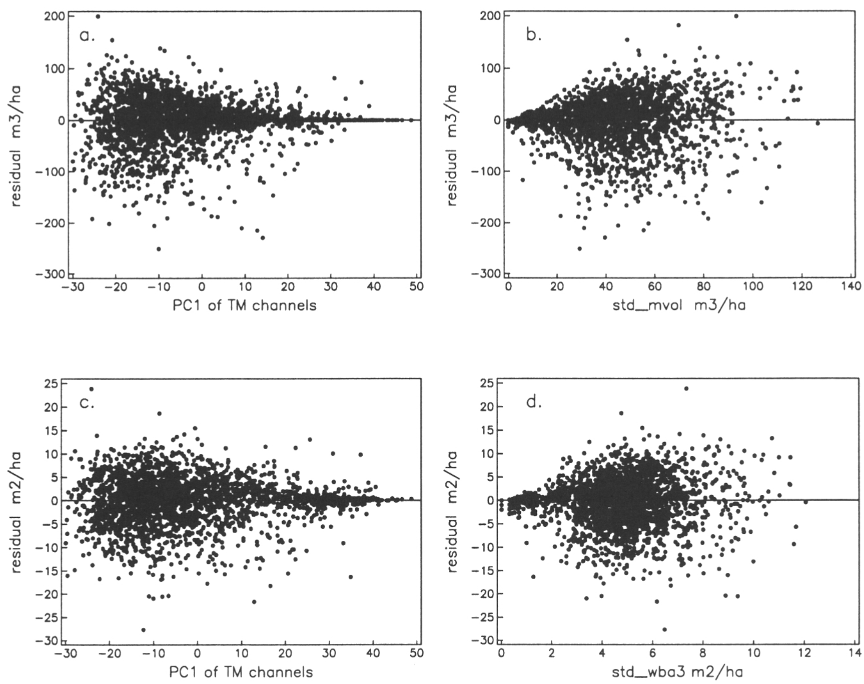


Fig. 6. Residuals e_i of the field plot volume estimate and the first principal component of the spectral channel values (PC1) (a) and standard deviation of the k neighbours' field plot volume (std_mvol) (b), residuals e_i of weighted mean of the basal area observations estimate and PC1 (c) and the standard deviation of the k neighbours' weighted mean of BA observations (std_wba3) (d). Western Finland study area, mineral soil stratum, reassigned training data , k= 10, 40 km horizontal reference area.

pixel values as well as in the spatial neighbourhood. Consequently this variable was not useful.

The delineation of stand boundaries in the field is often defined by criteria other than those visible on the Landsat PAN images, e.g. tree species composition or site index. However, the stand boundaries with other land use classes and between different development classes obtained high Sobel gradient magnitude values. On average, the k–NN volume estimates were biased downwards on the field plots with high edge magnitude. Apart from this trend, there was no clear dependence between the field plot volume estimate residuals and Sobel gradient magnitude (Fig. 7a).

The nearness of non-FRYL indicated by the non-FRYL map pixels in the 3×3 window caused systematic bias in the field plot volume estimates in the two study

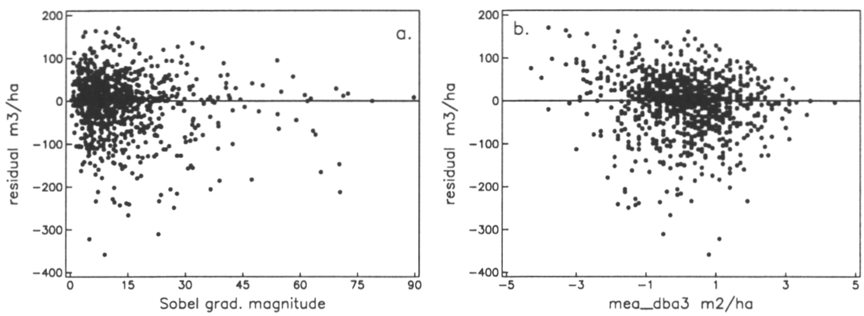


Fig. 7. Residuals e_i of the field plot volume estimate and Sobels edge gradient from a 5×5 window of Landsat ETM+ Pan channel (12.5×12.5 m^2) (a) and the mean of k neighbours' $G_{obs} - G_i$ (mea_dba3) (b). Eastern Finland study area, mineral soil stratum, minimum distance $\geq$ 30 m to the nearest stand boundary, k= 10, 50 km horizontal reference area.

areas. The field plot volume was underestimated with nearness of other land use mask pixels, e.g. agricultural or built up land. However, when the number of other land use pixels increased to over six pixels, the bias disappeared, partly due to the decreased volume of the target field plots. The few field plots close to water obtained overestimates of volume (Table 6).

Table 6. The average residual $\bar{e}$ of the field plot volume estimate and the number of non-forestry land pixels in a 3×3 window according to numerical map data, Western Finland study area, mineral soil stratum, all forestry land field plots.

	other land		water	
no. of pixels	n	$\bar{e}$ (m^3/ha)	n	$\bar{e}$ (m^3/ha)
0	2893	0	3210	-3
1 – 3	295	-23	39	24
4 – 6	45	-28	9	55
7 – 9	27	2	2	193
All	3260	-2	3260	-2

The distance between estimated and the true channel values $d_{p_i,\hat{p_i}}$ (Fig. 8a) was not correlated with the residuals of volume estimates and weighted mean of BA observations estimates. The distance $d_{p_i,\hat{p_i}}$ was correlated both with the distance to the first nearest neighbour and the standard deviation of the nearest neighbour

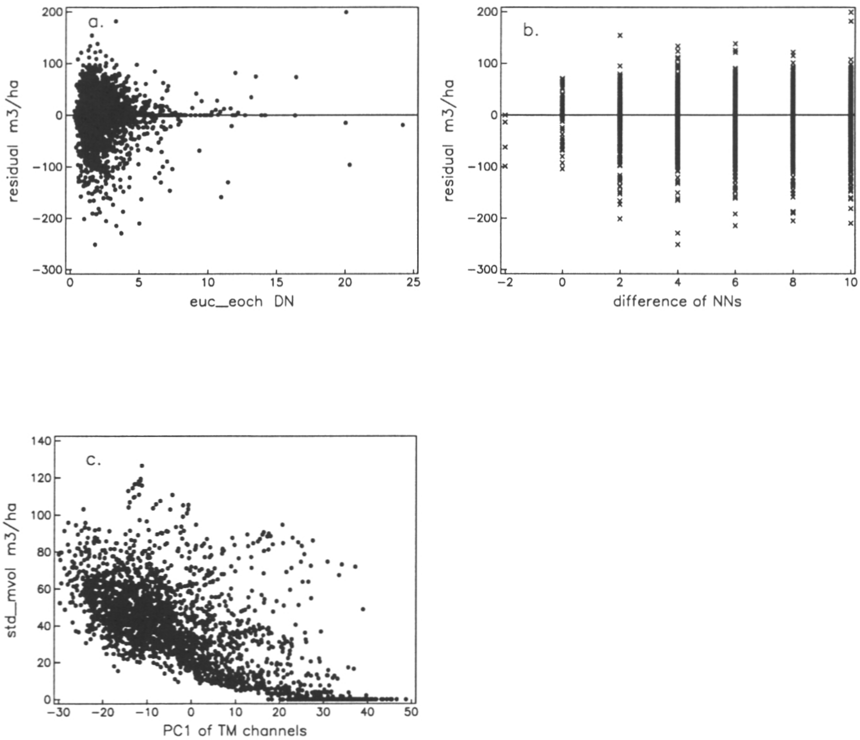


Fig. 8. Residuals e_i of the field plot volume estimate and the Euclidean distance between target field plot pixel and k–NN estimated spectral channel values (euc_eoch) (a), the difference in the number of neighbours between two halves of feature space (b) and the first principal component of the spectral channel values (PC1) and standard deviation of the k neighbours' field plot volume (std_mvol) (c). Western Finland study area, mineral soil stratum, reassigned training data , k= 10, 40 km horizontal reference area.

distances. These distance measures were, in turn, weakly correlated with PC1; with low PC1 values, the neighbour distances are small.

The difference in the number of nearest neighbours between half spaces and the small polar coordinate angles between nearest neighbours should indicate an uneven spatial distribution of the nearest neighbours in the feature space. However, there was no clear dependence between the number of nearest neighbours in the halfspaces (Fig. 8b) or the average polar coordinate angle $\theta_{i,j}$ of adjacent nearest neighbours and the residuals of volume or weighted mean of BA observations.

4.3. Specific error models

The spectral brightness value (PC1) of the target pixel was a dominating explanatory variable in the error analysis: low PC1 value pixels obtained the highest estimates and also had the highest variation in the residuals of the k–NN estimation. Possible explanatory variables based on the spectral channel values in the spatial neighbourhood or in the nearest neighbours contained little variation at the low PC1 field plot pixel values. For example, the standard deviation of the k neighbours' field plot volume was strongly correlated with PC1 of the spectral channel values (Fig. 8c).

Generalized linear models were used to regress the residual variance against PC1 of spectral features and its transformations. The aim was to remove the effect of the spectral brightness value from the volume residuals of the cross-validation. The Poisson regression models had a significant goodness of fit and parameter standard errors. The models captured the average trend between PC1 and the variance of the volume residuals.

The models explained most of the trend in error variance correlated with PC1 (Fig. 9a). The variation of the residual component δ_i had the strongest correlation with the standard deviation of the k neighbours' field plot volume (Fig. 9b).

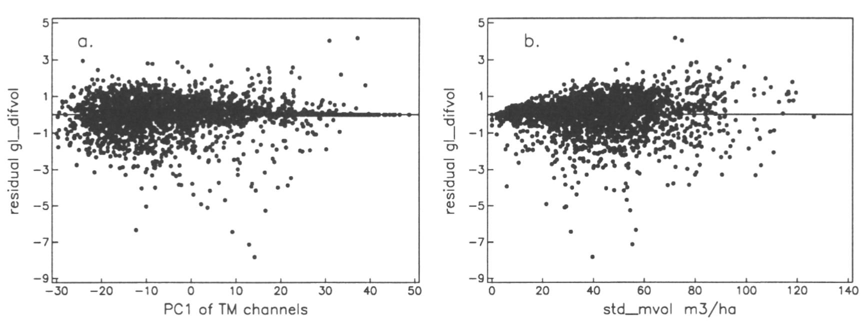


Fig. 9. Residual component δ_i of the field plot volume estimate with variance component of first principal component (PC1) removed by Poisson regression (gl_difvol) and PC1 of the spectral channel values (a) and standard deviation of the k neighbours' field plot volume (std_mvol) (b). Western Finland study area, mineral soil stratum, reassigned training data , k= 10, 40 km horizontal reference area.

5. Discussion

The residual variation of the k–NN estimates of the field plot volume and the weighted mean of the BA observations was studied employing the prediction errors from leave-one-out cross-validation. The appropriate estimation parameters, e.g. k and geographical HRA, were selected as a compromise between minimising the overall error and retaining some of the original variation in the field plot data in the spatial variation of the estimates (Katila & Tomppo 2001). Potential variables explaining the variation in prediction error were sought, based on knowledge of the error components in the MS-NFI estimation. The standard deviation of the k neighbours' field plot variable was a good measure of the estimation uncertainty and was correlated with the k–NN estimates of the variable. The nearness of the non-FRYL map mask increased the bias in the estimates. The spectral brightness of the field plot pixel (PC1) had a strong relation to the volume and BA estimate, and to the uncertainty of the k–NN estimate. The prediction errors were higher at the lower end of the spectral brightness values, as the correlation between the field variables and the remote sensing variables weakened. In a variance model of the field plot volume residuals, the PC1 value of the field plot pixel explained most of the non-random variation.

In the resampling methods used to estimate the prediction error, the observations at the edges of the feature space obtained neighbours from one direction only. However, the variables describing the spatial distribution of the k neighbours in the feature space did not clearly correlate with the volume or BA residuals in this study. Although this error component was not very distinct, an advanced non-parametric method could remove a part of this error, e.g. symmetrized k–NN estimator (Linton & Härdle 1998) or local adaption of non-parametric methods (Malinen 2003).

The k–NN estimates of forest stand border pixels have larger bias than those inside the stands (Tokola & Kilpeläinen 1999). Thematic map errors frequently occur at patch boundaries and are associated with the misregistration of map data and mixed pixels (Foody 2002). The errors in the pixel-level estimates are often spatially correlated (Congalton 1988, Flack 1995). In this study, only field plots with a minimum distance of 30 m to the stand boundary or field plots not intersecting stand boundaries (reassigned training data) were used. If all the FRYL field plots were applied in the cross-validation, the prediction error variation in the results would be higher and the dependencies between residuals and explanatory variables would weaken. The effect of the neighbouring pixel values on the estimation errors was analysed using the DN values of Landsat TM and ETM+, Sobel gradient magnitude from Landsat 7 Pan images and the number of non-FRYL map pixels in the 3×3 window. Only the map data was demonstrated to be useful in the error detection (Table 6). Katila & Tomppo (2002) applied MS-NFI by map strata, an idea that is supported by these results. Furthermore, new strata should be formed to estimate separately the boundary pixels of water, other land use and FRYL.

The Euclidean distances of field plot pixels in the feature space were not directly related to the differences in the field plot variable values (Fig. 4). A distance measure related to the variation in the field plot variable might be more easily interpreted, e.g. Tokola et al. (1996) employed differences between the regression estimates of forest stand characteristic.

Since the spectral brightness value of the field plot pixel was correlated with the residual variation and also with the other explanatory variables, an attempt was made to remove this trend from the residuals using a variance model (eq. 4). Poisson regression models were fitted to the residual variance and PC1. Although the parameters of the models were significant, separate models for the low and high values of PC1 might have worked better. The k–NN estimates themselves, e.g. from the produced thematic map data, can be used in posterior analyses of uncertainty in the estimates. The estimated volume could be employed as a dependent variable and modelling of the error variance could occur after the estimation.

Explanations of the magnitude and direction of residuals seemed to be case sensitive. When the field plot values and the potential explanatory variables were studied together with a display on the numerical map data and the remote sensing data, several explanations for the error presented themselves: mislocation of the field plot, the radiation from the surrounding land use classes or stands, the deviation of the target field plot from the surrounding forest and extreme field plot variable values (e.g. BA 40 m^2/ha or greater).

By reducing the main sources of error in the MS-NFI, e.g. in the field plot data, it should be possible to decrease the random error in the k–NN estimates. Reducing the effect of the field plot location error in the training data decreases the RMSE values of mean volume estimates obtained from the cross-validation (Halme & Tomppo 2001). It also corrects the typical shrinkage towards the mean in the k–NN estimates and better preserves the original variation of the field plot data in the spatial variation of the estimates. The use of BA observations from an area larger than a field plot decreased the random variation in the training data; $s(y)/\overline{y}$ provides the coefficient of variation in Table 5. The relative RMSE of the weighted mean of BA observations was 10 to 25 percentage points lower than the relative RMSE of the field plot volume estimates (Table 5).

In the MS-NFI, cross-validation has been applied assuming independent sampling. If the prediction errors from cross-validation are spatially correlated the parameters obtained may favour undersmoothing (Altman 1990). A solution is to apply 'leave-some-out' cross-validation (Linton & Härdle 1998) or to modify the cross-validation procedure (Altman 1990).

Although a larger field plot size employing weighted mean of BA observations and the reassignment of the training data removed some of the sampling error and the locational error, the random error component remained considerable in the estimates of this study. Since the larger k–NN estimates had also a larger residual variation and variation in the the selected nearest neighbours, it might be possible to decrease the prediction error by applying stronger smoothing to the pixels where high volume estimates will be produced. Again, a local adaptation of the non-parametric methods could be used, based on the selected nearest neighbours.

New very high resolution satellite data and high altitude aerial photographs are becoming increasingly available for remote sensing purposes. These data can be used to survey the location and the representativeness of a field plot by detecting forest stand edges and mixed pixels.

A future research task in the development of the MS-NFI method is to develop a reliable method for estimating the error at the pixel level and a method to derive error estimates for small areas. The error estimates obtained for single pixels cannot be directly combined to estimate the error in larger areas due to locational errors in the field plot data and the spatial autocorrelation both in the satellite image and field data. The error variance of the MS-NFI for small areas could be estimated by employing models describing the second order properties of the MS-NFI error estimates obtained from cross-validation for pixels (Lappi 2001). However, the field plot volume prediction error of the MS-NFI estimates not only depends on distance between pixels but, e.g. on the true volume. In addition, the k–NN prediction errors may not be treated as the residuals of a trend surface of a spatial model. The various sources of error in the MS-NFI can reduce the reliability of the spatial modelling of errors.

Acknowledgements

I thank prof. Erkki Tomppo and Ph.D. Juha Heikkinen and the pre-examiners, Prof. Michael Köhl and Ph.D. Ronald McRoberts for their valuabe comments and suggestions on the manuscript. The English language was edited by Ph.D. Ashley Selby. The study was partly funded by the Graduate School Program "Forests in GIS", University of Helsinki.

References

Altman, N. S., 1990. Kernel smoothing of data with correlated errors. J. Am. Stat. Assoc. 85:749–758.

Ardö, J., 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. Int. J. Remote Sens. 13(9):1779–1786.

Canters, F., 1997. Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification. Photogrammetric Engineering and Remote Sensing 63(4):403–414.

Congalton, R. G., 1988. Using spatial autocorrelation analysis to explore errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54:587–592.

DeBruin, S., 2000. Spatial uncertainty in estimates of the areal extent of land cover types. In: Heuvelink, G. B. M. & Lemmens, M. J. P. M. (eds.), Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 12–14July 2000, Amsterdam, The Netherlands. De Rode Hoed, Amsterdam, pp. 137–144.

Flack, J. C., 1995. On the Interpretation of Remotely Sensed Data Using Guided Techniques for Land Cover Analysis. Ph.D. thesis, Curtin University of Technology, School of Computing.

Foody, G. M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80:185–201.

Franco-Lopez, H., Ek, A. R. & Bauer, M. E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote Sens. Environ. 77:251–274.

Gonzales, R. & Woods, R., 1993. Digital Image Processing. Addison-Wesley Publishing Company, Reading (MA).

Halme, M. & Tomppo, E., 2001. Improving the accuracy of multisource forest inventory estimates by reducing plot location error – a multi-criteria approach. Remote Sens. Environ. 78:321–327.

Horler, D. N. H. & Ahern, F. J., 1986. Forestry information content of thematic mapper data. Int. J. Remote Sens. 7(3):405–428.

Kangas, A., 1996. Small-area estimates using model-biased methods. Can. J. For. Res. 26:758–766.

Katila, M. & Tomppo, E., 2001. Selecting estimation parameters for the Finnish multi-source national forest inventory. Remote Sens. Environ. 76:16–32.

—, 2002. Stratification by ancillary data in multisource forest inventories employing k-nearest neighbour estimation. Can. J. For. Res. 32(9):1548–1561.

Keller, J. M., Gray, M. R. & Givens Jr., J. E., 1985. A fuzzy k-nearest neighbor algorithm. IEEE Transactions on Systems, Man, and Cybernetics 15(4):580–585.

Lappi, J., 2001. Forest inventory of small areas combining the calibration estimator and a spatial model. Can. J. For. Res. 31:1551–1560.

Linton, O. & Härdle, W., 1998. Nonparametric regression. In: Kotz, S., Read, C. B. & Banks, D. L. (eds.), Encyclopedia of Statistical Sciences, update vol. 2, Wiley, New York, pp. 470–485.

Malinen, J., 2003. Locally adaptable non-parametric methods for estimating stand characteristics for wood procurement. Silva Fennica 37(1):109–120.

McRoberts, R. 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 Sens. Environ. 82:457–468.

Nilsson, M., 1997. Estimation of Forest Variables Using Satellite Image Data and Airborne Lidar. Ph.D. thesis, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics. Acta Universitatis Agriculturae Sueciae. Silvestria 17.

Polasek, W., Liu, S. & Neudecker, H., 1998. Heteroskedastic linear regression models. In: Kotz, S., Read, C. B. & Banks, D. L. (eds.), Encyclopedia of Statistical Sciences, update vol. 3, Wiley, New York, pp. 301–309.

Schreuder, H. T., Gregoire, T. G. & Wood, G. B., 1993. Sampling methods for multiresource forest inventory. Wiley, New York, U.S.A.

Steele, B. M., Winne, J. C. & Redmond, R. L., 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps. Remote Sens. Environ. 66(2):192–202.

Tokola, T. & Kilpeläinen, P., 1999. The forest stand margin area in the interpretation of growing stock using Landsat TM imagery. Can. J. For. Res. 29:303–309.

Tokola, T., Pitkänen, J., Partinen, S. & Muinonen, E., 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. Int. J. Remote Sens. 17(12):2333–2351.

Tomppo, E., 1987. An application of a segmentation method to the forest stand delineation and estimation of stand variates from satellite images. In: Proceedings of the 5th Scandinavian Conference on Image Analysis. Stockholm, pp. 253–260.

—, 1991. Satellite image-based national forest inventory of Finland. International Archives of Photogrammetry and Remote Sensing 28:419–424. Proceedings of the Symposium on Global and Environmental Monitoring, Techniques and Impacts, 17–21 Sept. 1990, Victoria, British Columbia, Canada.

—, 1996. Multi-source National Forest Inventory of Finland. In: Päivinen, R., Vanclay, J. & Miina, S. (eds.), New Thrusts in Forest Inventory. Proceedings of the subject group S4.02-00 'Forest Resource Inventory and Monitoring' and subject group S4.12-00 'Remote Sensing Technology', vol. 1, IUFRO XX World Congress, 6–12 Aug. 1995, Tampere, Finland. European Forest Institute, Joensuu, Finland, pp. 27–41.

Tomppo, E., Henttonen, H., , Ihalainen, A., Tonteri, T. & Tuomainen, T., 2001. Etelä-Savon metsäkeskuksen alueen metsävarat 1966–2000. Folia For. 2B/2001:309–388. (in Finnish).

Tomppo, E., Henttonen, H., Korhonen, K. T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., Mikkelä, H., Tonteri, T. & Tuomainen, T., 1998. Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968–97. Folia For. 2B/1998:293–374. (in Finnish).

Tomppo, E., Varjo, J., Korhonen, K., Ahola, A., Ihalainen, A., Heikkinen, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S. & Tuomainen, T., 1997. Country report for Finland. In: Study on European Forestry Information and Communication Systems. Reports on forestry inventory and survey systems, Office for Official Publications of the European Communities, Luxembourg, volume 1, pp. 145–226.

Zhang, J. & Foody, G. M., 1998. A fuzzy classification of sub-urban land cover from remotely sensed imagery. Int. J. Remote Sens. 19(14):2721–2738.

III

Calibration of small-area estimates for map errors in multisource forest inventory

Matti Katila, Juha Heikkinen, and Erkki Tomppo

Abstract: A multisource inventory method has been applied in the Finnish National Forest Inventory (NFI) since 1990. The method utilizes satellite images and digital map data, in addition to field measurements, and produces estimates of all field parameters for computation units as well as thematic maps. Information from base maps is employed in delineating forestry land from other land use classes. The map data are not necessarily up-to-date and often contain significant errors. This paper introduces a statistical calibration method aimed at reducing the effect of map errors on multisource forest resource estimates. The correction is based on the confusion matrix between land use classes of the field sample plots and corresponding map information. The proposed method is illustrated in a realistic setting using data from the ninth NFI.

Résumé : Une méthode d'inventaire multi-source a été appliquée dans le cadre de l'inventaire national des forêts en Finlande depuis 1990. La méthode utilise des images satellite et des données cartographiques digitales, en plus de mesures prises sur le terrain, et produit des estimés de tous les paramètres de terrain pour les unités de calcul aussi bien que pour les cartes thématiques. L'information des cartes de base est utilisée pour délimiter le territoire forestier et le distinguer du territoire affecté à d'autres usages. Les données cartographiques ne sont pas nécessairement à jour et contiennent souvent des erreurs significatives. Cet article présente une méthode statistique de calibration visant à réduire l'effet des erreurs cartographiques sur les estimés multi-source des ressources forestières. La correction est basée sur la matrice de confusion entre les classes d'utilisation des terres des places-échantillons sur le terrain et les informations géographiques correspondantes. La méthode proposée est illustrée ici dans un contexte réaliste à l'aide de données provenant du neuvième inventaire national des forêts.

[Traduit par la Rédaction]

Introduction

One of the greatest challenges to today's large-scale forest inventories is to produce accurate localized results. Estimates are required for small regions, such as municipalities or forest holdings, using sample sizes that yield adequately precise estimators only for larger regions, such as provinces or forestry centres. This problem is also familiar, and more widely studied in the context of official and demographic statistics, where various strategies have been proposed for small-area estimation, usually utilizing supplementary data from censuses or administrative records (e.g., Rao 1998).

For forest inventories, digital maps and satellite images are the most commonly available useful sources of supplementary data. Typical topographic map information is helpful in separating the area of interest, the forestry land, from water and areas of other land use, though the maps are seldom up-to-date. Other common problems with map data include location errors, missing or noncorresponding land use classes, and errors that arise during data processing, when rasterizing map themes of small or narrow area, for example.

In the National Forest Inventory of Finland (NFI), conducted by the Finnish Forest Research Institute, digital maps and satellite images have been used in small-area estimation since 1990. The applied multisource method (MS-NFI; Tomppo 1991, 1996), using the k-nearest-neighbours (k-nn) estimation, has been proven to yield reliable small-area statistics and to be practical for operational use. It has gained widespread interest and has been experimented in Sweden, Germany, Norway, China, and New Zealand (Tomppo et al. 1999*a*).

MS-NFI is essentially a two-stage procedure, where digital maps are applied in the first stage to delineate forestry land and to estimate its area. Estimation of the area of forestry land subclasses and the mean and sum of forest variables are then based on field observations and satellite data within the map-delineated forestry land. The reason for this is that all nonforestry land use classes cannot be separated from forestry land reliably enough with satellite image analysis (Tomppo 1996).

The direct use of digital maps typically yields overestimates of forestry land area, mostly because some land use masks (e.g., power lines and railways) are not always available in the applied digital maps. On the other hand, nonforestry land included within the map-based forestry strata reduces the mean timber volume estimates. In practice it has sometimes been necessary to calibrate the small-area estimates in such a way that their aggregation into large regions agrees with the corresponding estimates from pure field measurements.

In this paper a statistical calibration method is suggested to reduce the effect of map errors on small-area estimates using the confusion matrix estimated from a large region. For the land use class areas, the suggested calibration leads to synthetic estimators (Gonzalez 1973), whose aggregates over the whole region agree with unbiased post-stratification estimators (Holt and Smith 1979). The approach is also found in calibration and remote sensing literature (Brown 1982; Czaplewski and Catts 1992) as "inverse calibration for classification error," a method introduced in Tenenbein (1972). However, our application and the proposed method are slightly different from the usual

Received January 21, 1999. Accepted October 6, 1999.

M. Katila,[1] J. Heikkinen, and E. Tomppo. Finnish Forest Research Institute, Unioninkatu 40 A, FIN-00170 Helsinki, Finland.
e-mails: matti.katila@metla.fi; juha.heikkinen@metla.fi; erkki.tomppo@metla.fi

[1] Corresponding author.

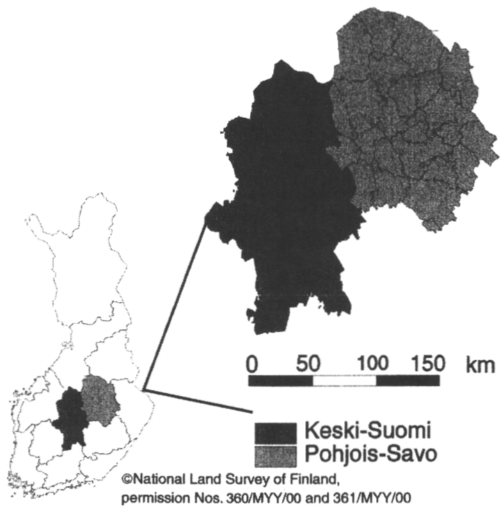


Fig. 1. Location of the study area, and the municipality boundaries of the forestry centres.

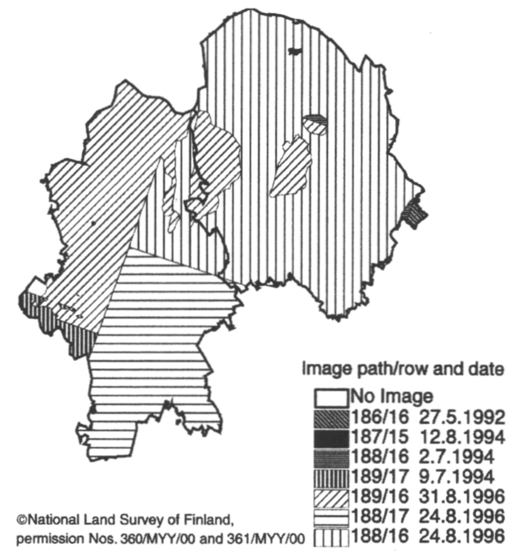


Fig. 2. Satellite image mosaic, Landsat 5 Thematic Mapper path and row and time of acquisition.

classification setting, because the map categories may differ from the statistics categories.

Objectives of the study

The aims of this study were to develop a calibration method that could be implemented into the current operational NFI and to study its behaviour in a realistic setting, using the computation units and the actual data available from the ninth NFI of Finland, with all its limitations. To fulfill these aims we could not always choose the most obvious alternatives. For example, different types of map errors had to be treated differently in order to integrate the calibration into the current MS-NFI. Also, we have not yet found a fully satisfactory method of assessing the standard errors of MS-NFI estimators, which makes it rather difficult to study the properties of calibrated small-area estimators.

Standard errors are available for large-area estimators from pure field data and post-stratified land use class area estimators (the available map data are not useful as a basis for poststratification in the estimation of forest variables). Accordingly, the main emphasis here is on the large-area properties: Does calibration improve the agreement of the aggregates of small-area estimates with unbiased field data estimates? The assumption of homogeneous map errors over the large regions may naturally lead to biases in the synthetic estimators for small regions. Their magnitude in the small regions of interest is difficult to determine, but to reveal significant biases we examined regions of intermediate area, for which pure field data estimates are reasonably reliable.

In this study, large regions are represented by forestry centres and small regions, by municipalities. The primary land use classes are forestry land (FRYL), arable land, built-up land, land claimed by traffic and power lines, and water. Forestry land is further divided into subclasses of forest land, other wooded land, and waste land. The union of forest and other wooded land (FOWL) is of particular importance, because only trees on FOWL are included in the NFI timber volume estimates; waste land consists of practically treeless open bogs and rocks. The statistics considered in this study are the area of the primary land use classes and of FOWL (essential in total volume estimation), and the mean and total volumes of growing stock of major tree species.

Material

Field measurements

The study area (Fig. 1) consists of the forestry centres of Keski-Suomi (total area 19 388 km^2) and Pohjois-Savo (total area 19 953 km^2) in central Finland, located approximately in the area bordered by 24°10′–28°50′E and 61°20′–64°00′N. Water covers 17% of the study area and forest 84% of the land area. The forests are typical boreal forests dominated by Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.), which also form mixed forests with birch (*Betula* spp.) and other deciduous species.

The study area contains 54 municipalities ranging from 68 to 1589 km^2; 30 municipalities are located in Keski-Suomi and 24 are in Pohjois-Savo. The total area and land area of each municipality were obtained from the National Land Survey of Finland (1997) and are assumed to be exact in this study.

The field data are from the ninth NFI; both Keski-Suomi and Pohjois-Savo were sampled during the 1996 field season. A systematic cluster sampling design was applied, where one cluster consists of 18 or 14 field plots[2] located along a rectangular tract

[2]Three out of four clusters consist of 18 temporary plots; in every fourth cluster, 14 permanent plots were established with few additional measurements.

Fig. 3. Topographic database (dark grey) and the Base map data areas.

300 m apart (Tomppo et al. 1999*c*, 1999*b*). The cluster reference points form a square lattice with 7 km between adjacent clusters. A total of 13 613 field plots, of which 11 275 were on land, were measured within our study area.

Trees were measured on parts of plots belonging to FOWL. They were picked up by sampling with probability proportional to size, the inclusion probability of a tree being proportional to its basal area. Relascopic factor two was applied on the study area. Diameter and distance of all boundary trees were measured to judge whether a tree should be included in the sample.

Supplementary data

Seven Landsat 5 Thematic Mapper images were needed to cover the whole study area (Fig. 2). Three main images were from the same year as the field inventory data and covered 90% of the area. Four additional images from years 1992 and 1994 were used to obtain full cloud-free coverage. Each image was rectified to the National Coordinate System using regression models of first- or second-order polynomials, fitted to 30–70 control points, which were identified from base maps. The nearest-neighbour method was applied for the resampling of the images to 25 m × 25 m pixel size.

The digital map data are mainly from the National Land Survey (FNLS) but vary in their quality and accuracy. The Topographic Database (TOPO) (National Land Survey of Finland 1998) is the most accurate and up-to-date data source, but it currently covers only 50% of our study area (Fig. 3). For the rest of the area (to be called 'BASE area'), the map data are from several data sources. For example, the arable land mask was scanned from the 1 : 50 000 basic maps, for which the field work dates from 1961 to 1985.

Table 1. Derived map stratification.

Code (h)	Stratum	Region
1	Forestry land	BASE
2	Forestry land	TOPO
3	Arable land	BASE
4	Arable land	TOPO
5	Buildings and urban area	BASE
6	Buildings and urban area	TOPO
7	Other built-up land	BASE
8	Other built-up land	TOPO
9	Roads	Whole
10	Water	BASE
11	Water	TOPO

By combining the various map data sources we produced a thematic map (or stratification) that classifies each 25 m × 25 m pixel of the study area into one of the 11 strata listed in Table 1. The stratification was designed in such a way that each stratum is as homogeneous as possible with respect to the land use class distribution. To achieve this, each theme, except for the uniform road mask, was split into BASE and TOPO areas. The forestry land strata consist of the areas that are left outside all other map themes.

As in the operational MS-NFI, we also used the digital elevation model and a peatland mask to supplement the satellite image data in the *k*-nn estimation, and the digital municipality boundaries (from FNLS), to delineate the computation units (Tomppo 1996).

Current NFI methods

The method presented in this paper has been developed as a modification of the currently operational NFI of Finland. Therefore the latter is used as the basis for comparisons whenever possible. The large-area NFI estimates for forestry centres are based on field data only, whereas municipality level (small-area) estimates are computed by means of the MS-NFI method using satellite images and digital map data.

Field data method

Area estimation from pure field data is based on ratios of field plot counts and on the known land area of forestry centres. The area of land use class l within forestry centre R is estimated by

$$[1] \qquad \widehat{A}_{R,l} = \frac{n_{R,l}}{n_{R,\text{land}}} A_{R,\text{land}}$$

where $n_{R,l}$ is the number of sample plots within R that represent land use class l, $n_{R,\text{land}}$ is the number of plots located on land, and $A_{R,\text{land}}$ is the total land area of R. The area of any subclass f whose indicator is observed in the field (e.g., forest – other wooded – waste land or pine – spruce – deciduous-dominated forest) can naturally be estimated in a similar manner.

The mean volume within forestry land subclass f of forestry centre R is estimated by the sample average

$$[2] \quad \widehat{v}_{R,f} = \frac{\sum_{i \in I_{R,f}} v_i}{n_{R,f}}$$

where v_i is the mean volume (m^3/ha) in field plot i, $I_{R,f}$ is the set of sample plots within R that represents subclass f, and $n_{R,f}$ is the (random) number of such plots. The total volume estimator $\widehat{V}_R$ is the product of the mean volume estimator ([2]) for f = FOWL and the FOWL area estimator obtained using [1], and it simplifies to

$$[3] \quad \widehat{V}_R = \frac{\sum_{i \in I_{R,\mathrm{FOWL}}} v_i}{n_{R,\mathrm{land}}} A_{R,\mathrm{land}}$$

These can all be considered as ratio-of-means estimators. However, because of the spatial correlation of forest variables combined with the systematic sampling design, the usual variance estimators (Cochran 1977), based on simple random sampling, are not valid in their accuracy assessment. Instead, the standard errors are estimated using local quadratic forms as suggested in Matérn (1960); details are given in the Appendix.

MS-NFI method

In the current multisource method, forestry land is separated from other land uses on the basis of the digital map data (strata 1 and 2 of Table 1; the term 'forestry land stratum' will refer to the union of these strata). Forestry land subclass areas and means, and totals of forest variables, are estimated by weighted sums or averages of field measurements in plot i belonging to the training data set $J \subset I_{\mathrm{FRYL}}$, which includes all FRYL plots except for those that are obviously poorly localized in the field and those that contain non-FRYL parts. In our study, this set contained 9417 field plots.

Weights for plot $i \in J$ are computed as sums of pixel weights over pixels in the forestry land stratum. The pixel weights, in turn, are determined by the k-nn method, the details of which are given in Tomppo (1991, 1996). The basic idea is to use the satellite image and other supplementary data to find, for each pixel p within the forestry land stratum, the k most similar in the training set. Let us denote the field plots corresponding to these "k nearest-neighbours" of p by $i_1(p), \ldots, i_k(p)$. Non-negative weights $w_{i,p}$ are defined according to the applied similarity measure in such way that $w_{i,p} > 0$, if and only if $i \in \{i_1(p), \ldots, i_k(p)\}$ and

$$[4] \quad \sum_{i \in J} w_{i,p} = a$$

where a is the area of one pixel. The weight of plot $i \in J$ to the forestry land of municipality U is then

$$[5] \quad c_{i,U} = \sum_{p \in U_{\mathrm{FRYL}}} w_{i,p}$$

where U_{FRYL} denotes the restriction of the forestry land stratum to U. Note that $c_{i,U}$ may be positive also for plots outside U, which leads to synthetic estimators borrowing strength from outside the computation unit.

The sum of weights $c_{i,U}$ over all training set plots is equal to the area of U_{FRYL}. This allows for the interpretation of $c_{i,U}$ as that area of the forestry land of U that is most similar to plot i. The natural estimator for the area of any forestry land subclass f within U is then

$$[6] \quad \widetilde{A}_{U,f} = \sum_{i \in J_f} c_{i,U}$$

where J_f contains the training set plots that belong to subclass f. The MS-NFI estimator of the mean volume within forestry land subclass f of U is the weighted average:

$$[7] \quad \widetilde{v}_{U,f} = \frac{\sum_{i \in J_f} c_{i,U} v_i}{\sum_{i \in J_f} c_{i,U}}$$

and that of the total volume is obtained by choosing subclass f = FOWL and omitting the denominator:

$$[8] \quad \widetilde{V}_U = \sum_{i \in J_{\mathrm{FOWL}}} c_{i,U} v_i$$

Calibrated MS-NFI estimators

It is obvious that the MS-NFI method is vulnerable to the failure of the forestry land map stratum in representing the true forestry land area. In practice, this stratum is usually too large, leading to the overestimation of the forestry land area. On the other hand, the nonforestry land pixels typically add to the weights of low volume plots, which leads to the underestimation of mean volume.

Here we propose a calibration to the MS-NFI estimators, based on large-area estimates of map errors. Although the quality of map data varies, it is often possible to define the map strata in such a way that each one is reasonably homogeneous with respect to the map errors and the land use class distribution. This enables the use of synthetic small-area estimation, using the proportions that have been estimated from a larger region. The stratification of Table 1 was applied in this study. The restriction of stratum h to forestry centre R or to municipality U will be denoted by R_h and U_h, respectively.

Land use class areas

Let us first consider the estimation of the area of land use class l in municipality U. Recall that the (uncalibrated) MS-NFI estimator would simply be the combined area of the map strata within U that correspond to class l. We propose a natural calibration for errors in the map strata using the field data from forestry centre R in which U belongs. First, the proportion of land use class l within each map stratum h is estimated by the corresponding plot count ratio:

$$[9] \quad \widehat{P}_{R_h,l} = \frac{n_{R_h,l}}{n_{R_h}}$$

computed over the entire forestry centre (ratios computed by municipalities are too variable, the very reason for the need of specific small-area estimation). The calibrated area estimator is then obtained by summing the corresponding proportions of municipality level stratum areas:

$$[10] \quad A^*_{U,l} = \sum_h \widehat{P}_{R_h,l} A_{U_h}$$

where A_{U_h} is the area of U_h.

Note that the aggregate of small-area estimates over forestry centre R

$$[11] \quad \sum_{U \in R} A^*_{U,l}$$

is equal to the unbiased poststratification estimator:

$$[12] \quad A^*_{R,l} = \sum_h \widehat{P}_{R_h,l} A_{R_h}$$

Properties of the synthetic municipality level estimators [10] depend highly on the homogeneity of map strata with respect to land use class distribution. If the true proportions $P_{U_h,l}$ were constant for all municipalities within R, then the estimators would be unbiased.

Since the areas A_{R_h} are known and proportions $\widehat{P}_{R_h,l}$ are based on field data, the sampling errors of the large-area estimators $A^*_{R,l}$ can be assessed by combining standard stratified sampling formulae (Cochran 1977) with the variance estimators for the field data method. Again, the details are given in the Appendix.

Calibrated plot weights

Map errors affect MS-NFI estimators [6]–[8] through the plot weights $c_{i,U}$, which are defined as sums over the forestry land map strata [5]. Calibration of these weights for the map errors is not straightforward in the MS-NFI context, essentially because nonforestry land field plots are excluded from the training set for satellite image processing and also because the map strata are different from the NFI land use classes. Here we propose a heuristically derived calibration, which is implementable in the currently operational system and has the important property that in analogy with the uncalibrated MS-NFI, the sum of the calibrated weights over all training data plots is equal to the calibrated forestry land area estimate $A^*_{U,\mathrm{FRYL}}$.

First, we wish to eliminate from the sum on the right hand side of [5] the contribution of the pixels that are falsely classified as forestry land on the basis of map data. Our proposal is to estimate the contributions separately for each nonforestry land use class l by the product of the estimates of the number and average weight of the forestry land stratum pixels that actually belong to l.

Using again the large-area estimate of the confusion matrix, the number of such pixels in municipality U can be estimated by

$$[13] \quad N^*_{U_{\mathrm{FRYL}},l} = \sum_{h \in \{1,2\}} \widehat{P}_{R_h,l} N_{U_h}$$

where N_{U_h} is the number of pixels in U_h.

We have no direct way to reliably assess the weights $(w_{i,p})$ of the forestry land stratum pixels (p) that actually belong to l.

Table 2. Representative map strata for nonforestry land use classes used in the plot weight correction.

Land use	Strata
Arable	3,4
Built-up	5,6
Traffic, etc.	9
Water	10,11

Therefore our estimation is based on the assumption that they are, on average, similar to those of pixels in such map strata that best represent land use class l. We selected the representative map strata for each nonforestry land use class as shown in Table 2; the union of the map strata that represent land use class l is denoted by $\boldsymbol{h}(l)$. The pixel weights $(w_{i,p})$ of training data set plots $i \in J$ were computed to all pixels (p) within these strata in the same manner (k-nn) as for those within the forestry land stratum in the ordinary MS-NFI. The average weight of training set plot i to a pixel, whose actual land use class is l, was then estimated by

$$[14] \quad \overline{w}_{i,U_{\boldsymbol{h}(l)}} = \sum_{p \in U_{\boldsymbol{h}(l)}} \frac{w_{i,p}}{N_{U_{\boldsymbol{h}(l)}}}$$

Our estimator for the total contribution of falsely classified forestry stratum pixels to $c_{i,U}$ is finally obtained by summing the products of [13] and [14] over all nonforestry land use classes:

$$[15] \quad c^-_{i,U} = \sum_{l \neq \mathrm{FRYL}} N^*_{U_{\mathrm{FRYL}},l} \overline{w}_{i,U_{\boldsymbol{h}(l)}}$$

This leads to "downwards calibrated" weights:

$$[16] \quad c'_{i,U} = c_{i,U} - c^-_{i,U}$$

which can be considered to represent the contribution from pixels in the forestry land strata that actually belong to forestry land. The calibrated estimator of the area of that part is

$$[17] \quad A^*_{U_{\mathrm{FRYL}},\mathrm{FRYL}} = \sum_{h \in \{1,2\}} \widehat{P}_{R_h,\mathrm{FRYL}} A_{U_h}$$

To account for the map errors to the other direction, that is, for pixels in the nonforestry land strata that actually belong to forestry land, we assumed that in each computation unit they are, on average, similar to pixels in the forestry land stratum of that unit. This leads to scaling the downwards calibrated weights $c'_{i,U}$ up by the area correction factor $A^*_{U,\mathrm{FRYL}}/A^*_{U_{\mathrm{FRYL}},\mathrm{FRYL}}$. As a result the calibrated weights are

$$[18] \quad c^*_{i,U} = \frac{A^*_{U,\mathrm{FRYL}}}{A^*_{U_{\mathrm{FRYL}},\mathrm{FRYL}}} \left(c_{i,U} - \sum_{l \neq \mathrm{FRYL}} N^*_{U_{\mathrm{FRYL}},l} \overline{w}_{i,U_{\boldsymbol{h}(l)}} \right)$$

Calibrated MS-NFI estimates are then obtained by replacing $c_{i,U}$ in [6]–[8] by $c^*_{i,U}$. It should be noted that although these weights add up to $A^*_{U,\mathrm{FRYL}}$, the positivity of individual weights is not guaranteed.

Table 3. Land use class distribution among field plots by map stratum in the forestry centre of Keski-Suomi.

	Land use class l										
	Forestry		Arable		Built-up		Traffic, etc.		Water		Total
Stratum h	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	n_{R_h}
Forestry/BASE	94.6	2763	1.9	55	1.3	37	1.3	39	0.9	25	2919
Forestry/TOPO	93.8	1766	2.0	37	2.4	46	1.3	24	0.5	9	1882
Arable/BASE	23.0	49	69.9	149	5.2	11	0.9	2	0.9	2	213
Arable/TOPO	6.1	13	91.9	195	1.4	3	0.0	0	0.5	1	212
Buildings,urban/BASE	22.6	7	0.0	0	67.7	21	9.7	3	0.0	0	31
Buildings,urban/TOPO	10.5	2	0.0	0	89.5	17	0.0	0	0.0	0	19
Other built-up/BASE	10.5	2	0.0	0	89.5	17	0.0	0	0.0	0	19
Other built-up/TOPO	0.0	0	0.0	0	91.7	11	8.3	1	0.0	0	12
Roads	39.3	100	7.9	20	20.1	51	32.3	82	0.4	1	254
Water/BASE	2.3	10	0.0	0	0.4	2	0.0	0	97.3	431	443
Water/TOPO	0.3	2	0.2	1	0.4	3	0.0	0	99.1	677	683
Total	70.5	4714	6.8	457	3.3	219	2.3	151	17.1	1146	6687

Table 4. Land use class distribution among field plots by map stratum in the forestry centre of Pohjois-Savo.

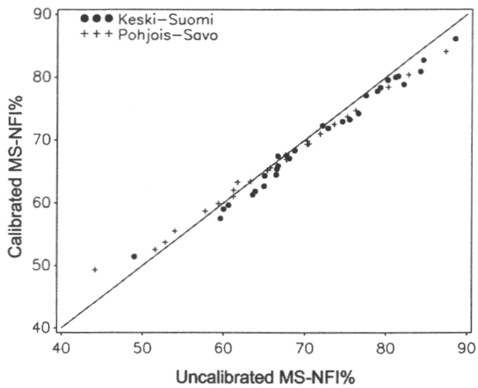
	Land use class l										
	Forestry		Arable		Built-up		Traffic, etc.		Water		Total
Stratum h	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	$\widehat{P}_{R_h,l}$ (%)	$n_{R_h,l}$	n_{R_h}
Forestry/BASE	93.5	1853	2.0	39	1.7	34	1.2	24	1.6	31	1981
Forestry/TOPO	94.7	2631	1.8	49	1.9	53	1.2	33	0.4	12	2778
Arable/BASE	25.4	63	69.8	173	3.2	8	0.8	2	0.8	2	248
Arable/TOPO	4.8	20	93.1	388	1.2	5	0.7	3	0.2	1	417
Buildings,urban/BASE	9.4	2	4.8	1	81.0	17	0.0	0	4.8	1	21
Buildings,urban/TOPO	12.5	2	31.2	5	56.3	9	0.0	0	0.0	0	16
Other built-up/BASE	19.2	5	3.9	1	76.9	20	0.0	0	0.0	0	26
Other built-up/TOPO	10.3	3	3.5	1	86.2	25	0.0	0	0.0	0	29
Roads	41.6	89	15.9	34	12.6	27	29.9	64	0.0	0	214
Water/BASE	3.8	20	0.0	0	1.1	6	0.0	0	95.1	503	529
Water/TOPO	3.8	25	0.0	0	0.0	0	0.0	0	96.2	642	667
Total	68.1	4713	10.0	691	2.9	204	1.8	126	17.2	1192	6926

Table 5. Land use class area estimates and their standard errors for forestry centres with and without poststratification.

		Area (1000 ha)		Absolute SE (1000 ha)		Relative SE (%)	
Centre	Land use	Field	Poststratification	Field	Poststratification	Field	Poststratification
Keski-Suomi	Forestry	1382	1378	12.6	7.4	0.9	0.5
	Arable	134	128	9.9	4.7	7.4	3.6
	Built-up	64	69	5.7	4.0	8.8	5.9
	Traffic, etc.	44	44	3.6	3.2	8.1	7.4
	Water	314	321	—	2.3	—	0.7
Pohjois-Savo	Forestry	1357	1360	12.4	7.3	0.9	0.5
	Arable	199	187	10.9	4.9	5.5	2.6
	Built-up	59	59	6.7	4.4	11.4	7.4
	Traffic, etc.	36	39	3.6	3.4	9.8	8.8
	Water	344	351	—	2.9	—	0.8

Note: Water area of the field inventory column was obtained from the official statistics of the National Land Survey of Finland.

Fig. 4. Uncalibrated vs. calibrated MS-NFI estimates for each municipality: percentage of FRYL of the total area (%).



Results

Confusion matrices

The confusion matrices (Tables 3 and 4) are quite similar in the two forestry centres, thus indicating homogeneity in the map quality. Forest map strata have nearly 95% co-occurrence with field data on both BASE and TOPO map areas. On the basis of field plot counts, the forestry land strata overestimate the actual FRYL area by 2% in Keski-Suomi and by 1% in Pohjois-Savo. The most notable differences between BASE and TOPO maps are in the arable land strata, the latter clearly being more accurate. The road stratum clearly overestimates the corresponding land use class and is quite inaccurate, as expected.

Calibrated area estimates for municipalities

Figure 4 shows the calibrated municipality level estimates for the proportional area of FRYL plotted against the corresponding MS-NFI estimates. The calibration reduces the FRYL area in most cases, although the changes are small compared with the absolute values (Fig. 5*a*). The stronger correction in the smallest municipalities of Pohjois-Savo could be due to the overlay operations of the original land use masks: roads are on top of all other masks. Since the estimated proportion of FRYL in the road stratum is relatively high, this may lead to the effect of "transferring" too much FRYL from the rural to built-up areas.

Calibrated volume estimates for municipalities

The mean volume in FOWL increases overall after calibration, as expected (Fig. 5*b*). The decrease of the small-area FRYL estimates compensate the mean volume increase, and on average, the total volume estimates remain unchanged (Fig. 5*c*). However, the relative correction of the mean volume increases together with the uncalibrated MS-NFI estimate $\widetilde{v}_U$ (Fig. 5*b*).

To understand the effect of calibration on the volume estimators, we must consider some results of the classification of non-FRYL map strata. On average, other wooded land and waste land field plots receive higher weights in the analysis of non-FRYL strata than in the analysis of FRYL. The mean volume estimates for non-FRYL strata were lower (7–91 m^3/ha) than those for FRYL, except those for water (160–204 m^3/ha). The high values for water are due to both water and highly stocked coniferous stands having low intensities of reflectance on all of the Landsat Thematic Mapper satellite channels applied in this study.

Further analysis showed that the increase in the proportion of area covered by the TOPO map increased the relative correction of mean volume on FOWL. The lower proportion of water area on TOPO forestry stratum leads to smaller subtraction of the weights of plots with high mean volumes.

Table 6. Area of FOWL and forestry land of forestry centres: pure field data estimate, MS-NFI estimates with and without calibration of plot weights.

		Area (1000 ha)		
Centre	Land use	Field	MS-NFI	Calibr.
Keski-Suomi	FOWL	1368	1389	1368
	FRYL	1382	1403	1378
Pohjois-Savo	FOWL	1333	1349	1344
	FRYL	1357	1372	1360

Aggregates of calibrated area estimates

Recall that the aggregates of the calibrated estimates of land use class areas are equal to the poststratified large-area estimates, which are unbiased and more precise than the pure field data estimates. Table 5 shows that the poststratification results in nearly half the standard error of forestry land area estimators. The variance reduction is smaller on more heterogeneous classes, e.g., roads, where the within-strata variation is large.

Table 6 demonstrates how the calibration draws the aggregates of FOWL and FRYL area estimates towards the pure field data estimates (large-area estimates in Tables 6 and 7 were obtained by replacing U in [6]–[8] and their calibrated versions, by R). Note that based on poststratification standard errors, there is a significant bias in the uncalibrated MS-NFI estimate of FRYL area for Keski-Suomi.

Aggregates of calibrated volume estimates

Calibration of MS-NFI plot weights gives the expected increase to the mean volume in Keski-Suomi (Table 7). Pohjois-Savo seems to benefit little from the calibration, but there was little need for the calibration in the first place. The effect of the calibration on the volume estimates varies by tree species. Almost all the calibrated MS-NFI estimates of total volume are within 1 SE of the field inventory estimate in Keski-Suomi, and within 2 SE in Pohjois-Savo. There are noticeable biases in the MS-NFI estimates of the volume of birch and other deciduous species. The discrimination of these species is not easy, because they occur mainly as mixed species on coniferous stands.

Calibrated weights are negative for 1.5% of the training set plots. The negative weights result from the spectral responses of non-FRYL pixels in the satellite image being concentrated near those of a few exceptional FRYL pixels; for example, arable land pixels are similar to FRYL pixels with very low timber volume and water pixels are similar to high volume FRYL pixels. The mean volume of the negatively weighted plots over the study area was 142 m^3/ha. These field plots had slightly

Fig. 5. Percent difference between calibrated and uncalibrated MS-NFI estimates for each municipality plotted against the uncalibrated estimates: (*a*) area of FRYL (km^2), (*b*) mean volume (m^3/ha), and (*c*) total volume (1000 m^3).

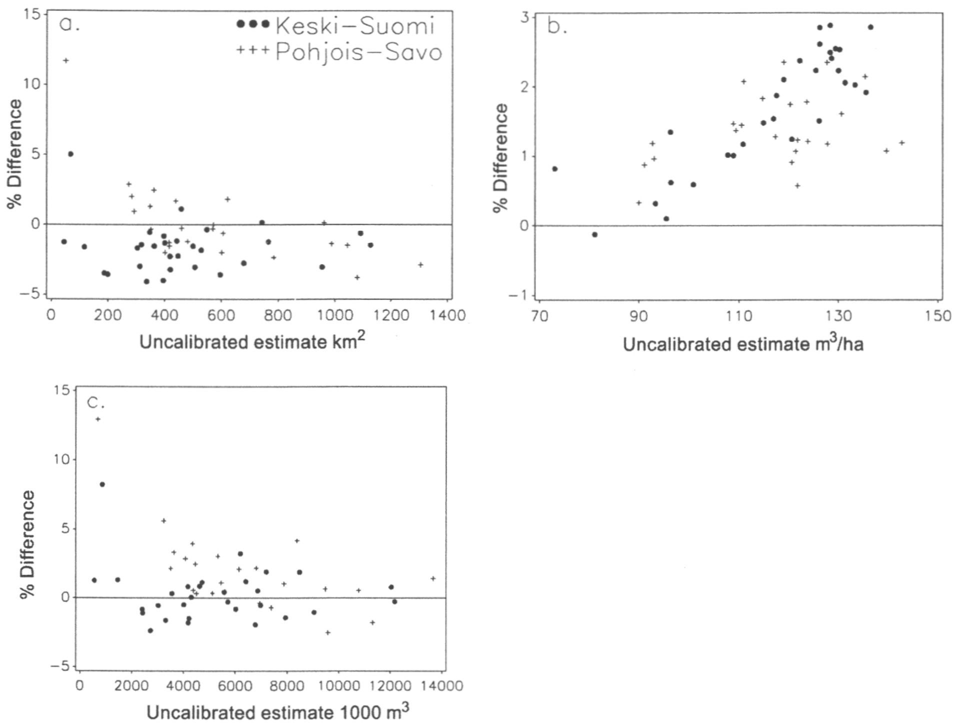


Table 7. Volume of growing stock on FOWL of forest centres: pure field data estimate with sampling error, MS-NFI estimates with and without calibration of plot weights.

		Mean volume (m^3/ha)				Total volume (10^6 m^3)			
Centre	Tree species	Field	MS-NFI	Calibr.	SE	Field	MS-NFI	Calibr.	SE
Keski-Suomi	Pine	48.6	47.2	48.0	1.1	66.4	65.5	65.7	1.6
	Spruce	47.7	47.4	48.4	1.5	65.2	65.8	66.3	2.1
	Birch	15.8	15.0	15.0	0.5	21.6	20.8	20.5	0.7
	Other deciduous spp.	4.2	4.1	4.1	0.3	5.7	5.7	5.6	0.4
	Total growing stock	116.2	113.7	115.5	1.8	158.9	157.9	158.1	2.9
Pohjois-Savo	Pine	36.7	37.5	37.9	1.0	48.9	50.5	51.0	1.4
	Spruce	52.1	52.5	53.6	1.5	69.5	70.9	72.1	2.1
	Birch	18.5	17.6	17.8	0.6	24.6	23.7	23.9	0.8
	Other deciduous spp.	6.6	5.5	5.5	0.4	8.8	7.5	7.4	0.5
	Total growing stock	113.9	113.1	114.8	1.7	151.8	152.6	154.3	2.7

higher pine volume estimates than the mean pine volume over the whole study area. Of the negative weights, 29% were on other wooded land and waste land field plots.

Small-area bias

To study a possible bias of small-area estimates the original and calibrated MS-NFI estimates of FRYL area, mean, and total volume were combined into five groups of municipalities in both forestry centres and compared with the pure field data estimates (Fig. 6). The FRYL areas of the subregions ranged from 1890 to 4160 km^2. The standard errors of the field inventory estimates were calculated and plotted (Figs. 7 and 8). The calibration of MS-NFI estimates did not cause notable systematic errors to the FRYL area and volume estimates compared with

Fig. 6. Groups of municipalities 1–10 in the study area.

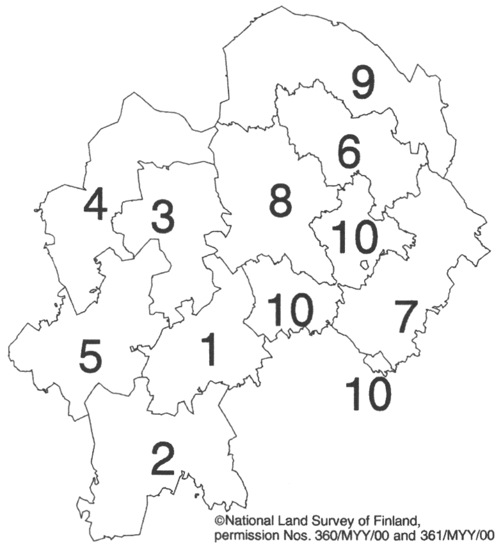
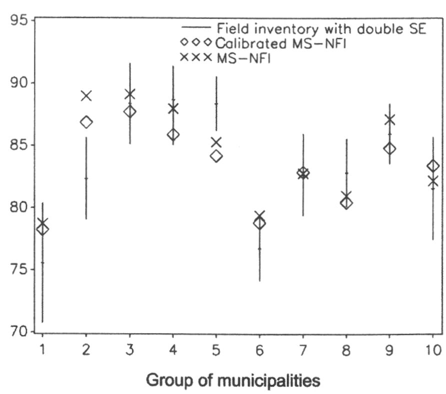


Fig. 7. Groups of municipalities: pure field data estimates ±2 SE, MS-NFI estimates and calibrated MS-NFI estimates; percentage of FRYL (%).

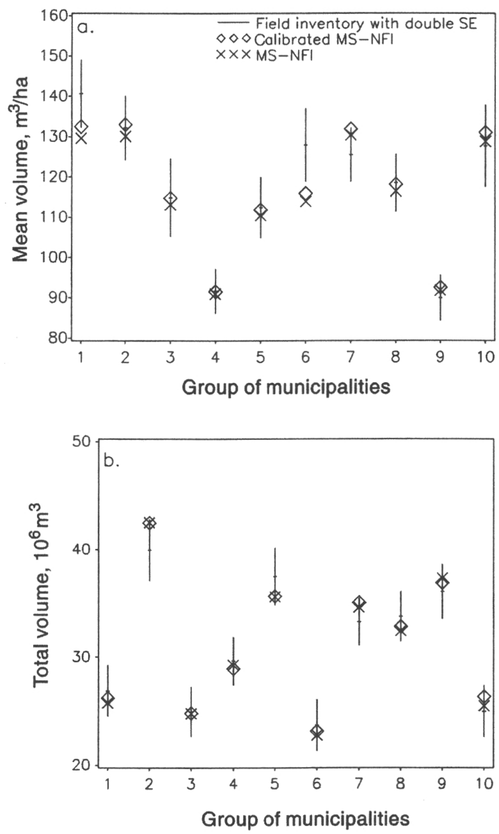


the pure field data estimates, although in two groups neither of the two estimates of FRYL area were within 2 SE of the pure field data estimates (Fig. 7). The corrections are, in most groups of municipalities, towards the field inventory estimates.

Discussion

We have presented a statistical calibration method for reducing the effect of the map errors in the MS-NFI estimates. The method uses a confusion matrix, estimated from two data sets. One of them is assumed to be a sparse sample yielding accurate unbiased estimates for large areas, but having too few observations for reliable small-area estimation. The other data set, on the other hand, is assumed to give complete coverage of the study area, but may contain systematic errors.

Fig. 8. Groups of municipalities: pure field data estimates ±2 SE, MS-NFI estimates and calibrated MS-NFI estimates; (*a*) mean volume (m^3/ha), (*b*) total volume (10^6 m^3).



The method is derived and described in a real multisource forest inventory setting, applying satellite images, digital map data, and a sparse grid of field data from the ninth NFI of Finland. A normal large-area forest inventory involves a large number of variables, typically 200 to 400. The presented method is applicable with all variables and parameters.

In general, the calibration method corrects the aggregates of MS-NFI estimates towards those based on field data (which are considered to be unbiased).

In our application, the total numbers of field plots for forestry centres were fairly large, 6687 and 6926. In spite of that, field plots were seldom observed on the smallest map strata, such as urban areas and other built-up land. Therefore the standard errors of the estimates of $\widehat{P}_{R_h,l}$ for these strata can be high. Czaplewski and Catts (1992) recommend a minimum of 500–1000 random sample plots for categorical data assuming that the probabilities of misclassification are constant over the region. Well-classified categories would need smaller samples. In our case, the map stratum of buildings and urban areas may have too few observations compared with the accuracies of maps.

The method assumes that pixels that are spectrally and actually similar to non-FRYL pixels can be found among FRYL strata pixels. We do not know how much the non-FRYL spectral values really differ from FRYL. These differences also vary seasonally because of phenology, and may cause different allocation of weights with images of different time points. For instance, the spectral responses of arable land are more distinct from those of forestry land in the early summer when there is no vegetation.

In our study area, the mean weights of field plots corresponding to non-FRYL concentrated to certain field plots in such a way that [18] gave negative weights to 1.5% of the field plots. Methods for avoiding negative weights will be studied.

In general, we have found the calibration to work reasonably well, and it has already been implemented as a part of the operative MS-NFI.

Acknowledgements

We are grateful to Dr. Charles T. Scott, Dr. Raymond L. Czaplewski, and the Associate Editor for their valuable comments. The study was partly funded by the Graduate School Program "Forests in GIS," University of Helsinki.

References

Brown, P.J. 1982. Multivariate calibration. J. R. Stat. Soc. Ser. B **44**: 287–321.

Cochran, W.G. 1977. Sampling techniques. 3rd ed. Wiley, New York.

Czaplewski, R.L., and Catts, G.P. 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sens. Environ. **39**: 29–43.

Gonzalez, M.E. 1973. Use and evaluation of synthetic estimators. *In* Proceedings of the Social Statistics Section. American Statistical Association, Washington. pp. 33–36.

Holt, D., and Smith, T.M.F. 1979. Post stratification. J. R. Stat. Soc. Ser. A **142**: 33–46.

Matérn, B. 1960. Spatial variation. Medd. Statens Skogsforskningsinst. **49**(5). Also appeared as No. 36 of Lecture Notes in Statistics. Springer-Verlag, New York, 1986.

National Land Survey of Finland. 1997. Suomen pinta-ala kunnittain 1.1.1997. National Land Survey of Finland, Helsinki, Finland. [In Finnish]

National Land Survey of Finland. 1998. Topographic database. WWW-pages, National Land Survey of Finland, ⟨http://www.nls.fi/kartta/maps/topodb.html⟩ (accessed January 5, 1999).

Rao, J.N.K. 1998. Small area estimation. *In* Encyclopedia of statistical sciences, update vol. 2. *Edited by* S. Kotz, C.B. Read, and D.L. Banks. Wiley, New York. pp. 621–628.

Tenenbein, A. 1972. A double sampling scheme for estimating from misclassified multinomial data with applications to sampling inspection. Technometrics, **14**: 187–202.

Tomppo, E. 1991. Satellite image-based national forest inventory of Finland. *In* Proceedings of the Symposium on Global and Environmental Monitoring, Techniques and Impacts, 17–21 Sept. 1990, Victoria, B.C. Int. Arch. Photogramm. Remote Sens. **28**: 419–424.

Tomppo, E. 1996. Multisource national forest inventory of Finland. *In* New thrusts in forest inventory. Proceedings of IUFRO XX World Congress, 6–12 Aug. 1995, Tampere, Finland. *Edited by* R. Päivinen, J. Vanclay, and S. Miina. International Union of Forest Research Organisations, Wien, Austria. pp. 27–41.

Tomppo, E., Goulding, C., and Katila, M. 1999*a*. Adapting Finnish multisource forest inventory techniques to the New Zealand preharvest inventory. Scand. J. For. Res. **14**(2): 182–192.

Tomppo, E., Henttonen, H., Korhonen, K.T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., and Tuomainen, T. 1999*b*. Keski-Suomen metsäkeskuksen alueen metsävarat ja niiden kehitys 1967–96. [In Finnish.] Metsätieteen aikakauskirja – Folia For. (Helsinki) **2B**: 309–388.

Tomppo, E., Henttonen, H., Korhonen, K.T., Aarnio, A., Ahola, A., Heikkinen, J., and Tuomainen, T. 1999*c*. Pohjois-Savon metsäkeskuksen alueen metsävarat ja niiden kehitys 1967–96. [In Finnish.] Metsätieteen aikakauskirja – Folia For. (Helsinki) **2B**: 389–462.

Appendix A: On the estimation of standard errors

Let us first consider the estimation of area proportions from pure field data. To be specific, let us choose some regions (R) and one land use class (l), and define indicator functions:

$$\text{[A1]} \quad x(r) = \begin{cases} 1 & \text{if point } r \in R \text{ belongs to land use class } l \\ 0 & \text{otherwise} \end{cases}$$

and

$$\text{[A2]} \quad y(r) = \begin{cases} 1 & \text{if point } r \in R \text{ is on land} \\ 0 & \text{if it is not} \end{cases}$$

Then the proportion

$$\text{[A3]} \quad P = \frac{A_{R,l}}{A_{R,\text{land}}}$$

of land use class l among the land area of region R can be rewritten as

$$\text{[A4]} \quad P = \frac{\int_R x(r)\,\mathrm{d}r}{\int_R y(r)\,\mathrm{d}r}$$

In large-area NFI, proportion P is estimated by the field plot ratio:

$$\text{[A5]} \quad \widehat{P} = \frac{\sum_{c \subset I} x_c}{\sum_{c \subset I} y_c}$$

where c refers to a cluster of field plots, $x_c = \sum_{i \in c} x(r_i)$ is the number of those plots in c whose centre belongs to land use class l (r_i is the location of the centre of plot i), and $y_c = \sum_{i \in c} y(r_i)$ is the number of those plots in c whose centre is located on land.

To estimate the variance of $\widehat{P}$ the cluster-wise residuals:

$$\text{[A6]} \quad z_c = x_c - P y_c$$

are assumed to form a partial realisation of a second-order stationary stochastic process on R. Letting n denote the number of clusters in I and

$$\text{[A7]} \quad \widehat{z} = \sum_{c \subset I} \frac{z_c}{n}$$

the variance per cluster

[A8] $\sigma_z^2 = n\,\mathrm{E}\,\hat{z}^2$

is estimated by the average of quadratic forms:

[A9] $$T_g = \frac{(z_{c1} - z_{c2} - z_{c3} + z_{c4})^2}{4}$$

in rectangular groups g of four clusters:

[A10]
$$\begin{array}{cc} c3 & c4 \\ c1 & c2 \end{array}$$

The residuals [A6] are naturally evaluated using the estimate $\widehat{P}$ instead of the unknown true value P. The variance estimator for

[A11] $$\widehat{P} = \frac{n\widehat{z}}{\sum_c y_c} + P$$

is obtained by applying the standard approximation for ratio estimators (Cochran 1977):

[A12] $$\sigma_{\widehat{P}}^2 = \mathrm{E}(\widehat{P} - P)^2$$

$$= \mathrm{E}\left(\frac{n\widehat{z}}{\sum_c y_c}\right)^2 \approx \frac{n^2\,\mathrm{E}\,\widehat{z}^2}{(\sum_c y_c)^2} = \frac{n\sigma_z^2}{(\sum_c y_c)^2}$$

Inserting the quadratic form estimator of σ_z^2 yields

[A13] $$\widehat{\sigma}_{\widehat{P}}^2 = \frac{n\widehat{\sigma}_z^2}{(\sum_c y_c)^2} = \frac{q\sum_g T_g}{(\sum_c y_c)^2}$$

where the grouping factor q is the ratio between the number of clusters and the number of groups. Usually all possible quadruples are included so that each cluster appears in four distinct groups and $q = 1$. Estimators given by [1] and [2] are essentially field plot ratios similar to that in [A5], and the approach described here was used to estimate their standard errors.

Let us then consider the poststratification estimator [12] of the area of land use class l within region R. Using vector notation $\boldsymbol{A} = (A_{R_1}, \ldots, A_{R_B})^{\mathrm{T}}$, $\widehat{\boldsymbol{P}} = (\widehat{P}_{R_1,l}, \ldots, \widehat{P}_{R_B,l})^{\mathrm{T}}$, where B is the number of map strata and superscript T denotes the transpose, we can rewrite [12] as

[A14] $$A_{R,l}^* = \boldsymbol{A}^{\mathrm{T}}\widehat{\boldsymbol{P}}$$

and the variance of $A_{R,l}^*$ is

[A15] $$\sigma_{A^*}^2 = \mathrm{Var}\,A_{R,l}^* = \mathrm{Var}\,\boldsymbol{A}^{\mathrm{T}}\widehat{\boldsymbol{P}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{A}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of $\widehat{\boldsymbol{P}}$.

To estimate $\boldsymbol{\Sigma}$ we have derived a direct generalisation of [A13] to the multivariate case. For a consistent notation each element of vector $\widehat{\boldsymbol{P}}$ $(\widehat{P}_{R_h,l})$, $h = 1, \ldots, B$ is expressed as

[A16] $$\widehat{P}_{R_h,l} = \frac{\sum_{c \subset I} x_{c,h}}{\sum_{c \subset I} y_{c,h}}$$

where $x_{c,h}$ and $y_{c,h}$ are the cluster-wise plot counts within stratum R_h, corresponding to those in [A5], and the cluster-wise residuals are defined as

[A17] $$z_{c,h} = x_{c,h} - \widehat{P}_{R_h,l}\,y_{c,h}, \quad h = 1, \ldots, B$$

The elements of covariance matrix $\boldsymbol{\Sigma}$ $\sigma_{h,h'}^2$, $h, h' = 1, \ldots, B$, that is, the covariances of $\widehat{P}_{R_h,l}$ and $\widehat{P}_{R_{h'},l}$, are then estimated by

[A18] $$\widehat{\sigma}_{h,h'}^2 = \frac{q\sum_g T_{g,h,h'}}{\sum_c y_{c,h}\sum_c y_{c,h'}}$$

where the "covariance per cluster" is estimated by generalising [A9] to

[A19] $$T_{g,h,h'} = (z_{c1,h} - z_{c2,h} - z_{c3,h} + z_{c4,h})$$

$$\times(z_{c1,h'} - z_{c2,h'} - z_{c3,h'} + z_{c4,h'})/4$$

The variance estimator $\widehat{\sigma}_{A^*}^2$ is then obtained by simply inserting the estimated covariances to [A15].

IV

Stratification by ancillary data in multisource forest inventories employing *k*-nearest-neighbour estimation

Matti Katila and Erkki Tomppo

Abstract: The Finnish multisource national forest inventory (MS-NFI) utilizes optical area satellite images and digital maps in addition to field plot data to produce georeferenced information, thematic maps, and small-area statistics. In the early version, forestry land (FRYL) was taken directly from the numerical map data. Such data may be outdated and can contain significant errors, for example, the FRYL area is typically overestimated and the mean volume is underestimated. A statistical calibration method has been introduced to reduce the map errors on multisource forest resource estimates. It is based on large-area estimates of map errors, a confusion matrix among land-use classes of the field sample plots, and corresponding map information. The method has some drawbacks: calculations are more complicated than in the original MS-NFI and some field plots may have negative expansion factors. The paper presents a new stratified MS-NFI method to reduce the effect of inaccurate map data on the forest-resource estimates. In this method, the *k*-nearest-neighbour (*k*-NN) estimation is applied by strata. All the field plots within each map stratum, independently of their land-use classification by field crew, are used to estimate the areas of land-use classes and forest variables of that stratum. The method was tested on two large areas containing three Landsat 5 TM scenes and field-inventory data from the ninth NFI. The stratified MS-NFI is essentially a different estimation method compared with the calibrated MS-NFI, which calibrates the MS-NFI estimates more or less systematically in one direction. The stratified MS-NFI was found to be statistically simpler and there were fewer significant errors in the estimates than in the calibrated MS-NFI.

Résumé : L'inventaire forestier national multisource de la Finlande (MS-NFI) utilise les images satellitaires optiques et les cartes numérisées, en plus des données provenant de parcelles terrestres, pour produire l'information à référence spatiale, les cartes thématiques et les estimations pour de petites surfaces. Dans la version antérieure, le territoire forestier était obtenu directement à partir des cartes numériques. Or ces cartes peuvent être obsolètes et contenir des erreurs importantes : par exemple, la superficie du territoire forestier est typiquement surestimée et le volume moyen est sous-estimé. Une méthode de calibration statistique a été développée pour réduire les erreurs d'estimation multisource des ressources forestières à partir des cartes. Cette méthode est basée sur l'estimation des erreurs des cartes sur de grandes superficies au moyen d'une matrice de confusion entre les classes d'affectation des terres obtenues à partir des parcelles terrestres et l'information correspondante provenant des cartes. Elle comporte certains inconvénients. Les calculs sont plus compliqués qu'avec le MS-NFI original et certaines parcelles terrestres peuvent avoir des facteurs d'expansion négatifs. Cet article présente une nouvelle méthode, le MS-NFI stratifié, pour réduire l'effet des données erronées dans les cartes sur l'estimation des ressources forestières. Avec cette méthode, l'estimation *k*-NN est appliquée à chaque strate. Toutes les parcelles terrestres dans chaque strate, indépendamment de leur classification pour l'affectation du sol par les équipes de terrain, sont utilisées pour estimer l'aire selon la classe d'affectation des terres et les données forestières de cette strate. La méthode a été testée sur deux grandes zones couvertes par trois images Landsat 5 TM et les données d'inventaire terrestre du neuvième inventaire national de la Finlande. Essentiellement, le MS-NFI stratifié est une méthode d'estimation différente du MS-NFI calibré qui calibre les estimations du MS-NFI plus ou mois systématiquement dans une direction. Le MS-NFI stratifié s'est révélé plus simple du point de vue statistique et les estimations comportent significativement moins d'erreurs qu'avec le MS-NFI calibré.

[Traduit par la Rédaction]

Introduction

The multisource forest inventories have been subject to increasing research in the countries with existing national forest inventories (NFI), e.g., the Scandinavian countries and the United States. The basic idea has been to combine objectively measured field-inventory data with available numerical map data and remote sensing data, most often from high-resolution optical satellites (Landsat TM, Spot HRV). Sampling-based methods (Poso 1972) and nonparametric estimation methods have been used for multisource forest inventories (Tomppo 1991; Tokola et al. 1996; Nilsson 1997; Franco-Lopez et al. 2000; Gjertsen et al. 2000)

The amount of available numerical map data is increasing. In forest inventories, the maps and remote sensing data have been used to delineate the forestry land (FRYL) (Loetsch and Haller

Received 13 July 2001. Accepted 19 February 2002. Published on the NRC Research Press Web site at http://cjfr.nrc.ca/ on 24 August 2002.

M. Katila[1] and E. Tomppo. Finnish Forest Research Institute, Unioninkatu 40 A, FIN-00170 Helsinki, Finland.

[1]Corresponding author (e-mail: matti.katila@metla.fi).

 DOI: 10.1139/X02-047

1973). The classification of land-use or land-cover classes has been one of the major applications of satellite image based remote sensing (Campbell 1996). In the remote sensing applications, there are three different ways to incorporate the auxiliary geographic information system (GIS) data into the classification: stratification, classifier modification, and post-classification sorting (Hutchinson 1982). Examples are stratification of the image areas prior to estimation (Tomppo 1996), application of the map data as a new feature in the conventional estimation methods (Tomppo et al. 1999; Poso et al. 1987), and probability- or knowledge-based models for the multisource data fusion (Benediktsson and Kanellopoulos 1999). In the Finnish multisource national forest inventory (MS-NFI), FRYL has been delineated directly from the numerical map data (Tomppo 1991).

The problem with the current MS-NFI map data is that it is not necessarily up-to-date, it includes locational errors, and it does not correspond exactly to the NFI land-use classes. Typically, FRYL area is overestimated, and consequently the mean volume is underestimated in the MS-NFI small-area estimates. Land-use masks give more accurate area estimates than an estimation from the optical high resolution satellite data alone. However, the accuracy of area estimates can be increased if map information and satellite image information are used together.

The Finnish MS-NFI utilizes optical satellite images and digital maps in addition to field plot data. A nonparametric k-nearest-neighbour method (k-NN) is used in the estimation (Tomppo 1991). One of the advantages of the k-NN method is that all the inventory variables can be estimated at the same time. Field data from surrounding calculation units (municipalities), in addition to the unit itself, are utilized when estimating results for that unit. This makes it possible to obtain estimates for smaller areas than would be possible with sparse field data only (Kilkki and Päivinen 1987; Tomppo 1991, 1996; Nilsson 1997). The method produces georeferenced information, thematic maps, and small-area statistics. In the original MS-NFI (oMS-NFI), only those field plots that are located entirely on FRYL, on the basis of field inventory, are used in the estimator. The estimates of FRYL area are derived from the digital map data (FRYL mask). Both the FRYL map area and FRYL field plots are usually divided into two strata on the basis of map data: mineral soil stratum and peatland stratum (Tomppo 1996).

Currently, a calibration method, denoted here by cMS-NFI, has been introduced to reduce the effect of map errors on multisource forest resource estimates (Katila et al. 2000). The method is based on the confusion matrix among land-use classes of the field sample plots and corresponding map information. The FRYL area estimates of the calibration method are consistent with post-stratified estimates for large regions (i.e., in areas of 500 000 ha or greater), while for small areas the estimator is synthetic (Rao 1998). Despite the rather simple idea of the calibration, it is quite laborious when applied to the MS-NFI: the calculation is more complicated than in the oMS-NFI and some field plots obtain negative weights (Katila et al. 2000). All weights are in turn used to calculate the small-area estimates, e.g., for municipality-level estimates.

The aim of the research

The paper presents a new multisource forest inventory method that employs accurate field plot measurements, satellite images, and inaccurate land-use map data. The method simultaneously produces land use class estimates and forest parameter estimates and reduces the effect of inaccurate map data. In this method, denoted by sMS-NFI, the k-NN estimation is applied by strata. The whole area to be analysed, including water and all land areas, is stratified on the basis of map data. Each field plot is assigned to its corresponding stratum. All the field plots within each map stratum, independently of the field measurement based land-use class, are used for estimating the areas of land-use classes and forest variables of the particular stratum. The final estimates are derived by combining the stratum-wise estimates. It is expected that the method will give more accurate forest-variable estimates for FRYL than the oMS-NFI and possibly also more accurate estimates than the calibration method. The number of FRYL plots within a certain non-FRYL map stratum may be small, and the errors of the forest-variable estimates will therefore be high within the stratum. However, the weight of these estimates (plots) on the final combined estimates is small.

The questions to be studied are (*i*) does the new method reduce the errors of FRYL area estimates and other forest resource estimates caused by errors in land-use map data, and (*ii*) what is the error of the forest-variable estimates compared with the estimates based on pure field data in large areas?

The MS-NFI estimates will be calculated in a realistic setting using data from the ninth NFI. Small-area estimates for municipalities (68–1577 km^2) are calculated using the three different MS-NFI methods. The pixel-level errors of FRYL and non-FRYL estimates of the new method are compared with the estimates based on the oMS-NFI method by applying a leave-one-out cross-validation method. The estimates for large- and medium-scale (group of municipalities) areas are compared with the field-inventory estimates to discover the magnitude of errors of different methods.

Materials

Field measurements

Two study areas, central Finland and western Finland, were employed. The central Finland study area was within the Landsat 5 TM images 188/16 and 188/17 (acquisition date: 24 August 1996), and the western Finland study area within the image 191/16 (acquisition date: 13 June 1997). The NFI field measurements employed were from the same year as the satellite images. The field plots, used in the k-NN estimation, are located approximately between 20°38′E, 28°50′E and 61°20′N, 64°00′N (Fig. 1). FRYL covers 82 and 73% of the land area in the central Finland and western Finland study areas, respectively. The central Finland study area is rich in mineral soil forests while the western Finland study area contains large peatland forest areas (Katila and Tomppo 2001). Both study areas consist of typical boreal forests dominated by Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.). Birch (*Betula* spp.) and other deciduous species occur often as mixed species.

A subset of municipalities that were covered by the satellite images and field plots were selected for evaluating the small-

Fig. 1. Location of the study areas and applied Landsat 5 TM path, row, and date of acquisition.

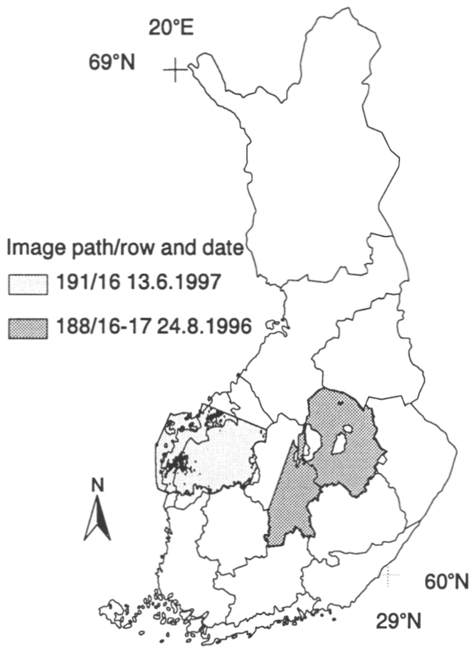
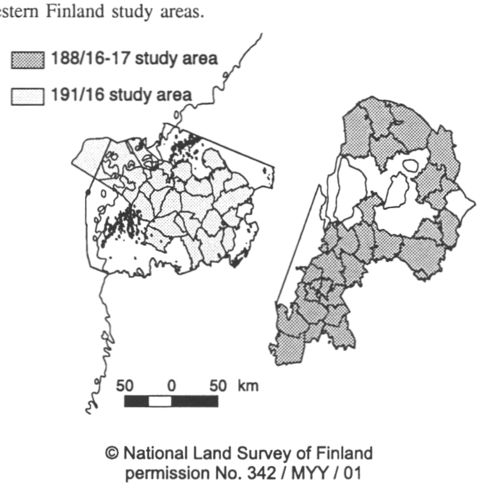


Fig. 2. Municipality boundaries in the central Finland and western Finland study areas.



area estimates. The two study areas contain 56 municipalities, 29 in central Finland and 27 in western Finland (Fig. 2). The municipalities range in area from 68 to 1577 km^2. The total areas and land areas of each municipality were obtained from the (National Land Survey of Finland 1997) and are assumed to be exact in this study.

The NFI field data were measured during the 1996 field season in central Finland and in 1997 in western Finland. A systematic cluster sampling design was applied. One cluster consists of 18 (temporary) or 14 (permanent) field plots located along a rectangular tract 300 m apart (Tomppo et al. 1998). The cluster reference points form a square lattice with a distance of 7 km between adjacent clusters. A total of 6816 and 7695 field plots were measured within the central Finland and western Finland training data areas, of which 4706 and 4832 were on FRYL, respectively.

Table 1. Derived map stratification.

Code	Stratum, *(h)*
1	Forestry land, mineral soil
2	Forestry land, peatland
3	Arable land
4	Built-up land and roads
5	Water

Trees were measured on parts of plots belonging to forest and other wooded land (FOWL). If a plot is cut by a stand or a land use class boundary, the entire plot is considered to consist of two or more parts. Trees were selected by probability proportional to size sampling, the inclusion probability of a tree being proportional to its basal area. A relascopic factor of two with a maximum distance of 12.52 m was employed. The diameter and distance of all boundary trees were measured to judge whether a tree should be included in the sample or not.

Supplementary data

The satellite images were rectified to the national coordinate system using regression models of first- or second-order polynomials, fitted to 30–70 control points, which were identified from base maps. The nearest-neighbour method was applied for the resampling of the images to 25 × 25 m pixel size (Tomppo 1996).

The digital map data comes mainly from the National Land Survey, but it varies in quality and accuracy. For the central Finland study area, the topographic database (National Land Survey of Finland 1996) was the most accurate and up-to-date data source, but it covered only 50% of the study area. For the rest of the area, the map data comes from several data sources (Katila and Tomppo 2001). In the western Finland study area, the areas covered by water and agricultural land were updated from the latest topographic database.

A stratification of the study area and field plots was produced by combining the various map data sources. Each 25 × 25 m pixel of the study area was delineated into one of the five strata listed in Table 1. The stratification was designed to form homogeneous strata with respect to the NFI-based land-use classes. On the other hand, the stratification was done in such a way that a high enough number of field plots, from the point of view of k-NN estimation, were included in each stratum. A more detailed stratification was used for the calibration method (Katila and Tomppo 2001).

A map of mineral soils and peatlands was used for stratifying the FRYL and the corresponding field plots already in oMS-NFI. This map was also used for the stratification in the new method. The digital elevation model was applied as in the oMS-NFI. Digital municipality boundaries were used to delineate the computation units (Tomppo 1996).

Methods

MS-NFI by strata

As explained below in more detail, new plot expansion factors, i.e., plot weights, are computed in the original multisource method for all the field plots i belonging to FRYL (Tomppo 1991, 1996). In the modified method, the weights are computed by land-use map strata. Some notations are introduced. Let us denote by H the set of the map strata, i.e., FRYL (subdivided into mineral soil stratum and peatland stratum), arable land, built-up land, and water, and by h its element. In the notations, no difference is made between a stratum, its ground elements, i.e., 25 × 25 m squares, and the image elements corresponding to the ground elements, i.e., the pixels. The set of pixels of municipality U is thus denoted by $U = \bigcup_{h\in H} U_h$, $U_h \bigcap U_{h'} = \emptyset$, if $h \neq h'$, where U_h is the set of pixels in stratum h. The set of field plot parts employed in the estimation is denoted by J and can be presented as a union of field plot parts of different land-use classes within different map strata: $J = \bigcup_{h\in H} \bigcup_{l\in G} J_{l,h}$, where l refers to land-use class on the basis of NFI field data (true land-use class), h is on the basis of a map stratum, and G is the set of land-use classes on the basis of NFI (Fig. 3). Instead of land-use class, the stratification l can be based on some subclass of FRYL, e.g., pine-dominated forests, and the rest of J. The area estimation with pure field data utilizes the information from the centre points of the field plots only, while the volume estimation uses information from the whole plot (Tomppo et al. 1997). The oMS-NFI utilizes all parts of the plots, also in area estimations (Tomppo 1996). However, the oMS-NFI usually uses only those field plots totally belonging to FRYL. A difference in the treatment of the plots by oMS-NFI and NFI field inventory exists only when the plot is shared among two or more different land-use class or forest stands (when the plot consists of two or more plot parts).

Plots on any land-use class and all parts of field plots are used in the estimation in the new method. Poorly localized field plots are, however, removed in both multisource methods because the correct image data can not be assigned to those plots.

As in the oMS-NFI, a distance measure d is defined in the feature space of the satellite image data. The k-nearest field-plot pixels (in terms of d), i.e., pixels which cover the centre of some field plot, are sought for each pixel p under the cloud-free satellite image area. Contrarily to the oMS-NFI, the neighbours are sought for each pixel within each U_h not only for FRYL map stratum pixels, and the neighbours can belong to any land-use class l or FRYL subclass f. Note that the neighbours must belong to the same map stratum as the target pixel.

The k-nearest field plots to pixel p_h, belonging to map stratum h, are denoted by $i_1(p_h), ..., i_k(p_h)$.

The weight w_{i,p_h} of field plot i to pixel p_h is defined as

[1]
$$w_{i,p_h} = \frac{1/d^t_{p_{h_i},p_h}}{\sum_{j\in\{i_1(p_h),...,i_k(p_h)\}} 1/d^t_{p_{h_j},p_h}} \quad \text{if } i \in \{i_1(p_h), ..., i_k(p_h)\}$$
$$= 0, \quad \text{otherwise}$$

where t is the power applied with the distance measure d.

The weight w_{i,p_h} of the plot i is shared among the (possible) plot parts in the proportions of the assessed areas of the plot parts. The total weight of a field plot part i_l (belonging to map stratum h), $i \in J_{l,h}$, to land-use class (or FRYL subclass) l for municipality U is therefore

[2]
$$c_{i_l,U_h} = a a_{i_l} \sum_{p_h\in U_h} w_{i,p_h}$$

where a is the area of a pixel and a_{i_l} is the share of field plot i belonging to field land-use class (or FRYL subclass) l with $\sum_l a_{i_l} = 1$.

The area estimator of (FRYL) class l within U is

[3]
$$\hat{A}_{U,l} = \sum_h \sum_{i\in J_{l,h}} c_{i_l,U_h}$$

and the mean volume estimator for timber assortment s of land-use class or FRYL class l

[4]
$$\hat{v}_{U,l} = \frac{\sum_h \sum_{i\in J_{l,h}} c_{i_l,U_h} v_{i_l,s}}{\sum_h \sum_{i\in J_{l,h}} c_{i_l,U_h}}$$

where $v_{i_l,s}$ is the volume per hectare of the timber assortment s on the plot part i_l.

All field plots, regardless of the land-use class, are used in the estimation process. A subset of the plots, for example, entirely belonging to either FRYL or non-FRYL, can also be used. The estimates of the land-use areas and forest variables are computed simultaneously.

Calibrated MS-NFI estimators

The oMS-NFI estimates are calibrated based on large-area estimates of map errors in the calibration method cMS-NFI (Katila et al. 2000). The applied map strata are assumed to be homogeneous with respect to the "map errors". The proportions of land-use classes for small areas U are estimated by applying the proportions $\widehat{P}_{R_h,l}$ estimated from a larger area R (synthetic estimation (Rao 1998)).

[5]
$$A^*_{U,l} = \sum_h \widehat{P}_{R_h,l} A_{U_h}$$

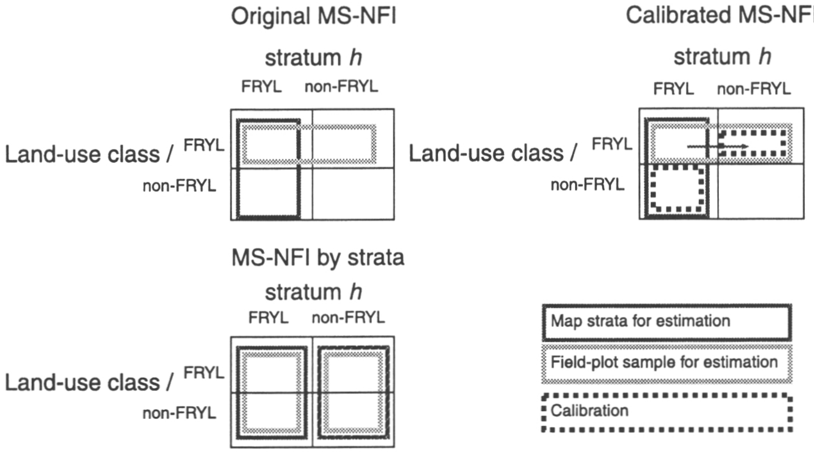
The aggregates of small-area estimates of land-use class areas over large areas are equal to unbiased post-stratification estimates (Holt and Smith 1979).

A method to compute the calibrated field-plot weights is found in Katila et al. (2000). The field plots for k-NN estimation are chosen for cMS-NFI in the same way as for the basic MS-NFI. The calibration typically increases the mean volume estimates and decreases the FRYL area estimates of small areas when the FRYL is overestimated by the map data (Katila et al. 2000).

Validation of the results

The main emphasis in this study is in the validation of the municipal-level estimates. An analytical method for estimating the standard errors of the MS-NFI small-area estimates has not yet been presented. The statistically validated estimates and their standard errors based on the field-inventory method are therefore used for comparisons. The relative standard error of

Fig. 3. Training data selection and map strata in the estimation in three different MS-NFI versions.



the area estimates of a stratum with an area of 2000 km^2 is usually not more than 5% and with an area of 10 000 km^2 is not more than 2% (Tomppo et al. 1998). The standard errors are estimated using the method of the operative NFI, which apply local quadratic forms (Matérn 1960). The aggregates of MS-NFI municipality estimates are compared with the field-inventory estimates and the standard errors from the same area (Katila et al. 2000).

Some estimation parameters have to be selected in the MS-NFI method. Examples are the value of k and the pixel-dependent geographical horizontal and vertical reference areas, that is, the area from which the field plots are applied in the estimation (Katila and Tomppo 2001). A leave-one-out cross-validation method and the root mean square error

$$[6] \qquad \mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

has been applied as a measure of reliability with the continuous variables in the selection. In eq. 6, y_i and $\hat{y}_i$ ($i = 1, ..., n$) are the observed and estimated values of the variables, respectively. The estimates of biases and the standard error of biases are used as further criteria (Katila and Tomppo 2001).

The two main goals with the new sMS-NFI method are to yield accurate FRYL and non-FRYL area estimates and accurate forest-variable estimates within FRYL. Thus the effect of the k-NN estimation parameters on the estimates of FRYL and non-FRYL classes within each stratum have to be examined. A 2×2 confusion matrix for FRYL and non-FRYL classes is calculated for the cross-validation. A fuzzy approach is used: the weights w_{i,p_h} obtained for the k spectrally nearest field plots are considered fuzzy membership values of the pixel p_h to be classified (Zhang and Foody 1998). The global sum of these weights are used in the 2×2 confusion matrices.

Within each stratum, the estimation parameters should yield a high overall classification accuracy (CC) and preserve the marginal distribution of the FRYL proportion in the field plot data, i.e., an unbiased estimate of FRYL area. The parameters should also minimize the MSE and give unbiased estimates of volumes.

Results

The accuracy of the applied land-use maps

The accuracy of the applied map data in stratifying different land-use classes is first discussed. The proportion of the FRYL field plot centre points within each stratum can be used as a measure of accuracy. The proportions and number n_{R_h} of the field plot centre points within each stratum are given in Tables 2 and 3. The FRYL area within the training data area was overestimated by 1.5 and 2.6% for central Finland and western Finland, respectively (Tables 2 and 3). The water stratum was the most accurate in separating the FRYL, while the combined built-up land and roads stratum was worst with 20–30% of FRYL field plots. The new agricultural area mask for the western Finland training data area increased the accuracy of the stratum compared with the central Finland training data area. The built-up land and road map stratum could be divided into a more specific stratum of houses, urban areas, and other built-up land and a second stratum for roads etc. However, this would decrease the number of field plots available for the training data set to considerably less than 500 for those strata.

The pixel-level accuracy of forestry-land estimates within map strata and the selected parameters for k-NN estimation

Parameters were selected on the basis of the pixel-level estimates. The goal was to obtain accurate FRYL and total volume estimations by strata. The parameters tested were the pixel-dependent geographical horizontal reference area radius (HRA), the number of nearest neighbours, k, and the power of spectral distances t. A suitable HRA was expected to be related to the proportion of stratum within the image area (Katila and Tomppo 2001). The results for selected HRA and k are summarized in Tables 4 and 5.

The power of the Euclidean distance measure $d_{p_{h,i},p_h}$ had only minor effects on the results. Weighting of the spectral

Table 2. Land-use class distribution among field plots by map strata in the central Finland study area.

		NFI land-use class, l								
		Forestry		Arable		Built-up, etc.		Water		Total
	Stratum, h	%	$n_{R_h,l}$	%	$n_{R_h,l}$	%	$n_{R_h,l}$	%	$n_{R_h,l}$	n_{R_h}
(1)	Forestry land, mineral	92.4	5009	4.1	132	2.4	222	1.1	58	5421
(2)	Forestry land, peatland	98.2	1220	0.4	5	0.7	9	0.6	8	1242
(3)	Arable	11.5	95	84.6	699	3.5	29	0.36	3	826
(4)	Built-up land and roads	30.8	148	10.2	49	58.5	281	0.4	2	480
(5)	Water	2.7	46	0.06	1	0.47	8	96.75	1635	1690
	Total	67.48	6518	9.17	886	5.68	549	17.66	1706	9659

Table 3. Land-use class distribution among field plots by map strata in the western Finland study area.

		NFI land-use class, l								
		Forestry		Arable		Built-up, etc.		Water		Total
	Stratum, h	%	$n_{R_h,l}$	%	$n_{R_h,l}$	%	$n_{R_h,l}$	%	$n_{R_h,l}$	n_{R_h}
(1)	Forestry land, mineral	92.0	3305	2.7	96	4.8	171	0.6	21	3593
(2)	Forestry land, peatland	97.6	1404	0.5	7	1.6	23	0.3	4	1438
(3)	Arable	3.0	37	95.6	1198	1.4	18	0	0	1438
(4)	Built-up land and roads	21.7	81	14.2	53	63.8	238	0.3	1	373
(5)	Water	0.5	5	0.4	4	0.4	4	98.8	1025	1038
	Total	62.8	4832	17.7	1358	5.9	454	13.7	1051	7695

bands was not used. The global estimates of FRYL and volume were not dependent upon the value of k. The overestimation of the FRYL area on the mineral soil and peatland strata remained despite the changes in the estimation parameters (Tables 4 and 5). FRYL area estimation within the built-up stratum performs better than the pure map stratification based estimate in the western Finland training data area (Table 5). The selected estimation parameters for the oMS-NFI and cMS-NFI are presented in the Table 6.

The errors of the minor land-use class at field-plot level, for either FRYL or non-FRYL, within each stratum were high (producer's accuracies were low), though the marginal distributions were more or less equal. Overall, CC percentages of the FRYL and non-FRYL classification from the cross-validation are 2.7 and 1.3% lower than the CC percentages from the map-based stratification for the central Finland and western Finland training data areas, respectively (Tables 4 and 5). However, the CC percentage of the map-based stratification is calculated from the field plot centre points, whereas all parts of the plots are used in the cross-validation. The FRYL classification within strata is in some cases poor, but a misinterpretation can be expected to occur between open FRYL plots and non-FRYL plots, whereas the map-based delineation may classify all kinds of forests to non-FRYL.

An important source of bias in the land use class estimation and other k-NN estimation is the mixed pixels between FRYL and non-FRYL. The cross-validation of field plots divided between FRYL and non-FRYL leads to a considerable overestimation of FRYL. Conversely, the mean volumes are significantly underestimated. (Tokola and Kilpeläinen 1999) reported slight overestimation of mean volume for NFI field plots nearest to the forest-stand boundaries in their cross-validation study applying only NFI field plots within FRYL. However, the volume for field plots was underestimated where the shape of the nearest stand edge was sharp. The divided land-use field plots have spectral values from mixed pixels, and small locational errors may change the spectral values attached to the field plots.

Estimates by municipalities

Both the cMS-NFI and the sMS-NFI slightly decreased the FRYL area estimates compared with the oMS-NFI estimates for all municipalities, except for the very small ones in central Finland (Fig. 4*a*). The relative decrease of the FRYL area is greater in the western Finland study area, –6.5 to 7.0% for sMS-NFI and –6.5 to 0.3% for cMS-NFI (Fig. 5*a*). The land-use map data also gives a greater overestimate of the FRYL area for this image area. The municipalities with small areas, e.g., cities and towns, have large proportions of non-FRYL areas, e.g., the built-up land stratum. The map usually overestimates this area wherefore the non-FRYL map strata contain a large amount of FRYL field plots. Consequently, both the sMS-NFI and cMS-NFI methods increase the FRYL area estimate in these municipalities.

The cMS-NFI systematically increases the mean volume estimates of FOWL by a few percentage points in both study areas, whereas the sMS-NFI changes the estimates both upwards and downwards compared with the oMS-NFI. Changes range from –3.9 to 5.5% (Figs. 4*b* and 5*b*). On average, the sMS-NFI does not increase the western Finland study area mean volumes.

In the western Finland study area, a gradual change takes place from pine- to spruce-dominated forests in an east–west direction. The mean volumes of pine and spruce therefore change by municipalities and groups of municipalities (cf. Figs. 9 and 10). The sMS-NFI seems to follow these changes better than the cMS-NFI and particularly better than the oMS-NFI (Figs. 4*c*, 4*d*, 5*c*, and 5*d*). Total volume estimates of the sMS-NFI method are, on average, smaller than the oMS-NFI estimates in the western Finland study area. This is due to the fact that area estimates

Table 4. Pixel-level errors by strata for central Finland: 2 × 2 confusion matrix and correctly classified (CC) FRYL and non-FRYL (nFRYL) percentages and mean volume ($\overline{m}$), root mean square error (RMSE), and bias of mean volume estimate and applied values of geographical horizontal reference area (HRA) and k.

		Cross-validation						
Stratum	NFI land use	$\widehat{\text{FRYL}}$, %	$\widehat{\text{nFRYL}}$, %	CC, %	$\overline{m}$, m³/ha	RMSE, m³/ha	Bias, m³/ha	HRA (k), km
Mineral soil	FRYL	87.0	5.2	88.5	115.2	99.4	0.39	50 (2)
	nFRYL	6.3	1.5					
Peatland	FRYL	96.3	1.6	96.3	84.7	69.1	1.01	60 (5)
	nFRYL	2.0	0.0					
Arable land	FRYL	4.3	7.1	85.8	9.8	38.0	–0.95	50 (3)
	nFRYL	7.1	81.5					
Built-up land, roads	FRYL	17.1	14.3	70.3	44.3	76.7	2.00	70* (5)
	nFRYL	15.4	53.2					
Water	FRYL	0.8	1.5	96.7	5.8	34.1	0.08	50 (6)
	nFRYL	1.7	95.9					
Overall	FRYL	62.6	4.7	89.8				
	nFRYL	5.5	27.2					
Overall from map†	FRYL	64.5	3.0	92.5				
	nFRYL	4.5	28.0					

*Euclidian distance weighting $t = 2$ applied.

†Estimate is based on map data.

Table 5. Pixel-level errors by strata for western Finland: 2 × 2 confusion matrix and correctly classified (CC) FRYL and non-FRYL (nFRYL) percentages and mean volume ($\overline{m}$), root mean square error (RMSE), and bias of mean volume estimate and applied values of geographical horizontal reference area (HRA) and k.

		Cross-validation						
Stratum	NFI land use	$\widehat{\text{FRYL}}$, %	$\widehat{\text{nFRYL}}$, %	CC, %	$\overline{m}$, m³/ha	RMSE, m³/ha	Bias, m³/ha	HRA (k), km
Mineral soil	FRYL	87.5	4.3	90.0	96.5	82.7	–0.46	40 (2)
	nFRYL	5.7	2.5					
Peatland	FRYL	95.8	1.7	96.0	57.1	52.9	–1.78	60 (3)
	nFRYL	2.3	0.1					
Arable land	FRYL	0.2	2.7	95.0	2.1	16.1	–0.75	40 (2)
	nFRYL	2.2	94.8					
Built-up land, roads	FRYL	12.9	7.4	84.8	24.2	45.9	1.00	70 (5)
	nFRYL	7.8	71.9					
Water	FRYL	0.0	0.6	99.0	0.0	7.7	0.26	50 (5)
	nFRYL	0.5	99.0					
Overall	FRYL	59.4	3.2	92.9				
	nFRYL	3.9	33.5					
Overall from map*	FRYL	61.2	2.5	94.2				
	nFRYL	4.2	33.0					

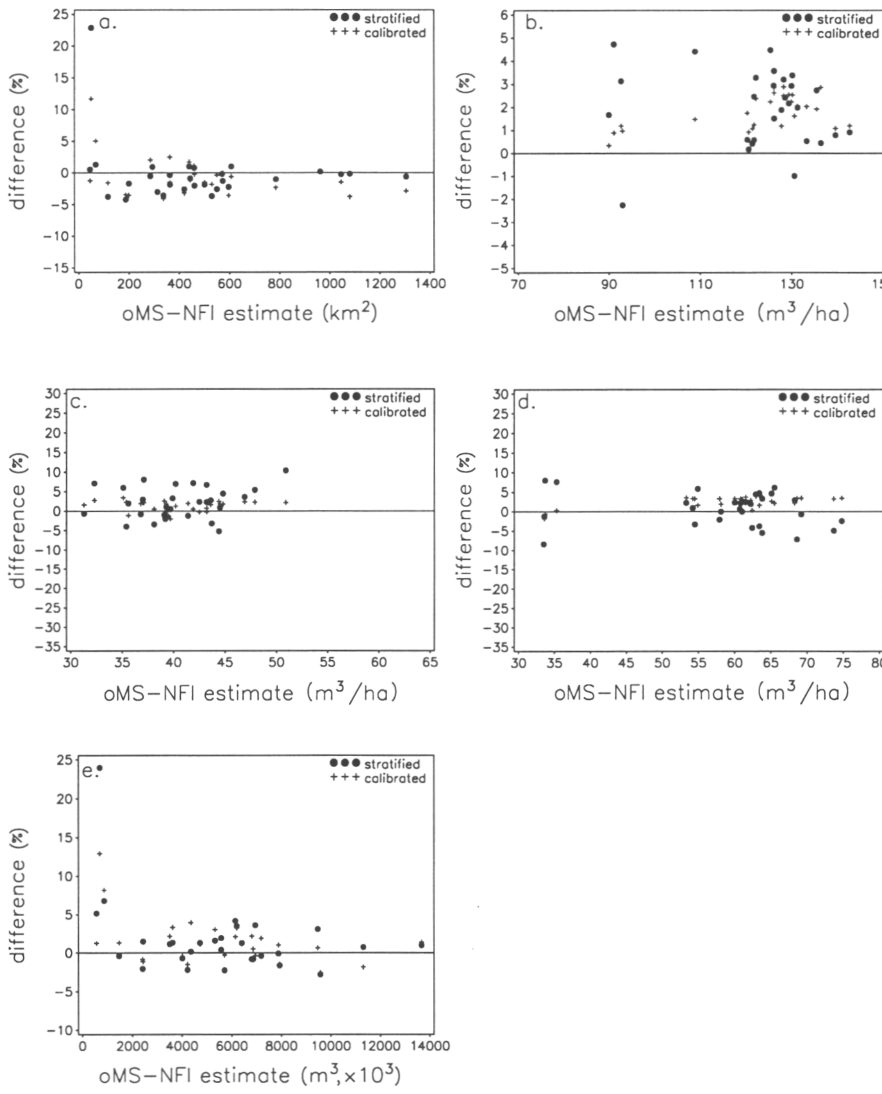
*Estimate is based on map data.

Table 6. Estimation parameters for oMS-NFI and cMS-NFI in central and western Finland study areas: geographical horizontal reference area radius (HRA) and value of k.

Study area, image	Stratum	HRA (km)	k
CF, 188/16	Mineral soil and peatland	75 × 45*	7
CF, 188/17	Mineral soil and peatland	60	7
WF, 191/16	Mineral soil	40	8
	Peatland	60	7

*A rectangular geographical HRA was applied, east–west × north–south distances.

Fig. 4. Percent difference between sMS-NFI and oMS-NFI estimates and cMS-NFI and oMS-NFI estimates for each municipality plotted against the oMS-NFI estimates for the central Finland study area for (*a*) area of FRYL (km^2), (*b*) mean volume (m^3/ha), (*c*) mean volume of pine (m^3/ha), (*d*) mean volume of spruce (m^3/ha), and (*e*) total volume (m^3, $\times 10^3$).



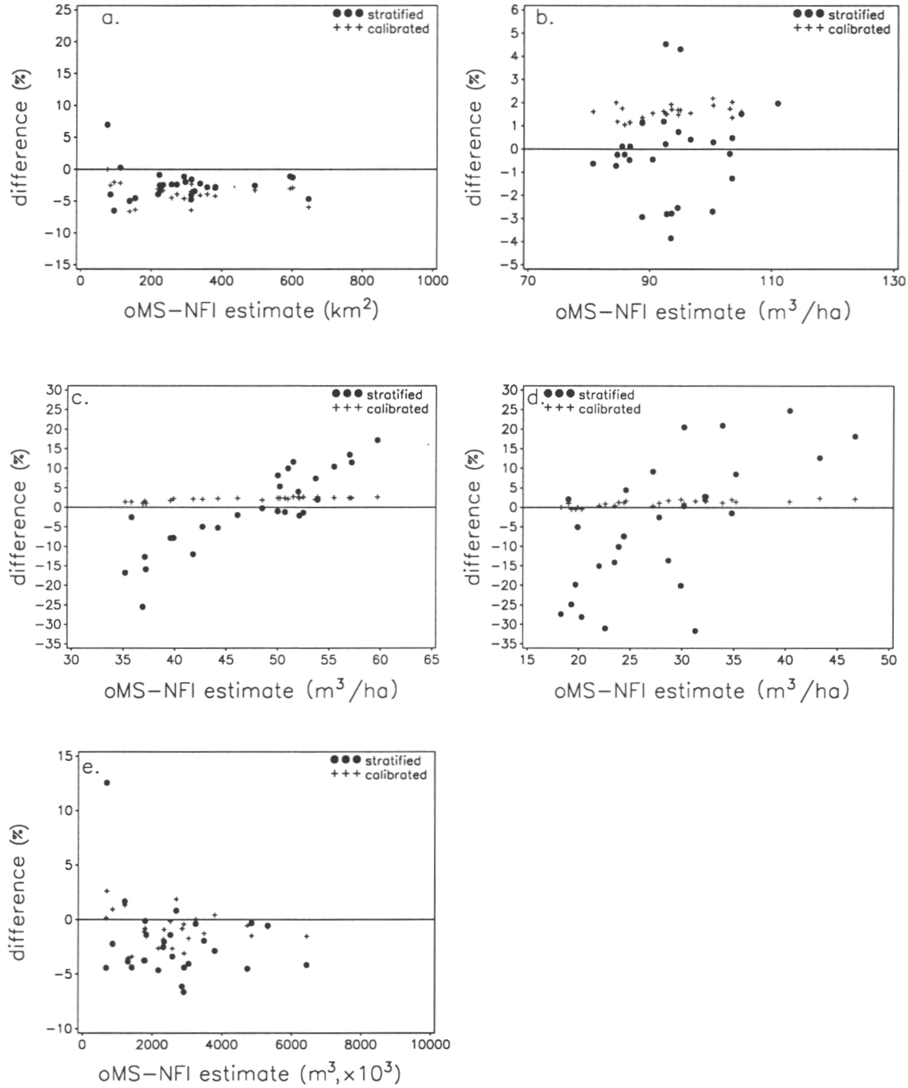
decrease, while the mean volume estimates do not increase correspondingly (Figs. 4*e* and 5*e*).

As a conclusion, the sMS-NFI volume estimates deviate from the oMS-NFI estimates more than the cMS-NFI estimates do. Figures 9 and 10 show that sMS-NFI is closer than the cMS-NFI estimates to the field data based estimates. This suggests that the sMS-NFI performs better than the cMS-NFI.

Bias by groups of municipalities

The subregions (groups of municipalities) are large enough to enable a comparison of the field-inventory error estimates with the MS-NFI estimates. A possible bias of the small-area estimates was studied in nine groups of municipalities within the central Finland (subregions 1–5) and western Finland (subregions 6–9) study areas (Fig. 6). The size of the subregions varied from 1738 to 4238 km^2 FRYL. The field-inventory estimates and standard errors of the percentage of FRYL area

Fig. 5. Percent difference between sMS-NFI and oMS-NFI estimates and cMS-NFI and oMS-NFI estimates for each municipality plotted against the oMS-NFI estimates for the western Finland study area for (*a*) area of FRYL (km^2), (*b*) mean volume (m^3/ha), (*c*) mean volume of pine (m^3/ha), (*d*) mean volume of spruce (m^3/ha), and (*e*) total volume (m^3, $\times 10^3$).



(Fig. 7), mean volume, total volume (Fig. 8), and mean and total volume for pine, spruce, and birch (Figs. 9, 10, and 11) were plotted for comparison.

The sMS-NFI and cMS-NFI did not show notable systematic errors in the percentage of FRYL, total volume, and mean volume estimates (Figs. 7, 8*a*, and 8*b*). The oMS-NFI estimate of the percentage of FRYL was significantly biased (4.4%) for subgroup 2. The sMS-NFI and cMS-NFI estimates corrected the percentage of FRYL, total volume, and mean volume estimates towards the field-inventory estimates for most subgroups.

In subregions 4 (central Finland), 6, 7, and 9 (western Finland) significantly biased estimates occurred for the mean and total volumes of pine and spruce with the oMS-NFI and cMS-NFI methods (Figs. 9*a*, 9*b*, 10*a*, and 10*b*). The sMS-NFI reduced the biases. The birch estimates were significantly biased for subregions 1, 7, and 8 with the oMS-NFI and cMS-NFI methods (Figs. 11*a* and 11*b*), while sMS-NFI slightly reduced the biases also in these cases.

The same systematic difference between sMS-NFI and the other two methods is clearly seen in the pine and spruce volume

Fig. 6. The nine groups of municipalities within the central Finland (1–5) and western Finland (6–9) study areas.

Fig. 7. Percentage of FRYL (%) of the land area obtained from pure field data estimates (±2 SE) and oMS-NFI estimates, cMS-NFI estimates, and sMS-NFI estimates for the groups of municipalities (Fig. 6) from the central Finland and western Finland study areas.

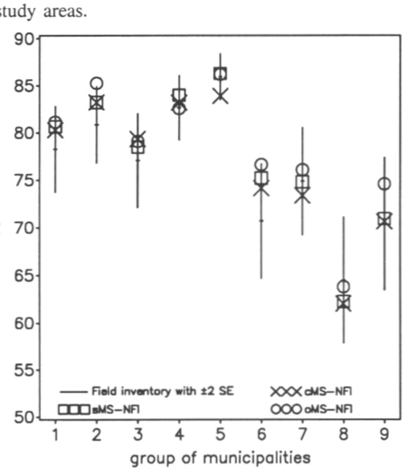
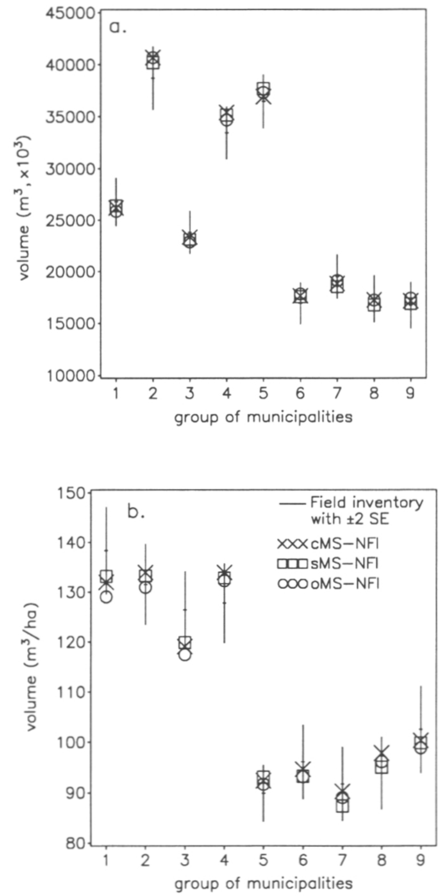


Fig. 8. Pure field data estimates (±2 SE) and oMS-NFI estimates, cMS-NFI estimates, and sMS-NFI estimates of (*a*) total volume (m^3, $\times 10^3$) and (*b*) mean volume (m^3/ha) for the groups of municipalities (Fig. 6) from the central Finland and western Finland study areas.



estimates for the western Finland study area subregions, as well as for the municipal-level estimates (Figs. 5*c* and 5*d*). The differences between estimates for sMS-NFI and for oMS-NFI are higher than those between oMS-NFI and cMS-NFI. This may be due to the different nature of the methods: the cMS-NFI is a kind of calibration method. The sMS-NFI produces less biased estimates for the volumes by tree species in the subregions than the oMS-NFI or the cMS-NFI. The mean volume estimates for pine ranged from 32.2 to 62.3 m^3/ha (sMS-NFI) and from 33.8 to 56.8 m^3/ha (cMS-NFI) and those for spruce ranged from 17.9 to 65.9 m^3/ha (sMS-NFI) and from 21.2 to 67.6 m^3/ha (cMS-NFI).

MS-NFI estimates for large regions

The estimates for the entire study areas based on the three methods (municipality estimates) were calculated and compared with the field inventory based estimates. The cMS-NFI and sMS-NFI shifted the area estimates of FOWL and FRYL towards the field-inventory estimate in the two study areas: decreases were 18 000 ha and 30 000 ha with cMS-NFI and 13 000 ha and 21 000 ha with sMS-NFI. The cMS-NFI estimates were closest to the field-inventory estimates. Both the cMS-NFI and sMS-NFI results were within two standard error of the field-inventory estimate (Table 7).

The cMS-NFI increases the regional mean volume estimates by 2.0 and 1.5 m^3/ha for the central Finland and western Finland study areas, respectively, compared with oMS-NFI, as well as increases the mean volume estimates by tree species. The sMS-NFI increases the mean volume estimate in the central Finland

Table 7. Area of FOWL and FRYL for the central Finland and western Finland study areas: pure field data estimate with sampling error and oMS-NFI, cMS-NFI, and sMS-NFI estimates.

	Central Finland area (ha, $\times 10^3$)				Western Finland area (ha, $\times 10^3$)			
Method	FOWL	SE	FRYL	SE	FOWL	SE	FRYL	SE
Field inventory	1343	13.8	1365	13.7	736	16.9	769	17.3
oMS-NFI	1375*		1394*		763		797	
cMS-NFI	1362		1376		742		767	
sMS-NFI	1362		1381		746		776	

Note: Estimates with asterisks deviate from the field data estimates by more than 2 SE.

Table 8. Mean volume of growing stock on FOWL for the study areas: pure field data estimate with sampling error and oMS-NFI, cMS-NFI, and sMS-NFI estimates.

		Mean volume (m^3/ha)									
Study area	Method	Pine	SE	Spruce	SE	Birch	SE	Other decidous	SE	Total growing stock	SE
Central Finland	Field inventory	38.6	1.0	54.3	1.5	18.7	0.5	6.7	0.4	118.4	1.7
	oMS-NFI	39.5		54.8		17.5*		5.6*		117.4	
	cMS-NFI	40.1		56.1		17.7		5.5*		119.4	
	sMS-NFI	40.1		54.9		18.0		6.4		119.4	
Western Finland	Field inventory	49.0	1.4	27.9	1.3	16.1	0.7	2.9	0.3	95.8	1.9
	oMS-NFI	48.6		27.8		15.2		2.4		94.1	
	cMS-NFI	49.8		28.1		15.3		2.4		95.6	
	sMS-NFI	49.4		26.7		14.8		2.7		93.5	

Note: Estimates with asterisks deviate from the field data estimates by more than 2 SE.

Table 9. Total volume of growing stock on FOWL for study areas: pure field data estimate with sampling error and oMS-NFI, cMS-NFI, and sMS-NFI estimates.

		Total volume (m^3, $\times 10^3$)									
Study area	Method	Pine	SE	Spruce	SE	Birch	SE	Other decidous	SE	Total growing stock	SE
Central Finland	Field inventory	51.8	1.5	73.0	2.1	25.2	0.7	9.0	0.5	159.0	2.8
	oMS-NFI	54.3		75.3		24.0		7.7*		161.4	
	cMS-NFI	54.6		76.4		24.1		7.5*		162.5	
	sMS-NFI	54.6		74.8		24.5		8.7		162.5	
Western Finland	Field inventory	36.0	1.3	20.5	1.1	11.8	0.6	2.1	0.3	70.5	2.1
	oMS-NFI	37.0		21.1		11.6		1.9		71.6	
	cMS-NFI	36.9		20.8		11.4		1.8		70.9	
	sMS-NFI	36.8		19.8		11.0		2.0		69.7	

Note: Estimates with asterisks deviate from the field data estimates by more than 2 SE.

by 2.0 m^3/ha and decreases it by 0.6 m^3/ha in the western Finland study area. The sMS-NFI did not systematically increase the mean volume estimates by tree species, as in the case of the cMS-NFI. The sMS-NFI gave the most accurate results for the broad-leaved volumes in the central Finland study area, while the oMS-NFI and cMS-NFI produced significantly biased results for the other deciduous species volume (Table 8).

The behaviour of the cMS-NFI and sMS-NFI total volume estimates followed that of the mean volume estimates. The other deciduous species volume estimate errors were also significant with the oMS-NFI and cMS-NFI in the central Finland study area (Table 9).

Discussion

A new multisource forest inventory method (sMS-NFI) is presented to produce forest parameter estimates and to reduce the effect of incorrect map data on the estimates. Estimates are computed by map strata. The new method has the advantage of including all the sample plots within each stratum in the training data. The method therefore resembles the one used in the field-inventory estimation. Only FRYL field plots were employed in the oMS-NFI and cMS-NFI and the plots intersecting FRYL and non-FRYL boundary were excluded (Fig. 3). The sMS-NFI estimates were compared with the ones from the oMS-NFI and the cMS-NFI (Katila et al. 2000).

The sMS-NFI reduced the bias in the FRYL area estimates of oMS-NFI that were caused by errors in the map data. The FRYL area estimates from sMS-NFI for large regions remained between the oMS-NFI estimates and the cMS-NFI estimates. The cMS-NFI region estimates were equal to the FRYL area estimates based on post-stratification (Katila et al. 2000). The sMS-NFI may either increase or decrease the mean volume estimates of large regions compared with the field-inventory

Fig. 9. Pure field data estimates (±2 SE) and oMS-NFI estimates, cMS-NFI estimates, and sMS-NFI estimates of (*a*) total volume of pine (m^3, $\times 10^3$) and (*b*) mean volume of pine (m^3/ha) for the groups of municipalities (Fig. 6) from the central Finland and western Finland study areas.

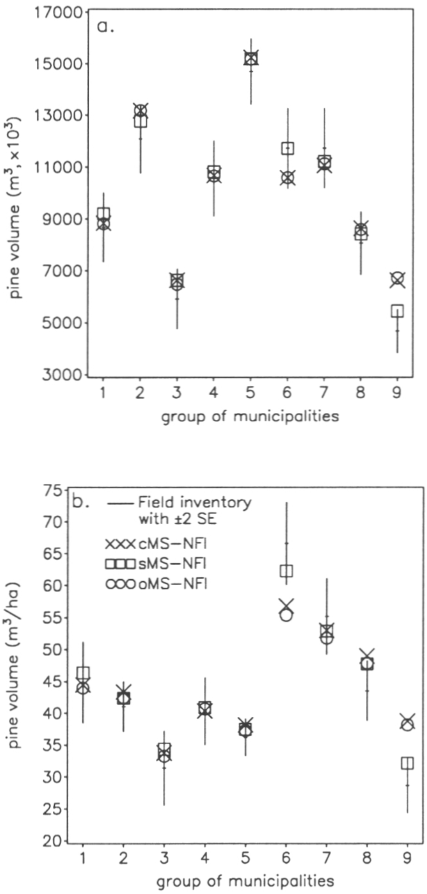
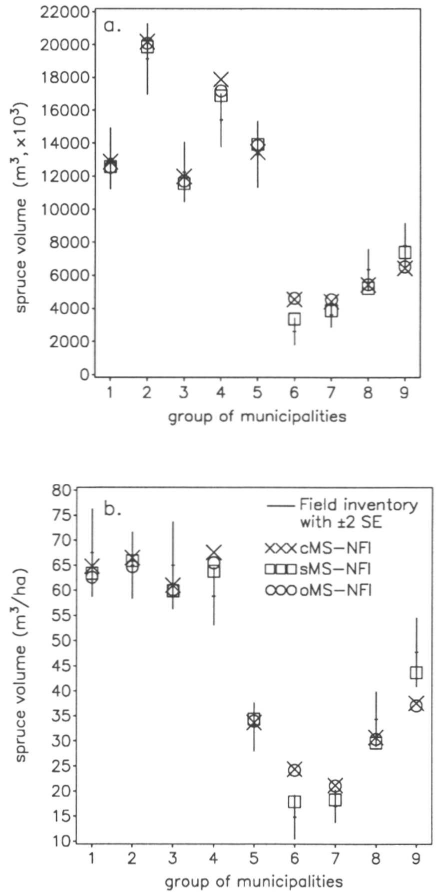


Fig. 10. Pure field data estimates (±2 SE) and oMS-NFI estimates, cMS-NFI estimates, and sMS-NFI estimates of (*a*) total volume of spruce (m^3, $\times 10^3$) and (*b*) mean volume of spruce (m^3/ha) for the groups of municipalities (Fig. 6) from the central Finland and western Finland study areas.



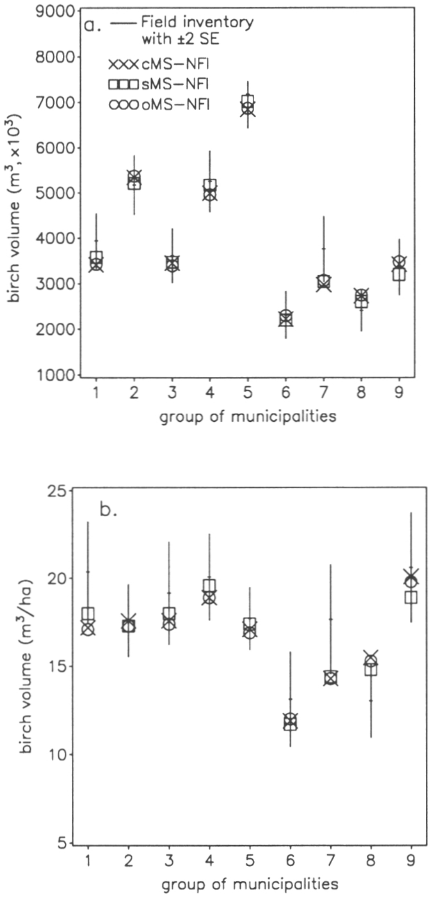
estimates, whereas the cMS-NFI typically increases these estimates. SMS-NFI volume estimates are within the two standard error of the field-inventory estimates. The volume estimates by tree species are more accurate for sMS-NFI than for oMS-NFI or cMS-NFI in large regions and subregions.

The municipal-level estimates of mean and total volumes, and mean and total volumes by tree species based on sMS-NFI differ more from the oMS-NFI estimates than those based on cMS-NFI; the cMS-NFI calibrates the MS-NFI estimates more or less systematically upwards or downwards from oMS-NFI. The sMS-NFI is essentially a different estimation method. The sMS-NFI and cMS-NFI estimates of FRYL area and volume did not produce significant errors when compared with the field-inventory estimates of subregions (1728–4238 km^2 FRYL). The sMS-NFI estimates of mean and total volume by tree species were more accurate compared with the field-inventory estimates than the two other MS-NFI methods in subgroups of municipalities: oMS-NFI estimates and cMS-NFI estimates failed to accurately follow the dominant tree species changes within the the western Finland study area.

The field-inventory estimates and their standard errors have proven to be useful in validating the MS-NFI estimates in large regions and subregions (i.e., in areas of 200 000 ha or greater) (Katila et al. 2000; Tomppo and Katila 1992). These estimates can be calculated for several combinations of municipalities to evaluate the MS-NFI estimates. The relative standard errors of the mean volume estimates, e.g., in the applied subregions, varied from 3.0 to 4.2%.

Since the sMS-NFI field plot data set contained all the field plots, the small-area estimates may be closer to the field inventory based estimates, even though the sMS-NFI estimator would not be very accurate.

Fig. 11. Pure field data estimates (±2 SE) and oMS-NFI estimates, cMS-NFI estimates, and sMS-NFI estimates of (*a*) total volume of birch (m^3, $\times 10^3$) and (*b*) mean volume of birch (m^3/ha) for the groups of municipalities (Fig. 6) from the central Finland and western Finland study areas.



The sMS-NFI gave unbiased results compared with large area field inventory results. Although it removed the bias at the region and subregion levels, it did not necessarily improve the accuracy at field-plot level. The oMS-NFI (FRYL delineation from map) gave a field-plot level FRYL %CC of 93–94%, but it overestimated the FRYL area. The overall %CC of FRYL and non-FRYL based on *k*-NN estimation and cross-validation at the pixel level was slightly poorer, 90–93%. The producer's and user's accuracies of the minor land-use class within strata were poor. However, the marginal distributions of the percentage of FRYL remained almost unchanged in most strata. Also, the mean volume estimates within strata were unbiased. In the cMS-NFI, the post-stratification probabilities of the FRYL proportions within each stratum in large areas were used to correct the FRYL area estimates afterwards (Katila et al. 2000), while in the sMS-NFI it is expected that the *k*-NN estimation will correct the FRYL area estimates directly in the estimation phase.

The pixel-level cross-validation results should be considered in a comparative way rather than in terms of absolute measures of reliability. The doubled effect of the locational error of field plots introduces conservative error estimates (cf., Verbyla and Hammond 1995; Halme and Tomppo 2001). On the other hand, cross-validation may underestimate errors in some cases (cf. Hammond and Verbyla 1996). The prediction error estimates of the cross-validation method may have a high variance. (Franco-Lopez et al. 2000) recommended bootstrap methods to obtain more stable variances (Efron and Tibshirani 1997).

In this test, only five strata in the sMS-NFI were employed, while in the cMS-NFI the number of strata was 6–11. The need for a sufficient amount of field plots in the training data limits the possibility to increase the number of strata in the sMS-NFI, whereas in the cMS-NFI, well-classified categories can have smaller field samples (Czaplewski and Catts 1992; Katila et al. 2000).

The suitable geographical HRA for each stratum was expected to be related to the proportion of stratum within the image area and the value of *k* to the number of field plots in the training data (Katila and Tomppo 2001). However, the FRYL area estimates within map strata in the cross-validation tests were not very sensitive to the value of *k* or the geographical HRA.

Region and subregion level estimates of sMS-NFI applying $k = 1$ or larger values of *k* were quite similar. In the operative inventory, reasonable small-area estimates are often obtained by estimating only a sample of pixels, e.g., every 10th pixel along lines and elements of satellite image and $k = 5$–10. Consequently, it is most probable that an unsampled MS-NFI estimation applying only $k = 1$ would also yield a sufficient amount of estimated neighbours for pixels for the municipal-level estimation of forest variables.

The method presented is statistically sound for removing the effect of the erroneous map data. The method is also computationally straightforward. It is flexible and can be utilized with ancillary data of varying quality. For instance, if the land-use map data is initially created using satellite image information, the effect of land use class errors on the final forest parameter estimates can be reduced by using the method presented.

Our first results are encouraging. An independent reliable forest inventory data would be needed to study the different MS-NFI methods in detail. The study areas should preferably cover different geographical and land-use combinations, as contradictory results from the two study areas were sometimes obtained.

Acknowledgements

We are grateful to Dr. Charles T. Scott for the initial suggestion of separate multisource estimations in each strata and we thank Dr. Juha Heikkinen, Lic. Tech. Kai Mäkisara, and two anonymous referees for their valuable comments. The English language was edited by Dr. Ashley Selby. The study was partly funded by the Graduate School Program "Forests in GIS", University of Helsinki.

References

Benediktsson, J.A., and Kanellopoulos, I. 1999. Classification of multisource and hyperspectral data based on decision fusion. IEEE Trans. Geosci. Remote Sens. **37**: 1367–1377.

Campbell, J.B. 1996. Introduction to remote sensing. 2nd ed. The Guilford Press, New York.

Czaplewski, R.L., and Catts, G.P. 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sens. Environ. **39**: 29–43.

Efron, B., and Tibshirani, R. 1997. Improvements on cross-validation — the .632+ bootstrap method. J. Am. Stat. Assoc. **92**: 548–560.

Franco-Lopez, H., Ek, A.R., and Bauer, M.E. 2000. Estimation and mapping of forest stand density, volume and cover type using the k-nearest neighbors method. Remote Sens. Environ. **77**: 251–274.

Gjertsen, A.K., Tomter, S., and Tomppo, E. 2000. Combined use of NFI sample plots and Landsat TM data to provide forest information on municipality level. *In* Remote Sensing and Forest Monitoring, Proceedings of IUFRO conference, 1–3 June 1999, Rogow, Poland. *Edited by* T. Zawila-Niedzwiecki and M. Brach. Office for Official Publications of the European Communities, Luxembourg. pp. 167–174.

Halme, M., and Tomppo, E. 2001. Improving the accuracy of multisource forest inventory estimates — a multi-criteria approach. Remote Sens. Environ. **78**: 321–327.

Hammond, T.O., and Verbyla, D.L. 1996. Optimistic bias in classification accuracy assessment. Int. J. Remote Sens. **17**: 1261–1266.

Holt, D., and Smith, T.M.F. 1979. Post stratification. J.R. Stat. Soc. Ser. A, **142**: 33–46.

Hutchinson, C.F. 1982. Techniques for combining landsat and ancillary data for digital classification improvement. Photogramm. Eng. Remote Sens. **48**: 123–130.

Katila, M., and Tomppo, E. 2001.Selecting estimation parameters for the Finnish multi-source national forest inventory. Remote Sens. Environ. **76**: 16–32.

Katila, M., Heikkinen, J., and Tomppo, E. 2000. Calibration of small-area estimates for map errors in multisource forest inventory. Can. J. For. Res. **30**: 1329–1339.

Kilkki, P., and Päivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. *In* Remote Sensing-Aided Forest Inventory. Proceedings of Seminars Organised by SNS, Dec. 10-12, 1986, Hyytiälä, Finland. University of Helsinki, Department of Forest Mensuration and Management, Helsinki, Finland. Res. Notes 19. pp. 209–215.

Loetsch, F., and Haller, K.E. 1973. Statistics of forest inventory and information from aerial photographes. BLVVerlagsgesellschaft, Munich.

Matérn, B. 1960. Spatial variation. Medd. Fr. St. Skogsf. Inst. **49**(5). [Also appeared as Lect. Notes Stat. 36.]

National Land Survey of Finland. 1996. Maastotietojen kohdemalli, Feature catalog of the topographic database. Maanmittauslaitoksen julkaisuja 71. National Land Survey of Finland, Helsinki. [In Finnish.]

National Land Survey of Finland. 1997. Suomen pinta-ala kunnittain 1. 1. 1997. National Land Survey of Finland, Helsinki. [In Finnish.]

Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne lidar. Ph.D. thesis, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics. Acta Univ. Agric. Suec. Silv. 17.

Poso, S. 1972. A method for combining photo and field samples in forest inventory. Commun. Inst. For. Fenn. **76**: 1–133.

Poso, S., Paananen, R., and Similä, M. 1987. Forest inventory by compartments using satellite imagery. Silva Fenn. **21**: 69–94.

Rao, J.N.K. 1998. Small area estimation. *In* Encyclopedia of statistical sciences, update Vol. 2. *Edited by* S. Kotz, C.B. Read, and D.L. Banks. Wiley, New York. pp. 621–628.

Tokola, T., and Kilpeläinen, P. 1999. The forest stand margin area in the interpretation of growing stock using Landsat TM imagery. Can. J. For. Res. **29**: 303–309.

Tokola, T., Pitkänen, J., Partinen, S., and Muinonen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. Int. J. Remote Sens. **17**: 2333–2351.

Tomppo, E. 1991. Satellite image-based national forest inventory of Finland. *In* Proceedings of the Symposium on Global and Environmental Monitoring, Techniques and Impacts, 17–21 Sept. 1990, Victoria, B.C. Int. Arch. Photogramm. Remote Sens. 28. pp. 419–424.

Tomppo, E. 1996. Multi-source national forest inventory of Finland. *In* New Thrusts in Forest Inventory. Proceedings of the Subject Group S4.02-00 "Forest Resource Inventory and Monitoring" and Subject Group S4.12-00 "Remote Sensing Technology". Vol. 1. IUFRO XX World Congress, 6–12 Aug. 1995, Tampere, Finland. *Edited by* R. Päivinen, J. Vanclay, and S. Miina. European Forest Institute, Joensuu, Finland. pp. 27–41.

Tomppo, E., and Katila, M. 1992. Multi-source national forest inventory of Finland. Int. Arch. Photogramm. Remote Sens. **29**: 671–674.

Tomppo, E., Varjo, J., Korhonen, K., Ahola, A., Ihalainen, A., Heikkinen, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S., and Tuomainen, T. 1997. Country report for Finland. *In* Study on European forestry information and communication systems. Reports on forestry inventory and survey systems. Vol. 1. Office for Official Publications of the European Communities, Luxembourg. pp. 145–226.

Tomppo, E., Henttonen, H., Korhonen, K.T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., Mikkelä, H., Tonteri, T., and Tuomainen, T. 1998. Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968–97. Folia For. 2B/1998: 293–374. [In Finnish.]

Tomppo, E., Goulding, C., and Katila, M. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scand. J. For. Res. **14**: 182–192.

Verbyla, D.L., and Hammond, T.O. 1995. Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. Int. J. Remote Sens. **16**: 581–587.

Zhang, J., and Foody, G.M. 1998. A fuzzy classification of sub-urban land cover from remotely sensed imagery. Int. J. Remote Sens. **19**: 2721–2738.

ISBN 951-40-1904-0
ISSN 0358-4283

Hakapaino Oy