7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.
Edited by M. Caetano and M. Painho.

Sampling simulation on multi-source output forest maps - an application for small areas

Matti Katila[1] and Erkki Tomppo

Finnish Forest Research Institute
Unioninkatu 40 A, FIN-00170 Helsinki, Finland
Tel.: +358 10 2111 ; Fax: +358102112104
matti.katila@metla.fi; erkki.tomppo@metla.fi

Abstract

Systematic samples were simulated on multi-source National Forest Inventory (MS-NFI) output thematic forest maps obtained by k-nearest neighbours method (k-NN) for areas of 1 km^2 and 100 km^2 . The standard deviations based on the simulated sample means were used as estimates for the standard errors (SE) of the particular designs. The variables of interest were the mean tree stem volume (m^3/ha) and the volumes by tree species. The effect of the temporal and thematic accuracy of the forest maps was studied with respect to the precision of the SE estimates from simulations. The simulated SEs were validated against an independent field inventory data measured from three test sites of 1 km^2 and seven sites of 100 km^2 in Eastern Finland. The effect of the estimation parameters of k-NN method on the error estimates was also studied. The study showed that the Finnish MS-NFI thematic maps can be used to estimate sampling errors for systematic field sampling designs with a plot distances down to 75-100 m on areas of 1 km^2 and larger.

Keywords: multi-source forest inventory, sample simulation, standard error, Landsat TM

1 Introduction

The Finnish multisource national forest inventory (MS-NFI) has utilised optical area satellite images and digital maps in addition to field plot data to produce geo-referenced information, thematic maps and small area statistics since 1990. A non-parametric *k*-nearest neighbour method (*k*-NN) is used in the estimation. Field data from surrounding calculation units (municipalities), in addition to the unit itself, are utilised when estimating results for one unit (Tomppo, 1996; Tomppo and Halme, 2004).

Currently there is no analytical method in the Finnish MS-NFI for operatively estimating the error at the pixel level or for deriving the error estimates for small areas. However, recent developments in estimating these errors are promising (Kim and Tomppo, 2006; McRoberts *et al*., 2005). A field inventory was planned and carried out to get reliable and independent reference data to assess empirically the errors of MS-NFI small area estimates in 2000.

A sampling simulation is an efficient tool to study different sampling strategies for a particular inventory problem. It presumes that either statistical (Mackisack and Wood, 1990) or map form (Tomppo *et al*.1998a; Tokola and Shestra, 1999) models of the variables of interest exist. The MS-NFI volume maps have been successfully used to compare the sampling errors of different designs for large area field inventories since the 8th NFI (Henttonen, 1991; Tomppo *et al*., 1998*a*). A concern related to sampling simulation and the associated error estimates using map format predictions is whether or not the average variation and the covariance structure of the predictions are same as those of the original field variables.

The objectives of this study is to compare the error estimates derived from the sampling simulations based on MS-NFI output thematic maps and the measured field sample data. The effect of the *k*-NN estimation parameters and the temporal accuracy of the MS-NFI predictions on the simulated errors are studied. The ultimate goal is to test the relevance of the sampling simulation and error estimation method under consideration in seeking 'optimal' sampling design for variables in question. The variables of interest were the mean total volume and the volumes by tree species. The objective was to obtain a sampling scheme with a relative standard error (SE) of 2.5 % for volume estimates for the municipal level size test units, i.e. for units with an area of some 100 km^2. For smaller test units of 100 ha, a larger SE was acceptable. The test data should contain repeated observations (5-10) of the above size areas on a large geographical area, preferably of a size approximately of a Landsat TM image.

2 Materials

2.1 MS-NFI data sources

The study area is located between longitudes 29°39′E and 31°36′E and latitudes 61°21′N and 63°50′N (Figure 1). The study area consists largely of medium fertile mineral soils. The main tree species are Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.), with small proportions of birch (*Betula* spp.) and other deciduous species.

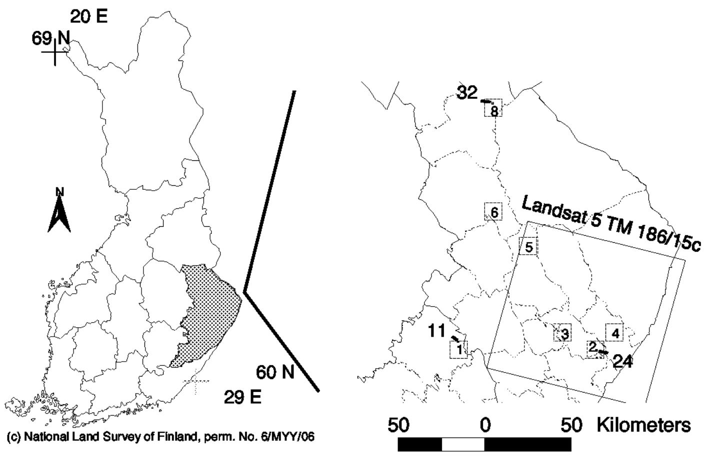

Figure 1 Location of the study area, the large and small test units measured in 2000 and the area covered by Landsat 5 TM image 186/15c, 29.7.1994.

The field samples of the 'updated' 8th NFI and 9th NFI were measured from systematically located clusters of sample plots in 1994 and 2000, respectively. In the updated 8th inventory, the sample plots (16 per cluster) were located along a L-shape tract at 200 m intervals. The distance between clusters was 14×8 km (Tomppo et al., 1998b). In the 9th inventory, there were 10-18 sample plots per cluster located along a rectangular or L-shape tract at 250 or 300 m intervals, depending on the area. The distance between clusters was 6×6 km or 7×7 km. Trees were measured from field plots belonging to forest land and poorly productive forest land stands. The tally trees were selected using PPS-sampling (sampling with probability proportional to size), applying a basal area factor of two. A maximum radius of 12.52 m was used in the 9th NFI while in the 8th NFI the PPS sampling was unrestricted. The inclusion of 'border' trees was carefully checked (Tomppo, 2006).

Images from Landsat 5 TM and Spot 2 XS satellites were used in the 8th MS-NFI (Tomppo *et al*., 1998b) whereas in the 9th MS-NFI the images were from Landsat 7 ETM+. 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 re-sampling of the images to 25 m × 25 m pixel size (Tomppo, 1996). The digital map data including the digital elevation model was obtained from the National Land Survey.

2.2 MS-NFI method

In the MS-NFI estimation forestry land (FRYL) is separated from other land uses on the basis of the digital map data. The satellite images and other supplementary data are used to find, for each pixel p within the FRYL, k most similar field plots in the training data set using the k nearest neighbours' method. Forest variable estimates are weighted sums or averages of field measurements in plots i belonging to the training data set. The weights of the selected field plots for the pixel p are defined applying the k-NN method (Tomppo 1996).

2.3 MS-NFI output thematic maps for sampling simulation

Three different sets of multi-source predictions were employed in simulations: 1) 'operative' MS-NFI8 (k=5-10, no reassignment), 2) reassigned MS-NFI8 (k=1, reassigned satellite image information) (Halme and Tomppo, 2001), and 3) MS-NFI9. The second set was used to test the effect of the reduced plot location errors on the simulated error estimates. The MS-NFI9 data was up-to-date and more accurate than MS-NFI8 data but available only after the field sample in the study area was measured and was used for comparison to MS-NFI8 data in simulations.

The operative MS-NFI8 thematic maps of stem volume and volume by tree species (pine, spruce, birch and other broadleaved trees) (Tomppo *et al*., 1998b) from the 'updated' 8th inventory were used in the basic sample simulation. A new multisource estimation was carried out for the image Landsat 5 TM 186/15c dated 29.7.1994 using updated 8th NFI data (Figure 1). The satellite image information was reassigned to field plot data in a 3×3 pixel window by a multi-criteria approach (Halme and Tomppo, 2001). This procedure reduces the effect of the locational errors in the training data and decreases the prediction errors, particularly for the total volume estimates. In addition, the use of only one nearest neighbour, i.e. one field plot, was tested in the estimation. The operative MS-NFI9 stem volume and volume by tree species maps were estimated using the improved k-NN method (Tomppo and Halme, 2004).

3 Methods

3.1 Test setting

Sampling was studied in two types of areas: an area (100 km^2) equivalent to a municipality, and an area (1 km^2) equivalent to a forest holding. The population to be sampled was the 9th NFI field inventory area in 2000, which was divided into 10×10 km^2 cells. The 103 grid cells that contained less than 20 % water were chosen for the population. The grid cells were sorted on the basis of the mean volume in the MS-NFI8 total volume map, and divided into eight strata with an equal number of cells. One cell of each stratum was selected to be a large test unit. The 1 km^2 test units were to be located within the larger units.

3.2 Sampling method for the test data and sample simulation

Systematic field plot samples were simulated from the MS-NFI thematic maps of total volume and volume by tree species. The samples were simulated for the eight test units and to smaller test units of size 100 ha within the eight units. Due to the pixel size in the MS-NFI maps, there were at maximum (line distance m)/25 × (plot distance m)/25 possible starting pixels, i.e. samples *i*, for the particular design. All possible samples (without replacement) were taken. The estimate of mean value $\bar{x}_{U,i}$ for the variable of interest was computed for the test units *U* from each sample. The sampling error (or SE) of the mean volumes on FRYL for a particular sampling design in the test unit was estimated from the standard deviation of the sample means $s(\bar{x}_{U,i})$. The ratio-of-means estimators used in the Finnish NFI were employed to calculate the forest variable estimates of each simulated sample (Tomppo, 2006). The estimator of the land use class *l* is

$$\hat{A}_{U,l} = \frac{n_{U,l}}{n_U} A_U \qquad (1)$$

where:

$n_{U,l}$ – is the (random) number of plots located on the FRYL mask within *U*,
n_U – is the number of plots located on the test unit,
A_U – is the total area of *U*.

In this case the interesting values for *l* were FRYL and non-FRYL. The mean volume within a land use class *l* (FRYL) of test unit *U* is estimated by the sample average

$$\hat{v}_{U,l} = \frac{\sum_{i \in I_{U,l}} v_i}{n_{U,l}} \qquad (2)$$

where :

v_i – is the mean volume in field plot *i* (MS-NFI pixel value),
$I_{U,l}$ – is the set of those sample plots within *U* that represent subclass *l (FRYL)*.

The total volume estimator is the product of the mean volume estimator and the FRYL area estimator.

3.3 Data processing and forest variable estimation of the measured field data

The field instructions of the 9th NFI were followed in the measurement of field plot data (Tomppo *et al*., 1998a). The volume estimation for tally trees and sample trees in the field plots was carried out as given in Tomppo (2006). The forest variable estimates for both the large and small measured test units were calculated by applying the ratio-of-means estimators presented above.

The SEs are estimated using local quadratic forms as presented by Matérn (1960) and as employed in the Finnish NFI (Heikkinen, 2006). The quantity $E(\hat{M} - M)^2$ is used as a measure of reliability of the ratio estimator $\hat{M} = \sum_i y_i / \sum_i x_i$, c.f. equations 1 and 2. Its estimator can be computed as follows. Instead of the cluster-wise residuals employed in the operative NFI, only one plot *i* per 'cluster' is employed and plot-wise residuals $z_i = y_i - \hat{M}x_i$ are used. The variance estimator for $\hat{M}$ finally takes the form

$$V\hat{a}r(\hat{M}) = \frac{q\sum_g T_g}{(\sum_i x_i)^2} \tag{3}$$

where the average of quadratic form $T_g = (z_{i1} - z_{i2} - z_{i3} + z_{i4})^2/4$ is the variance estimator per plot in a rectangular group *g* of four plots and *q* is the ratio between the number of plots and the number of groups (q=1 used in this study). See Heikkinen (2006) for more details.

4 Results

4.1 Sample simulation

4.1.1 100 km^2 test units

Systematic samples were simulated from the MS-NFI8 data for the selected eight test units using line distances ranging from 325 to 900 m and plot distances ranging from 200 to 475 m. The number of field plots on FRYL mask was expected to be between 500 and 1000 per test unit. The coefficient of variation (CV, i.e., an estimate of the relative SE) of mean volume estimates from all possible samples for each line and plot distance combination in the selected 100 km^2 test units are presented in Figures 2a and 2b. The CVs obtained from simulations from MS-NFI8 thematic maps predicted using reassigned training data and with *k* =1 were approximately 50 % larger than the CVs from operative MS-NFI8 data (Figure 2a) for the test units 268, 287 and 290. CVs of 3-3.5 % for the mean volume were obtained when samples of 600-700 field plots were simulated from the FRYL mask on the three test units (*k*= 1 and reassigned training data). The CV for the mean volume of pine, spruce and birch were 4.7-5.2 %, 6.5-7.5 % and 6.9-9.2 %, respectively.

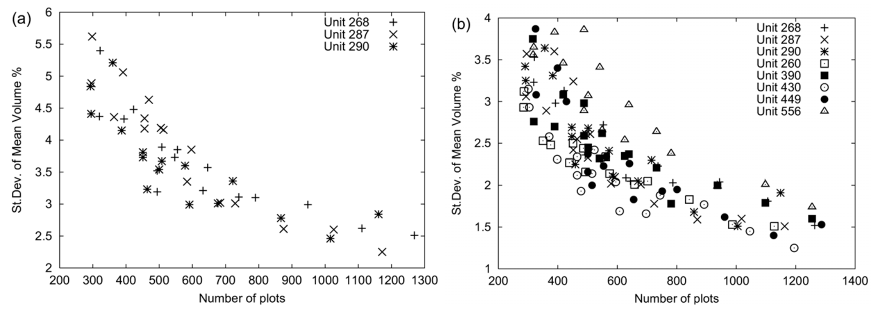

Figure 2 Coefficient of variation of mean volume estimates of all possible samples for various sample densities. MS-NFI8 estimates with reassigned training data and *k*=1 (three test units) (a), and operative MS-NFI8 data (b).

4.1.2 1 km^2 test units

Systematic samples from six areas of size 1.1×1.2 km^2 were simulated within each of the larger test units 268, 287 and 290; the MS-NFI8 thematic maps estimated employing reassigned training data and value of $k=1$ were employed. The distances between plots were between 75 m and 125 m in an east-west direction and between 50 m and 125 m in a north-south direction. Only 6-25 different samples could be taken from the MS-NFI thematic maps. Only the test units with FRYL area greater than 80 ha were accepted for further analysis. Approximately 140 field plots on FRYL was considered to be a minimum for the stated objectives (Figure 3). Naturally, the required sample size depends on the growing stock and its spatial variation within the specific forest area.

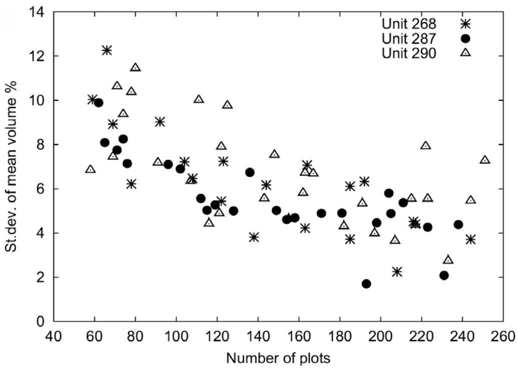

Figure 3 Coefficient of variation of mean volume estimates from all possible samples for various sample densities on 1.1×1.2 km^2 areas. 15 areas within test units 268,287 and 290, MS-NFI8 estimates with reassigned training data and k=1.

4.2 The measured field sample

Table 1 The measured field sample on large and small test units, number of field plots, total and on forestry land and total area.

new (old)index	No. of field plots total	No. of field plots FRYL	Total area, ha
1 (260)	825	680	9 900
2 (268)	808	724	9 696
3 (287)	712	607	8 544
4 (290)	825	686	9 900
5 (390)	825	746	9 900
6 (430)	825	710	9 900
8 (556)	825	722	9 900
11 a	182	178	117.6
24 b	140	137	147
32 b	140	138	147

The actual field sample was measured between August and October 2000. The number of plots by test units for the selected sampling designs are given in Table 1. In the end, field measurements were carried out on seven of the eight large test units using 400 m×300 m (north-south and east-west) plot distances corresponding to a total area of 9.9×10.0 km^2. Three small test units were measured using 80 m×75 m (unit 11) and 100 m×75 m (units 24 and 32) plot distances in north-south and east-west directions covering total areas of 1.05×1.12 km^2 and 1.05×1.4 km^2 respectively. The three small test units were purposely selected to represent lowest and highest mean volumes in the potential small test units. The location of the test units is presented in Figure 1 using new index numbers (Table 1). The samples were incomplete on test units 2 and 3. A total of 42 667 trees were tallied of which 5486 were sample trees.

4.3 Comparison of error estimates for the test units

Table 2 Mean volume of growing stock on FRYL and the SEs for the test units: field data-based estimates and corresponding simulation results using MS-NFI8 and MS-NFI9 thematic maps.

	Field inventory			MS-NFI8			MS-NFI9		
unit	Mean volume m^3/ha	SE, m^3/ha	SE, %	Mean volume m^3/ha	SE, m^3/ha	SE, %	Mean volume m^3/ha	SE, m^3/ha	SE , %
1	146.4	4.4	3.0	150.2	2.5	1.7	129.7	2.5	1.9
2	112.4	3.5	3.1	116.5	2.5	2.2	116.3	2.5	2.2
2[a]				103.8	3.3	3.2			
3	113.8	3.6	3.2	109.8	3.0	2.7	111.5	2.8	2.5
3[a]				91.3	3.5	3.8			
4	114.7	3.6	3.1	124.1	2.5	2.0	109.0	2.1	1.9
4[a]				116.4	3.7	3.2			
5	100.8	3.4	3.4	103.3	2.2	2.1	100.0	2.2	2.2
6	110.9	3.5	3.1	112.1	2.1	1.8	113.6	2.1	1.8
8	69.6	2.4	3.4	72.0	1.5	2.1	67.4	1.5	2.2
11	195.7	6.3	3.2	173.2	4.4	2.6	170.7	3.2	1.9
24	116.3	5.1	4.4	118.5	3.6	3.0	135.1	3.6	2.6
24[a]				104.9	4.1	3.9			
32	73.3	3.7	5.0	75.2	1.5	2.0	75.5	2.1	2.7

a MS-NFI8 estimates with reassigned training data and k= 1.

The SEs estimated for the test units were compared to the corresponding error estimates obtained from the simulation that employed operative and reassigned MS-NFI8 data and MS-NFI9 data.

There was a little variation in the SE estimates of the mean volume for the field data of the large test units; the relative standard error estimates were in the range of 3.0-3.4 % (Table 2. The SE estimates of mean volume from the simulations using operative MS-NFI8 and MS-NFI9 thematic maps were similar, but underestimated the sampling error (obtained from the field data) by approximately 25 %. SE estimates closer to the field data based SEs were obtained employing the MS-NFI8 thematic maps estimated using reassigned training data and k =1 for both large and small test units. However, the relative SE estimates were more suitable because the volumes were underestimated in these data. The mean volume estimates from MS-NFI8 agreed quite well with the field sample based estimates considering that the MS-NFI8 data was six years old.

The relative SE estimates of the mean volumes by tree species were of the same magnitude both from the MS-NFI8 and MS-NFI9 data for the large units (Table 3). The SE of volumes by tree species was underestimated in the simulation-based estimates of MS-NFI8 and MS-NFI9 thematic maps. The CVs of the sample means for the volume of birch and other broadleaved trees were particularly underestimates. In the small test units, there was more variation between the relative SE estimates obtained from the MS-NFI8 and MS-NFI9 data. As for the large test units, the relative SE estimates obtained from the MS-NFI8 data estimated using the reassigned training data and *k* =1 were closest to those estimated for the field sample.

Table 3 Relative SE of mean volume by tree species on FRYL for the test units: field data-based estimates and corresponding simulation results using MS-NFI8 and MS-NFI9 thematic maps for large and small test units.

	Field inventory, %				MS-NFI8, %				MS-NFI9, %			
	pine	spr.	birch	other broadl	pine	spr.	birch	other broadl.	pine	spr.	birch	other broadl.
1	5.1	4.6	7.1	18.5	2.9	3.1	4.3	8.4	2.9	3.4	3.5	9.0
2	4.6	5.6	7.0	16.2	3.0	4.5	4.1	7.8	2.5	4.5	3.3	7.9
2[a]					4.9	6.8	7.2	18.6				
3	4.5	6.9	8.9	15.5	3.1	5.7	4.7	8.4	2.7	4.9	4.0	6.5
3[a]					4.8	9.0	8.5	21.0				
4	5.9	5.0	6.3	15.5	2.9	3.8	4.0	8.4	3.3	3.7	3.2	6.9
4[a]					5.4	5.4	6.5	15.2				
5	4.8	6.2	6.2	15.9	2.8	4.3	4.0	7.8	3.0	4.5	3.3	7.5
6	5.0	5.5	6.0	11.5	2.3	3.6	3.3	6.5	3.1	4.1	3.0	5.9
8	4.3	7.6	7.2	20.1	2.4	6.1	4.6	9.9	2.3	5.3	4.0	11.2
11	11.3	4.6	11.0	29.1	3.3	5.2	4.4	8.0	3.0	4.0	8.2	11.6
24	10.4	6.4	14.4	28.3	4.5	4.9	4.0	16.4	5.8	6.4	8.1	12.0
24[a]					11.7	9.7	15.1	43.2				
32	6.7	16.5	21.1	49.6	3.1	8.6	7.7	19.1	2.1	8.7	5.2	16.7

a MS-NFI8 estimates with reassigned training data and *k*= 1.

The *k*-NN estimation was carried out using various parameter and training data combinations for the Landsat 5 TM 186/15c image. The aim was to explore the effect of the value of *k* and the use of reassigned training data on the SEs obtained from the simulation. Using *k*= 1 increased the SE estimates for mean volume, even to overestimates, especially for small test unit 24 (Table 4). When the reassigned training data was also used SE estimates were comparable to those calculated for the field data in all test units. Using *k*= 3 and the reassigned training data resulted in an underestimation of the SEs. The drawback of the reassigned training data is that it underestimates mean volumes.

Table 4 Mean volume of growing stock and its SE on FRYL for the large test units 2, 3 and 4 and small unit 24: field data-based estimates, corresponding simulation results from MS-NFI8 image data using various estimation parameters (*k* and reassigning of the training data).

Data	Parameters	variable	Unit 2	Unit3	Unit4	Unit 24
field inventory		Mean volume	112.4	113.8	114.7	116.3
		SE m^3/ha	3.5	3.6	3.6	5.1
		SE %	3.1	3.2	3.1	4.4
MS-NFI8	operative	Mean volume	116.5	109.8	124.1	118.5
		SE m^3/ha	2.5	3.0	2.5	3.6
		SE %	2.2	2.7	2.0	3.0
MS-NFI8	*k*= 1	Mean volume	117.2	109.5	128.0	117.8
		SE m^3/ha	4.3	4.0	4.0	7.7
		SE %	3.6	3.6	3.1	6.5
MS-NFI8	*k*= 1,	Mean volume	103.8	91.3	116.4	104.9
	reassigned	SE m^3/ha	3.3	3.5	3.7	4.1
		SE %	3.2	3.8	3.2	3.9
MS-NFI8	*k*= 3,	Mean volume	104.4	92.0	116.0	108.0
	reassigned	SE m^3/ha	3.1	3.0	3.1	2.6
		SE %	2.9	3.3	2.7	2.3

5 Discussion

Due to the sources of error in the field plot and remote sensing data the pixel level prediction errors of MS-NFI are considerable and the range of the predictions and the variance of the predictions are smaller than those of the original variable of the interest (in sampling simulation), e.g., mean volume. However, with a suitable parameter selection (*k*= 1) the original variation of the large area training data (area size of an Landsat Image) can be retained in the predicted maps (Franco-Lopez et al., 2001). Reducing the field plot location error in the training data by reassigning the spectral values (Halme and Tomppo, 2001) decreases the prediction errors of the output maps and helps to preserve the variation of the field plot data. When the simulation results from operative MS-NFI8 and MS-NFI9 were compared it was found that the average variation of forests does not change in a short time like 6 years.

The MS-NFI predictions and prediction errors are spatially correlated. If sampling is such that the geographical distance of observations in successive sample is small compared to the range of the correlation, the variance of the sample means may underestimate the sampling error. Although only 9 samples could be simulated from the MS-NFI thematic maps for the 1 km^2 test units SE estimates comparable to those calculated for the field data were obtained when reassigned training data and *k*= 1 was used.

A priori information of the area to be inventoried is required when planning a sampling design for a forest inventory. This study indicates that the Finnish MS-NFI thematic maps can be used to estimate sampling errors of field inventories for areas of sizes larger than 1 km^2 when using systematic sampling designs with a plot distances down to 75-100 m. One prerequisite for the application of this method is that the pixel level MS-NFI predictions and the original field variables have similar ranges, similar spatial correlation and that the predictions have also a sufficient accuracy.

References

Franco-Lopez H., Ek A.R., and 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, pp. 251-274.

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

Heikkinen J., 2006, Assessment of uncertainty in spatially systematic sampling. In *Forest inventory - methodology and applications. Managing forest ecosystems vol. 10*. Annika Kangas, and Matti Maltamo. (Ed.), pp. 155-176, Dodrecht, The Netherlands: Springer.

Henttonen H., 1991, Vmi8:n Pohjois-Suomen otanta-asetelmien vertailu satelliittikuvatulkinnan avulla. Finnish Forest Research Institute, Manuscript, in Finnish.

Kim H.J., and Tomppo E. 2006. Model-based prediction error uncertainty estimation for k-nn method. *Remote Sens. Environ.* in Press.

Mackisack M.S., and Wood G.B., 1990, Simulating the forest and the point sampling process as an aid in designing forest inventories. *Forest Ecology and Management* , 38, pp. 79-103.

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., Tomppo E., Finley A., and Heikkinen J., 2005, Model-based estimates for k-nearest neighbours technique. Forest Inventory and Analysis, North Central Research Station, USDA Forest Service, Manuscript.

Tokola T., and Shestra S.M., 1999, Comparison of cluster-sampling techniques for forest inventory in southern nepal. *Forest Ecology and Management*, 116, pp. 219-231.

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. Risto Päivinen, Jerry Vanclay, and Saija Miina (Ed.), pp. 27-41,European Forest Institute, Joensuu, Finland.

Tomppo E. 2006. The finnish national forest inventory. In *Forest inventory - methodology and applications. Managing forest ecosystems vol. 10*. Annika Kangas, and Matti Maltamo. (Ed.), pp. 179-194, Dodrecht, The Netherlands: Springer.

Tomppo E., and Halme M., 2004, Using coarse scale forest variables as ancillary information and weighting of variables in k-nn estimation: a genetic algorithm approach. *Remote Sens. Environ.,* 92, pp. 1-20.

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

Tomppo E., Katila M., Moilanen J., Mäkelä H., and Peräsaari J., 1998b, Kunnittaiset metsävaratiedot 1990-94. *Folia For.*, 4B/1998, pp. 619-839. (in Finnish).