Author(s): Minna Räty, Juha Heikkinen, Kari T. Korhonen, Jouni Peräsaari, Antti Ihalainen, Juho Pitkänen & Annika Susanna Kangas

Title: Effect of cluster configuration and auxiliary variables on the efficiency of local pivotal method for national forest inventory

Year: 2019

Version: Published version

Please cite the original version:

Minna Räty, Juha Heikkinen, Kari T. Korhonen, Jouni Peräsaari, Antti Ihalainen, Juho Pitkänen & Annika Susanna Kangas (2019): Effect of cluster configuration and auxiliary variables on the efficiency of local pivotal method for national forest inventory, Scandinavian Journal of Forest Research. https://doi.org/10.1080/02827581.2019.1662938.

Taylor & Francis
Taylor & Francis Group

Scandinavian Journal of Forest Research

ISSN: 0282-7581 (Print) 1651-1891 (Online) Journal homepage: https://www.tandfonline.com/loi/sfor20

Effect of cluster configuration and auxiliary variables on the efficiency of local pivotal method for national forest inventory

Minna Räty, Juha Heikkinen, Kari T. Korhonen, Jouni Peräsaari, Antti Ihalainen, Juho Pitkänen & Annika Susanna Kangas

To cite this article: Minna Räty, Juha Heikkinen, Kari T. Korhonen, Jouni Peräsaari, Antti Ihalainen, Juho Pitkänen & Annika Susanna Kangas (2019): Effect of cluster configuration and auxiliary variables on the efficiency of local pivotal method for national forest inventory, Scandinavian Journal of Forest Research

To link to this article: https://doi.org/10.1080/02827581.2019.1662938

Published online: 16 Sep 2019.

SCANDINAVIAN JOURNAL OF FOREST RESEARCH
https://doi.org/10.1080/02827581.2019.1662938

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Effect of cluster configuration and auxiliary variables on the efficiency of local pivotal method for national forest inventory

Minna Räty [a], Juha Heikkinen [a], Kari T. Korhonen [b], Jouni Peräsaari[a], Antti Ihalainen [a], Juho Pitkänen [b] and Annika Susanna Kangas [b]

[a]Natural Resources Institute Finland (Luke), Helsinki, Finland; [b]Natural Resources Institute Finland (Luke), Joensuu, Finland

ABSTRACT
Planning a forest inventory comprises making decisions related to the sampling strategy: cluster configuration, sample size and sample allocation within the survey area. Cluster configuration includes deciding on the number of sample plots within the cluster and distances between them. Available resources set the limit for field work in terms of man-days. If the time consumption for measurements is known, the sample size can be determined under the constraint. In this study, we simulated the second phase of inventory sampling with fixed time resources by replicating sample selection with a spatially balanced sampling utilizing local pivotal method (LPM) for different cluster configurations to find the most efficient. As a result, the temporary cluster configuration was changed from 9 to 5-sample plot configuration in a pilot inventory. Further, the sample selection was performed with LPM having total growing stock volume and broadleaf volume proportion as auxiliary information. The pilot results were aligned with the time series in respect to forest area and total growing stock volume, but in tree species groups deviations were observed in growing stock volume. A more comprehensive optimization should include the travelling routes, the plot-to-plot distances and the plot design. In any case, the result is region specific.

ARTICLE HISTORY
Received 19 July 2018
Accepted 23 August 2019

KEYWORDS
Forest resources; limited inventory resources; optimization; relative efficiency; sampling

Glossary

Cluster size	= number of sample plots in a cluster
Cluster configuration	= cluster size + the actual chosen sample plots which determine the dimensions of a cluster and within cluster plot-to-plot distances
Auxiliary information	= MS-NFI10 forest resource maps
Auxiliary variable	= from auxiliary information derived cluster level variable
Sampling strategy	= unique combination of set of auxiliary variables and cluster configuration (and window size used when auxiliary information was gathered)
Forested land	= forest land and poorly productive forest land

Introduction

National forest inventories (NFI) collect information on variables related to the forest state and health, mean growing stock volume and forest area being the most commonly reported variables (Tomppo et al. 2010). Lately, most attempts to enhance NFI with the help of remote-sensing data have concentrated on the estimation phase (McRoberts et al. 2010). Methods like post-stratification (McRoberts et al. 2012; Tipton et al. 2013; Magnussen et al. 2015; Myllymäki et al. 2017) and model-assisted estimation (Särndal et al. 1992; Opsomer et al. 2007; Gregoire et al. 2011; McRoberts et al. 2013; Tipton et al. 2013; Saarela et al. 2015; Kangas et al. 2016) have been widely studied in connection to forest inventory. However, spatially balanced sampling has recently been demonstrated to offer great potential to enhance the design phase (Grafström et al. 2017; Räty et al. 2018). Samples that are well spread in the space generated by any given set of auxiliary variables can be obtained with the local pivotal method (LPM, Grafström et al. 2012).

Operational NFIs often apply systematic cluster sampling, where clusters of sample plots are spread out over the inventory region, for example, in the form of a square grid (Tomppo et al. 2010). NFI budget defines how many man-days each year can be allocated to the field work. The limit to the number of clusters is thus determined as a function of the cluster configuration (number and locations of plots in a cluster) and the resulting time needed to measure one cluster.

LPM can be implemented to a cluster-based operational inventory through a two-phase design (Grafström et al. 2017): A dense grid of clusters is first overlaid to the inventory region. Values of auxiliary variables are then obtained for each cluster of this first-phase sample. The final second-phase sample to be measured is then a subset of the first-phase sample such that the distribution of the auxiliary variables is matched as well as possible to that of the first-phase sample. In other words, LPM sample is drawn from the finite population of first-phase clusters.

Motivated by the promising results obtained in earlier simulation studies (Grafström et al. 2017; Räty et al. 2018), Finnish NFI decided to implement a pilot project where LPM is tested as a part of the operational inventory. The number of second-phase clusters to sample was limited by the resources, but choices had to be made concerning the

CONTACT Minna Räty minna.raty@luke.fi Natural Resources Institute Finland (Luke), PO Box 2, FI-00791 Helsinki, Finland

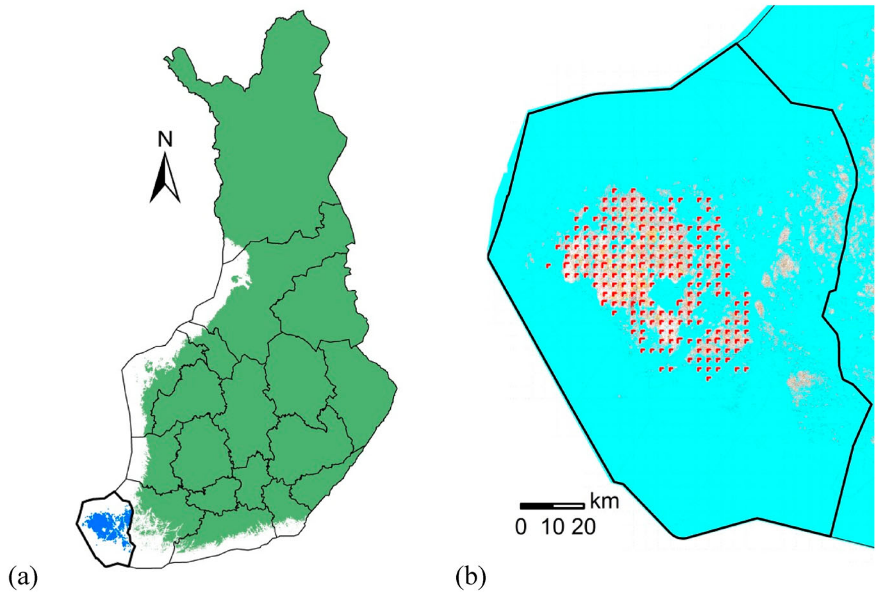

Figure 1. (a) The Åland region in the southwestern Finland (in blue) and (b) the locations of the subset of NFI11 sample plots (red dots) within the region used in the simulation study.

cluster configuration and the auxiliary data. In case of wall-to-wall multilayer raster map of auxiliary data (in our case, thematic maps of forest resources produced as a result of the previous inventory), the latter includes (1) selection of map layers used as auxiliary variables, (2) selection of pixels from which the cluster-level auxiliary variable is computed and (3) selection of cluster-level parameters computed from pixel-level values. The second question is necessary because the auxiliary information used does not fully correspond to one fixed location in the field due to georeferencing errors and influence of atmospheric conditions and geometrical correction of satellite image that enable the sample plot to spread over more than one pixel (Katila 2004; Tomppo et al. 2008).

The aim of this study was to analyse the effects of the choice of cluster configuration and auxiliary variables on the efficiency of the resulting LPM sampling design. The analysis was based on sampling simulation, where the role of the first-phase sample was played by real field data from the previous inventory and auxiliary variables were derived from an earlier multi-source inventory forest resources map. Such an approach was made possible by the fact that extra resources available during the previous inventory of the study region had led to exceptionally intensive sampling. The most applicable sampling strategy from the above analysis was applied in a pilot inventory in summer 2018. In this paper, we present the selected sampling strategy and preliminary inventory results.

Material and methods

Different LPM designs were compared by simulating subsamples from the whole set of field sample plots of the 11th National Forest Inventory (NFI11) within the 1227 km^2 study region located in the Southwestern archipelago of Finland in Åland region and inventoried in the summer of 2013 (Figure 1). The target variables were the proportion of forested land and the growing stock volumes on forested land by tree species groups (Table 1). "Forested land" in this study refers to a combination of two forestry land classes in the national system, i.e. classes "forest land" and "poorly productive forest land" (Tomppo et al. 2011), which is close to the United Nations Food and Agriculture Organisation (FAO) definition of forest (FAO 2012).

The auxiliary information was derived from the results of the 10th Multi-Source National Forest Inventory (MS-NFI10),which was based on field data of the 10th NFI of Finland (Korhonen et al. 2017) and Landsat 5 TM images, both from year 2007, as well as digital map data used for delineating other land uses from forestry land (Tomppo et al. 2008; Tomppo et al. 2012). The forest resource maps of MS-NFI10 were available with 20 m × 20 m pixel size. Six map layers were utilized: (1) total growing stock volume, (2) pine volume including all other conifers except spruce, (3) spruce volume, (4) birch volume, (5) other broadleaf growing stock volumes and (6) land use class.

Two different auxiliary data sets were subsampled from each thematic map. The first set contained only those pixels within which a sample plot centres was located. The second set was composed of 5-pixel windows around each plot also including pixels adjacent to the plot centre pixels in the cardinal directions, i.e. so-called Rook's case contiguity (e.g. Lloyd 2009). The cluster level auxiliary variables derived from the selected pixels included means, total values, variances and ratios (Table 2; Appendix Tables 1–2). Different combinations

Table 1. Proportion of forest land, mean and total volume of growing stock by tree species group and combined according to NFI11 in simulation study material.

		Mean volumes			
Forested land	Total volume	Pine	Spruce	BL	All
62.9%.	9.5×10^6 m^3	72.2 m^3 ha^{-1}	25.5 m^3 ha^{-1}	25.1 m^3 ha^{-1}	122.7 m^3 ha^{-1}

Table 2. Estimation of auxiliary variables from the forest resource MS-NFI map layers in two different ways: estimating mean volumes and total volumes. Note that the land class proportion variables (For) and total growing stock variance (Var) are the same for both auxiliary variables.

Auxiliary variable	Forest resource map	Mean estimation	Total estimation
Vol	Total growing stock volume of all tree species	Mean[1]	Total[1]
Pi	Pine growing stock volume[2]	Mean[1]	Total[1]
Spr	Spruce growing stock volume	Mean[1]	Total[1]
BL	Birch and other broadleaf growing stock volume	Mean[1]	Total[1]
BLprop	BL_{Mean}/Vol_{Mean} or BL_{Total}/Vol_{Total}	Proportion	
Var	Total growing stock volume	Variance[1]	
For	Land class	Proportion of forested land[3]	

[1] Mean, total and variance are calculated over forested land pixels within plots included in the cluster.
[2] Includes all conifer species except Norway spruce.
[3] Forested land consists of forestry land and shrub land classes.

Table 3. Tested auxiliary variable combinations in sampling simulations with both mean volume and total volume based variables. X and Y are geographic coordinates in Finland uniform coordinate system.

	Vol	Pi	Spr	BL	Var	For	BLprop	X	Y
1	X					X		X	X
2	X				X	X		X	X
3	X	X				X		X	X
4	X		X			X		X	X
5	X			X		X		X	X
6		X	X	X		X		X	X
7	X	X	X	X	X			X	X
8	X	X	X	X		X		X	X
9	X	X	X	X	X	X		X	X
10								X	X
11	X						X	X	X
12	X					X	X	X	X

of these variables and geographic coordinates in the Finnish uniform coordinate system (epsg: 2393) were tested (Table 3). The set of combinations was based on previous experience and simulations(Räty et al. 2018; Räty and Kangas 2019).

Cluster configurations compared in this study were subsets of the NFI11 clusters of nine sample plots with 200 m intervals (Figure 2 and Figure 3). Only those clusters were sampled that contained at least one plot at least partially on land according to the digital map data (Table 4).

The costs for each cluster configuration were defined as a sum of three time consumption components that were needed to reach the sample plot: (1) the travelling time on land to reach the cluster centre (30 min for each cluster), (2) travelling with boat to the cluster centre (0–100 min) and (3) time to reach the sample plot from the cluster centre (0–40 min). Obviously, boat cost above equalled to 0 if the cluster was reachable without a boat, i.e. located on one of the main islands connected with bridges and road network. Additionally, if the sample plot located in water the last cost component was also 0 because it would have not been visited in the field survey. The cost did not include an estimate of the actual field measurement time demanded, just the reaching time. All the field sample plots were assumed to be measured following the same principles and procedures and thus their time consumption would have been the same regardless of the cluster configuration.

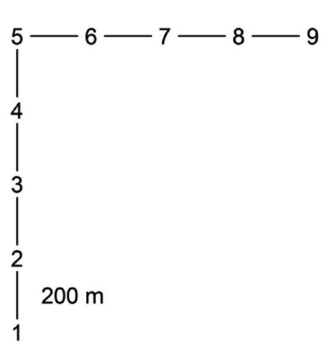

Figure 2. Temporary cluster configuration used in the 11th National Forest Inventory in Åland.

Subsampling of NFI11 field plots under each sampling strategy obtained as some combination of the two auxiliary data sets, 23 sets of auxiliary variables and 10 cluster configurations was replicated $T = 5000$ times. Number of sampled clusters varied between replications, most remarkably between cluster configurations, as a result of limiting the anticipated cost to 7000 ± 70 min.

The sample selection was based at LPM, which aims at a sample as close as possible to its distribution in population by excluding the most similar unit when promoting the other (Grafström et al. 2012). While the selection process is based on updating the inclusion probabilities of the clusters iteratively, the actual inclusion probabilities (both first and second order) are fixed.

Estimates $\hat{y}$ for population parameters k in sampling strategy j were derived using standard NFI estimators (Tomppo et al. 2011, Ch. 3) from each replication, which allowed us to estimate a mean squared error (MSE). MSE presents the variation around the true NFI11 population mean value Y_k

$$MSE_{k,j} = \frac{\sum_{i=1}^{T} (\hat{y}_{k,j} - Y_k)^2}{T}, \quad (1)$$

where k, $k = 1, \ldots, 6$, is the estimated population parameter, T is the number of replications, and j, $j = 1, \ldots, 360$ is the combination of the set of auxiliary variables (23), cluster configuration (10) and pixel window size (2) combination. The smaller the MSE the more accurately the sampling strategy was capable to estimate the population parameter. Efficiencies of the strategies were compared with relative measure:

$$RE_{k,j} = \frac{MSE_{k,ref}}{MSE_{k,j}}, \quad (2)$$

where $MSE_{k,ref}$ is the MSE of the reference sampling strategy: one-pixel windows when comparing auxiliary data sets, coordinates only when comparing auxiliary variable combinations, and nine-plot cluster when comparing cluster configurations.

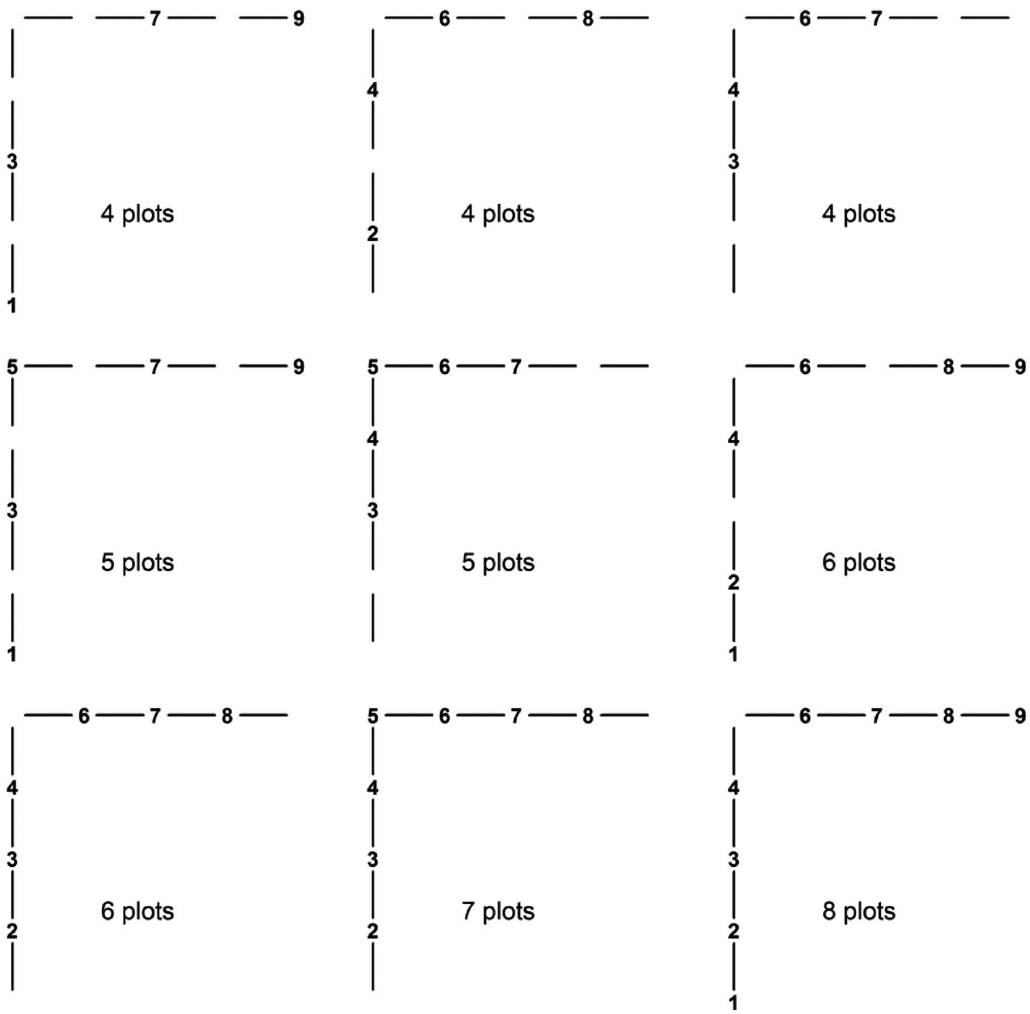

Figure 3. Different cluster configurations tested in the simulations. The numbered sample plots were included to the designs. Also the original 9 sample plot configuration was included.

Table 4. Sampling population sizes as total number of clusters and number of sample plots in land for different cluster configurations. Sampling population was limited to the clusters where at least one sample plot would have been visited in the field measurements.

Cluster size	Sample plots	Number clusters	Land plots
4	1, 3, 7, 9	181	494
4	2, 4, 6, 8	180	494
4	3, 4, 6, 7	171	497
5	1, 3, 5, 7, 9	192	619
5	3–7	175	622
6	1, 2, 4, 6, 8, 9	196	745
6	2–4, 6–8	188	742
7	2–8	189	863
8	1–4, 6–9	201	989
9	1–9	203	1113

From the pilot survey, the inventory results were estimated for the area and growing stock volumes. These were compared to the corresponding results of the same region in the previous NFIs.

Results

Number of clusters that can be measured within the given time limit varied from 66 (using the nine-plot clusters) to 99 (using the four-plot clusters, Table 5). With larger clusters, we obtained larger numbers of plots, but they are also more correlated due to stronger clustering. The five-pixel window was slightly but consistently better than one-pixel window (Table 6) and therefore the rest of the results are based on auxiliary data set estimated from a five-pixel window.

Utilizing of auxiliary information did enhance the sampling efficiency with most of the simulated auxiliary variable combinations and with total volume based auxiliary variables in all

Table 5. Average sample sizes (= number of chosen temporary clusters) and land and forested land sample plots in the samples with the predefined total cost of 7000 ± 70 min for different cluster sizes and configurations.

		Sampled clusters			
Cluster size	Sample plots	Number	% of population	Land plots	Forested land plots
4	1, 3, 7, 9	98	54	266	172
4	2, 4, 6, 8	99	55	271	165
4	3, 4, 6, 7	97	57	282	174
5	1, 3, 5, 7, 9	89	46	286	185
5	3–7	87	50	308	192
6	1, 2, 4, 6, 8, 9	82	42	309	193
6	2–4, 6–8	81	43	318	195
7	2–8	74	39	337	210
8	1–4, 6–9	70	35	343	215
9	1–9	66	33	357	225

Table 6. The ratio of mean squared errors when utilizing only the centre pixel auxiliary information against the 5-pixel-window around the sample plot centre to the estimation.

	Forested land	Total volume	Mean volumes			
			Pine	Spruce	BL	All
Min	0.80	0.81	0.74	0.67	0.73	0.74
1stQ	1.03	1.01	0.95	0.93	0.94	0.95
Median	1.10	1.07	1.01	1.02	0.99	1.02
Mean	1.13	1.10	1.02	1.04	1.00	1.04
3rdQ	1.19	1.15	1.09	1.11	1.04	1.09
Max	1.93	1.79	1.60	1.79	1.55	1.63

cases (Table 7). The rest of REs were within the range of ±0.3 for both the mean and total types of auxiliary variables. In cluster configurations, the sparsest four-plot configuration was clearly less efficient than the original nine-plot configuration, and many configurations were close to the original design. The densest cluster configurations with 5 and 6 sample plots provided highest REs (Table 8). Thus one could expect to gain on average a RE of 1.0–1.5 over the population parameters in respect to the original 9-sample plot configuration due to higher number of clusters. Also one of the 4-sample plot configurations had on average as high relative efficiency, but it resulted from very efficient growing stock estimation on expense of forested land proportion estimation which was regarded as not acceptable.

To favour simplicity, a 5-sample plot cluster configuration with plots number 3–7 in the original cluster configuration (Figure 2) and LPM utilizing only two auxiliary variables besides geographic coordinates, namely total growing stock volume and proportion of broadleaved volume of total growing stock volume in the sample selection, were chosen for the pilot of operational NFI.

We utilized the most recent multi-source forest inventory result maps dated at year 2015 (Mäkisara et al. 2019). For the sample selection, a primary cluster grid at 96 m intervals from a centre of cluster corner sample plot to the next cluster's corner was span over Åland. In the newest MS-NFI map, the pixel size was 16 × 16 m and therefore we collected four pixels corresponding to a square of 32 × 32 m from location of each sample plot, the cluster-level auxiliary variables were estimated from pixels, and a sample of 175 cluster was selected (Figure 4).

The clusters were measured during summer 2018 as a part of the 12th NFI (years 2014–2018). According to primary analysis, no changes have happened in land use class area estimates. The estimate of total volume of growing stock has increased to the level of the NFI10 (2007 in Åland) after a clear decrease in NFI11 (Figure 5). Both the area and volume estimates are credible as compared to the previous inventories and expected variance of the estimates.

Discussion

The aim of this study was to find an optimal cluster configuration for the southwestern archipelago in Finland, and a method to select a sample, and then finally apply the derived sampling design in selection of temporary clusters for the 12th NFI in the Åland region (Figure 1). The

Table 7. Average relative efficiencies for auxiliary variable combinations for estimated population parameters and average over all of them. Note that the geographic x and y coordinates were always included to the auxiliary variables. On the left are the results for auxiliary variables estimated using mean volumes and on the right with total volumes.

Auxiliary variables based on mean volumes													Auxiliary variables based on total volumes						
Auxiliary variables								Mean volumes							Mean volumes				
Vol	Pi	Spr	BL	Var	For	Forested land	Total volume	Pi	Spr	BL	All	Mean RE	Forested land	Total volume	Pi	Spr	BL	All	Mean RE
X					X	1.25	1.31	1.06	1.41	1.09	1.29	1.24	1.29	1.36	1.10	1.31	1.09	1.29	1.24
X				X	X	1.00	0.99	0.97	0.99	1.00	0.94	0.98	1.25	1.28	1.08	1.26	1.08	1.21	1.19
X	X				X	1.31	1.32	1.06	1.39	1.21	1.25	1.26	1.34	1.38	1.13	1.28	1.16	1.27	1.26
X		X			X	1.33	1.37	1.16	1.28	1.18	1.28	1.27	1.28	1.36	1.07	1.43	1.11	1.32	1.26
X			X		X	0.91	0.95	0.96	0.98	0.95	0.92	0.95	1.35	1.34	1.09	1.35	1.23	1.26	1.27
	X	X	X		X	1.30	1.32	1.08	1.41	1.21	1.27	1.27	1.33	1.33	1.08	1.39	1.25	1.30	1.28
X	X	X	X	X		1.55	1.41	1.06	1.30	1.11	1.24	1.28	1.18	1.21	1.10	1.30	1.12	1.15	1.18
X	X	X	X		X	1.64	1.48	1.11	1.38	1.17	1.28	1.34	1.35	1.34	1.09	1.34	1.24	1.29	1.28
X	X	X	X	X	X	1.23	1.24	1.10	1.32	1.13	1.16	1.20	1.38	1.32	1.14	1.37	1.17	1.25	1.27
X			Xprop[1]			1.33	1.34	1.10	1.35	1.10	1.30	1.25	1.13	1.15	1.12	1.24	1.21	1.16	1.17
X			Xprop[1]		X	1.39	1.33	1.09	1.36	1.20	1.26	1.27	1.42	1.38	1.11	1.42	1.20	1.33	1.31
						1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

[1]Auxiliary variable is the proportion of broadleaf growing stock volume of the total growing stock volume, BL/Vol.

Table 8. Average relative efficiencies of cluster configurations for estimated population parameters over auxiliary variable combinations excluding the sampling design with only geographic coordinates. On the left are the results for auxiliary variables estimated using mean volumes and on the right with total volumes.

		Auxiliary variables based on mean volumes							Auxiliary variables based on total volumes						
Cluster design		Forested	Total	Mean volumes				Mean	Forested	Total	Mean volumes				Mean
Size	Sample plots	land	volume	Pi	Spr	BL	All	RE	land	volume	Pi	Spr	BL	All	RE
4	1, 3, 7, 9	1.07	0.45	0.54	0.63	0.99	0.49	0.70	1.08	0.44	0.51	0.62	0.99	0.47	0.69
4	2, 4, 6, 8	0.82	0.78	1.33	1.72	1.54	1.29	1.25	0.86	0.79	1.37	1.69	1.42	1.29	1.24
4	3, 4, 6, 7	1.25	1.43	0.54	1.36	0.47	1.07	1.02	1.28	1.52	0.53	1.42	0.45	1.10	1.05
5	1, 3, 5, 7, 9	0.89	0.77	1.03	0.88	1.07	0.97	0.94	0.90	0.74	0.99	0.92	1.03	0.94	0.92
5	3–7	1.43	1.24	1.17	1.27	0.94	1.05	1.18	1.51	1.33	1.13	1.23	0.94	1.03	1.20
6	1, 2, 4, 6, 8, 9	1.23	1.16	1.10	1.37	1.24	1.05	1.19	1.28	1.14	1.08	1.33	1.19	1.01	1.17
6	2–4, 6–8	0.85	1.00	0.88	1.13	0.83	1.05	0.96	0.87	1.03	0.87	1.12	0.83	1.09	0.97
7	2–8	1.02	0.92	0.91	1.19	1.03	0.97	1.01	1.09	0.98	0.88	1.22	1.05	0.97	1.03
8	1–4, 6–9	1.05	0.96	0.82	1.05	0.92	0.85	0.94	1.03	0.96	0.82	1.00	0.94	0.83	0.93
9	1–9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

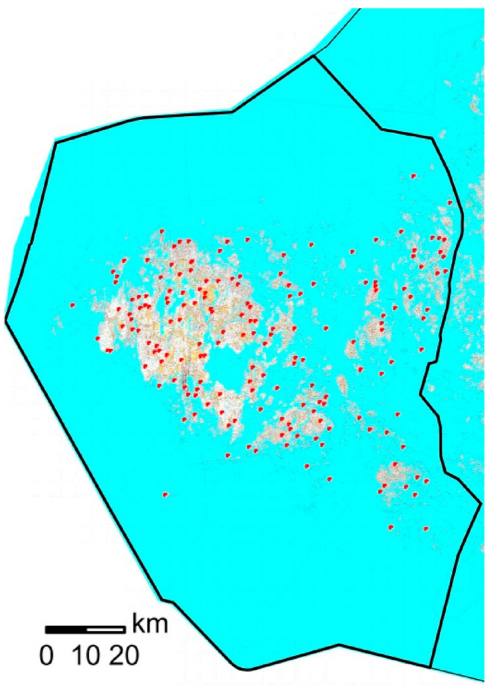

Figure 4. Temporary sample clusters (red) in the 12th Finnish National Forest Inventory in southwestern archipelago.

optimization of cluster configuration was conducted by simulating the second stage of two-stage sampling with spatially balanced sampling method utilizing local pivotal method (LPM) in sample selection. The number of sampled clusters was limited with time consumption cost to reach the plots, i.e. fixed inventory resources. This optimization was the novel part of this study. As far as we know, no one has applied LPM in a similar set-up before us. Thus the fundamental question was whether it could be connected to the decision making concerning the overall sampling strategy.

As we initially assumed, instead of only one single pixel within which the sample plot centre fell in, it was better to use the larger 5-pixel window in the auxiliary information extraction (Table 6). Apparently, the cluster level auxiliary variables which were totals, means or variances of the single pixel values of the MS-NFI forest resource maps benefitted from having larger pixel window from the positions of sample plots within the cluster. Further studies on the cluster level pixel selection could be carried on by enlarging the window or weighting the pixels according to the assumed proportion of sample plot that belongs to that pixel (Grafström et al. 2017).

In the auxiliary variables, we had both growing stock volume mean and total based and land use related variables (Table 7). In general, it was a beneficial to use auxiliary variables in sampling but between the set of used auxiliaries the differences were not large. Therefore even really simple sampling designs with few auxiliary variables would come into question. The auxiliary variables are the crucial part of the method since they are supposed to have explanatory power over the inventory variables and the distances are calculated between them. In fact, the auxiliary variables and inventory variables were linearly correlated (Appendix Table 1), which indicate the explanatory power of the chosen auxiliary variables. Compared to the previous study (Räty et al. 2018), we introduced in this work also the total value based auxiliary variables. Thus we had sum, mean and variance statistics at the cluster level. Further investigations on describing and comparing the clusters with the auxiliary variables could be tried out, for example, by introducing other descriptive statistics like quantiles.

We had reference field data only from the locations of the previous inventory's sample plot locations, and therefore we were only able to alter the set of sample plots that existed in the original cluster configuration. Thus we could neither simulate any larger clusters than the original cluster (Figure 2 and Figure 3) nor include the plot-to-plot distance optimization to the study (except for the comparison between 200 m distance to 400 m distance in plot types 1-3-5-7-9 and 3-4-5-6-7, Figure 3). The latter could have been used to evaluate spatial autocorrelation (see e.g. Schabenberger and Gotway 2004). In cluster sampling, to gain the most out of an inventory the sample plots are placed at distance where spatial autocorrelation between the plots vanishes (Tomppo et al. 2001).

Our time constraint was based on the visual interpretation of an expert. Those can be considered to be as realistic as possible taking into account that the resulting sample is not known in advance and hence the time consumption was defined for a generic case without travelling route optimization. Our time cost constraint included only the travelling time, but in more realistic estimation also the time to measure the sample plots would have been included. Typically the measurement time of a plot in simulation studies is

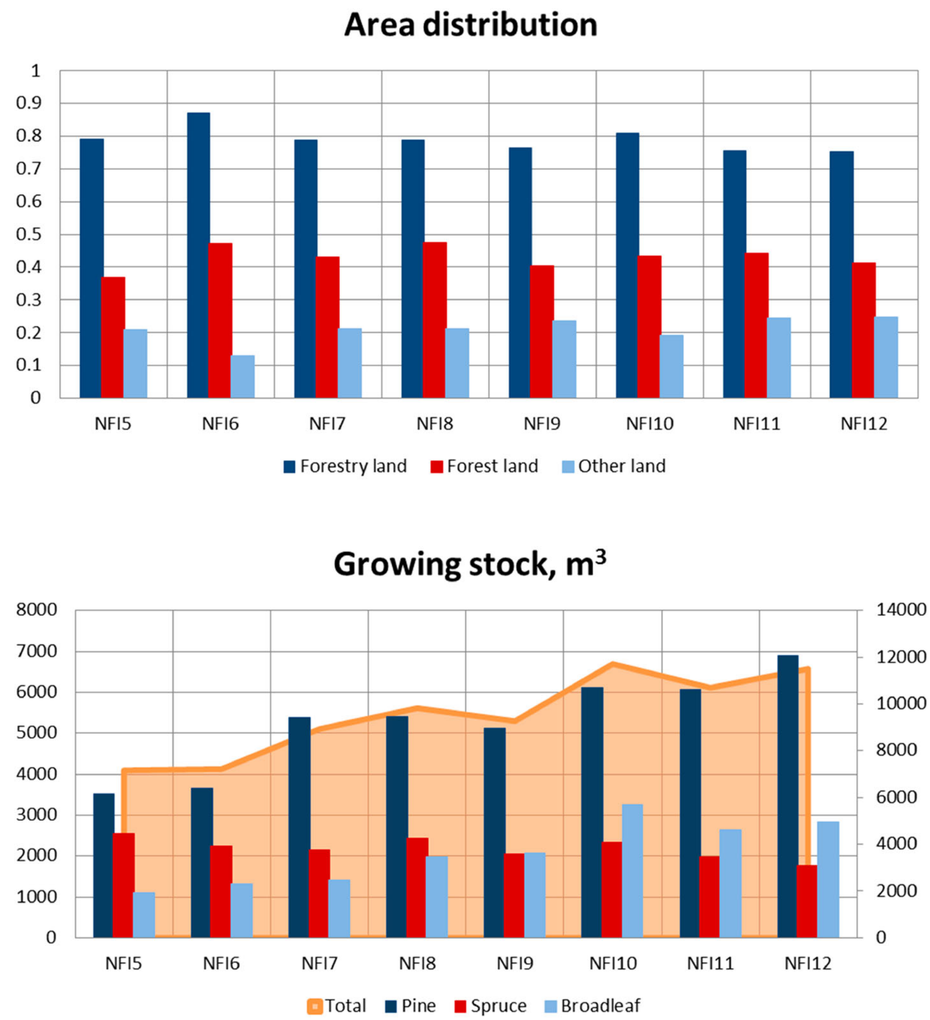

Figure 5. National forest inventory result time series of land area distribution (above) and growing stock volume (below) in Åland including the 12th inventory carried out with spatially balanced sampling for the temporary clusters.

an average value used for all plots, as more detailed information is not available, and therefore that would have only a slight impact on the results in our case. Introducing the route optimization (e.g. Yim et al. 2015) would have made the study more realistic but also more tedious as we compared 360 sampling designs using 5000 repetitions for each.

For a comprehensive sampling planning (e.g. Tomppo et al. 2014) also the considerations of plot configuration should have been included (e.g. Henttonen and Kangas 2015). Including the plot type optimization into the problem, however, would require mapped data of trees in the simulation. Connecting above-mentioned components into one and same optimization would actually allow to adjust the sampling strategy locally, in an optimal case seamlessly but at least for smaller regions than is currently the case (Figure 1 in Korhonen et al. 2017), but the adjustment should not depend on the measurements done in a plot. An optional approach would be to base the decision of sampling design on Cost + Loss analysis (Gilabert and McDill 2010; Barth and Ståhl 2012).

The result including both the cluster configuration and auxiliary variable combination in sample selection holds only for the study region. For example, spatial variation differs in regions (Ranneby et al. 1987; Tomppo et al. 2001), which impacts on the within cluster distances, cluster size and sampling intensity. Thus we expect that for other regions similar analysis should be carried out to find the best sampling strategy for those other conditions. Particularly in our case, the study region we had is exceptional compared to the other regions in Finland. On the other hand, this opens up the chance to use different datasets in different regions that are locally relevant and perhaps limitedly available such as laser-scanning data.

Based on the previous results on the cluster configuration (Table 8), it was clear that either 5 or 6 sample plot configuration would be the most efficient. Additionally, the auxiliary variables would most probably be based on the total growing stock volumes than means as the former proved to be clearly more efficient (Table 7). This might have been due to the way how mean volumes were estimated just over the forested land: in a scattered landscape as in the archipelago where the most remarkable feature is the variation in the size of forest patches. Most of the clusters included only few forested sample plots and with mean value auxiliaries there is no link to the actual amount of wood. With the total growing stock volume based auxiliary variables, the samples describe more efficiently the variation in the forested area and therefore also in the amount of wood. In the mainland, where the forest areas are more contiguous, the differences may not be as large. So, we eliminated all mean value based auxiliary variable combinations from our considerations.

In conclusion, the simulations utilizing auxiliary remote-sensed data and spatially balanced sampling with local pivotal method could make a distinction between different cluster configurations. If the target is to get equally good estimates for all tested population parameters, the cluster size could be smaller than in previous survey – 5–6 sample plots. Further, when estimating the auxiliary variables from auxiliary information (forest resource maps) a bigger window around the sample plots is preferable and the auxiliary variables should be based rather on total than mean growing stock volume. The best sampling design from practical point of view was applied over our survey area and the results showed similar trends to previous inventories. The resulting sampling strategy is a region specific and thus similar analysis should be carried out if applied elsewhere.

Acknowledgements

We thank Dr. Helena Henttonen for her comments and suggestion on the manuscript draft.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Ministry of Agriculture and Forestry Finland key project "Puuta liikkeelle ja uusia tuotteita metsästä" ("Wood on the move and new products from forest") Maa- ja metsätalousministeriö.

ORCID

Minna Räty http://orcid.org/0000-0001-9898-8712
Juha Heikkinen http://orcid.org/0000-0003-3527-774X
Kari T. Korhonen http://orcid.org/0000-0002-6198-853X
Antti Ihalainen http://orcid.org/0000-0003-0461-6649
Juho Pitkänen http://orcid.org/0000-0002-7583-6297
Annika Susanna Kangas http://orcid.org/0000-0002-8637-5668

References

Barth A, Ståhl G. 2012. Determining sample size in national forest inventories by cost-plus-loss analysis: an exploratory case study. Eur J For Res [Internet]. [Cited 2018 Jul 12]. 131:339–346. Available from: https://link.springer.com/content/pdf/10.1007%2Fs10342-011-0505-5.pdf.

FAO. 2012. Forest Resources Assessment 2015: Terms and Definitions. FAO Rep [Internet]. [cited 2019 Feb 1]. Available from: http://www.fao.org/docrep/017/ap862e/ap862e00.pdf.

Gilabert H, McDill ME. 2010. Optimizing inventory and yield data collection for forest management planning. For Sci [Internet]. [Ccited 2018 Jun 6]. 56:578–591. Available from: https://academic.oup.com/forestscience/article/56/6/578/4604486?searchresult=1.

Grafström A, Lundström NLP, Schelin L. 2012. Spatially balanced sampling through the pivotal method. Biometrics. 68:514–520.

Grafström A, Zhao X, Nylander M, Petersson H. 2017. A new sampling strategy for forest inventories applied to the temporary clusters of the Swedish NFI. Can J For Res. 47:1161–1167.

Gregoire TG, Ran G, Hl S, Naesset E, Gobakken T, Nelson R, Ren Holm S. 2011. Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. Can J For Res. 41:83–95.

Henttonen HM, Kangas A. 2015. Optimal plot design in a multipurpose forest inventory. For Ecosyst. 2:31.

Kangas A, Myllymäki M, Gobakken T, Naesset E. 2016. Model-assisted forest inventory with parametric, semiparametric, and nonparametric models. Can J For Res [Internet]. [Cited 2017 May 17]. 46:855–868. Available from: http://www.nrcresearchpress.com/doi/pdf/10.1139/cjfr-2015-0504.

Katila M. 2004. Controlling the estimation errors in the Finnish multisource National forest inventory [Internet]. Vantaa: Finnish Forest Research Institute, Research Papers 910. Available from: http://ethesis.helsinki.fi/julkaisut/maa/mvaro/vk/katila/controll.pdf.

Korhonen KT, Ihalainen A, Ahola A, Heikkinen J, Henttonen HM, Hotanen J-P, Nevalainen S, Pitkänen J, Strandström M, Viiri H. 2017. Suomen metsät 2009–2013 ja niiden kehitys 1921–2013 [Internet]. Helsinki. Available from: http://urn.fi/URN:ISBN:978-952-326-467-0.

Lloyd C. 2009. Spatial data analysis - An introduction for GIS users. [place unknown]: Oxford University Press.

Magnussen S, Andersen HE, Mundhenk P. 2015. A second look at endogenous poststratification. For Sci. 61(4):624–634.

Mäkisara K, Katila M, Peräsaari J. 2019. Multi-Source National Forest Inventory of Finland – methods and results 2015. Natural resources and bioeconomy studies 8/2019 [Internet]. Helsinki. Available from: http://urn.fi/URN:ISBN:978-952-326-711-4.

McRoberts RE, Gobakken T, Næsset E. 2012. Post-stratified estimation of forest area and growing stock volume using lidar-based stratifications. Remote Sens Environ [Internet]. [Cited 2017 Aug 4]. 125:157–166. Available from: http://linkinghub.elsevier.com/retrieve/pii/S003442571200274X.

McRoberts RE, Næsset E, Gobakken T. 2013. Inference for lidar-assisted estimation of forest growing stock volume. Remote Sens Environ [Internet]. [Cited 2017 Aug 4]. 128:268–275. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0034425712003896.

McRoberts RE, Tomppo EO, Naesset E. 2010. Advances and emerging issues in national forest inventories. Scand J For Res. 25:368–381.

Myllymäki M, Gobakken T, Naesset E, Kangas A. 2017. The efficiency of poststratification compared with model-assisted estimation. Can J For Res. 47:515–526.

Opsomer JD, Jay Breidt F, Moisen GG, Kauermann G. 2007. Model-assisted estimation of forest resources with generalized additive models. J Am Stat Assoc. 102:400–409.

Ranneby B, Cruse T, Hagglund B, Jonasson H, Sward J. 1987. Designing a new national forest survey for Sweden. Stud For Suec [Internet]. [Cited 2018 Jun 7]. 1–29. Available from: https://pub.epsilon.slu.se/4634/1/SFS177.pdf.

Räty M, Heikkinen J, Kangas AS. 2018. Assessment of sampling strategies utilizing auxiliary information in large-scale forest inventory. Can J For Res. 48:749–757.

Räty M, Kangas AS. 2019. Effect of permanent plots on the relative efficiency of spatially balanced sampling in a national forest inventory. Ann For Sci. 76:20.

Saarela S, Grafström A, Ståhl G, Kangas A, Holopainen M, Tuominen S, Nordkvist K, Hyyppä J. 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. Remote Sens Environ [Internet]. [Cited 2017 May 17]. 158:431–440. Available from: http://ac.els-cdn.com/S0034425714004726/1-s2.0-S0034425714004726-main.pdf?_tid=ca48bf16-3ae4-11e7-b2e5-00000aab0f6c&acdnat=1495014177_ab0233683e19d3f337a48c444dde70a3.

Särndal C-E, Swensson B, Wretman J. 1992. Model assisted survey sampling. New York, NY: Springer-Verlag Publishing.

Schabenberger O, Gotway CA. 2004. Statistical methods for spatial data analysis, 1st Ed. New York: Chapman & Hall/CRC Press.

Tipton J, Opsomer J, Moisen G. 2013. Properties of endogenous post-stratified estimation using remote sensing data. Remote Sens Environ. 139:130–137.

Tomppo E, Gschwantner T, Lawrence M, McRoberts RE, editors. 2010. National forest inventories: Pathways for common reporting. Springer.

Tomppo E, Haakana M, Katila M, Peräsaari J. 2008. Multi-Source National Forest Inventory - Methods and applications. Series: Managing Forest Ecosystems 18 [Internet]. [place unknown]: Springer. Available from: http://link.springer.com/10.1007/978-1-4020-8713-4.

Tomppo E, Heikkinen J, Henttonen HM, Ihalainen A, Katila M, Mäkelä H, Tuomainen T, Vainikainen N. 2011. Designing and Conducting a forest inventory - case: 9th National forest inventory of Finland. [place unknown]: Springer Netherlands.

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. Metsätieteen Aikakausk [Internet]. [cited 2018 Jun 7]:99–248. Available from: https://metsatieteenaikakauskirja.fi/pdf/article6434.pdf.

Tomppo E, Katila M, Mäkisara K, Peräsaari J. 2012. The multi-source National forest inventory of Finland –methods and results 2007 [online]. Work Pap Finnish For Res Inst [Internet]. 233. Available from: http://www.metla.fi/julkaisut/workingpapers/2012/mwp227.htm.

Tomppo E, Malimbwi R, Katila M, Mäkisara K, Henttonen HM, Chamuya N, Zahabu E, Otieno J. 2014. A sampling design for a large area forest inventory: case Tanzania. Can J For Res. 44:931–948.

Yim J-S, Shin M-Y, Son Y, Kleinn C. 2015. Cluster plot optimization for a large area forest resource inventory in Korea. Forest Sci Technol [Internet]. [Cited 2018 Jun 6]. 11:139–146. Available from: http://www.tandfonline.com/action/journalInformation?journalCode=tfst20.

Appendices

Appendix Table 1. Descriptive statistics of field survey data (NFI) and auxiliary variables and correlations between datasets for different population parameters in cluster configurations.

	Configuration		NFI						AUX						
	Size	Plots	Min	1Q	Med	Mean	3Q	Max	Min	1Q	Med	Mean	3Q	Max	Cor
Forest proportion	4	1,3,7,9	0.25	0.25	0.50	0.54	0.75	1	0.05	0.25	0.50	0.51	0.75	1	0.83
	4	2,4,6,8	0.25	0.25	0.50	0.60	0.75	1	0.05	0.25	0.50	0.53	0.75	1	0.84
	4	3,4,6,7	0.25	0.50	0.50	0.61	0.75	1	0.05	0.30	0.55	0.56	0.80	1	0.81
	5	1,3,5,7,9	0.20	0.20	0.60	0.52	0.80	1	0.04	0.20	0.48	0.49	0.72	1	0.85
	5	3–7	0.20	0.40	0.60	0.59	0.80	1	0.04	0.28	0.60	0.55	0.80	1	0.83
	6	1,2,4,6,8,9	0.17	0.33	0.50	0.52	0.67	1	0.03	0.20	0.47	0.47	0.70	1	0.88
	6	2-4,6-8	0.17	0.33	0.50	0.54	0.83	1	0.03	0.23	0.50	0.50	0.77	1	0.88
	7	2–8	0.14	0.29	0.57	0.53	0.71	1	0.03	0.23	0.51	0.50	0.77	1	0.89
	8	1–4,6–9	0.13	0.25	0.50	0.49	0.75	1	0.03	0.18	0.48	0.46	0.69	1	0.89
	9	1–9	0.11	0.22	0.44	0.49	0.67	1	0.02	0.18	0.44	0.46	0.72	1	0.90
Mean volume, m^3 ha^{-1}	4	1,3,7,9	3	87	115	125	164	355	0	56	96	97	131	251	0.38
	4	2,4,6,8	11	82	122	125	164	310	0	52	93	91	130	229	0.54
	4	3,4,6,7	4	69	125	127	172	310	0	43	93	91	135	222	0.49
	5	1,3,5,7,9	1	78	116	122	159	358	0	54	92	91	123	251	0.46
	5	3–7	5	71	119	123	156	301	0	44	91	88	124	223	0.42
	6	1,2,4,6,8,9	15	89	121	125	158	327	0	52	94	92	128	206	0.49
	6	2–4,6–8	10	84	124	129	168	310	0	46	93	87	125	223	0.47
	7	2–8	10	79	122	124	163	286	0	52	95	87	122	225	0.43
	8	1-4,6-9	10	87	123	125	159	302	0	52	92	89	123	199	0.55
	9	1–9	10	88	121	123	156	316	0	53	92	88	121	199	0.52
Mean pine volume, m^3 ha^{-1}	4	1,3,7,9	3	46	72	81	110	219	0	33	52	53	69	170	0.41
	4	2,4,6,8	2	48	69	79	103	310	0	28	49	50	69	139	0.33
	4	3,4,6,7	4	41	68	80	112	310	0	23	47	49	72	141	0.43
	5	1,3,5,7,9	1	39	68	75	100	255	0	31	47	49	68	170	0.47
	5	3–7	5	42	66	75	103	237	0	24	45	47	68	140	0.48
	6	1,2,4,6,8,9	5	46	68	78	100	283	0	31	51	51	68	125	0.36
	6	2-4,6-8	2	44	68	80	105	310	0	25	45	47	67	140	0.43
	7	2–8	2	43	67	75	98	260	0	30	46	47	66	138	0.45
	8	1–4,6–9	5	45	71	76	101	193	0	29	49	49	67	120	0.46
	9	1–9	5	45	69	73	97	189	0	30	48	48	66	123	0.50
Mean spruce volume, m^3 ha^{-1}	4	1,3,7,9	1	11	22	34	39	193	0	4.8	16	20	29	97	0.19
	4	2,4,6,8	1	10	21	32	48	123	0	3.7	12	18	28	71	0.45
	4	3,4,6,7	1	11	26	37	47	232	0	3.0	13	17	28	70	0.35
	5	1,3,5,7,9	1	11	20	32	36	256	0	4.3	15	18	27	80	0.24
	5	3–7	1	6	23	33	44	232	0	4.0	15	17	24	59	0.37
	6	1,2,4,6,8,9	1	11	22	31	42	184	0	5.4	14	18	28	81	0.35
	6	2-4,6-8	1	10	22	33	48	232	0	4.3	14	16	25	70	0.44
	7	2–8	1	8	20	31	44	232	0	4.9	14	17	25	64	0.45
	8	1–4,6–9	1	11	23	31	42	184	0	4.4	15	18	28	72	0.36
	9	1–9	1	11	23	29	40	162	0	5.9	14	17	26	62	0.39
Mean broadleaf volume, m^3 ha^{-1}	4	1,3,7,9	0	4	9	15	19	146	0	8.6	19	25	34	146	0.32
	4	2,4,6,8	0	5	8	13	15	94	0	6.2	17	23	36	108	0.42
	4	3,4,6,7	0	4	9	14	20	93	0	4.4	19	24	38	109	0.42
	5	1,3,5,7,9	0	4	8	14	17	125	0	7.8	19	24	34	125	0.38
	5	3–7	0	3	7	13	14	93	0	4.7	19	24	39	109	0.45
	6	1,2,4,6,8,9	0	4	7	11	13	113	0	6.7	18	23	32	113	0.43
	6	2–4,6–8	0	4	7	12	14	94	0	5.6	19	24	36	109	0.45

(Continued)

Appendix Table 1. Continued.

	Configuration		NFI						AUX						
	Size	Plots	Min	1Q	Med	Mean	3Q	Max	Min	1Q	Med	Mean	3Q	Max	Cor
	7	2–8	0	3	6	11	13	71	0	6.0	18	23	35	109	0.47
	8	1–4,6–9	0	3	5	10	11	83	0	7.1	18	22	31	97	0.48
	9	1–9	0	3	5	9	10	71	0	6.3	19	22	32	84	0.51
Volume variance, m^6	4	1,3,7,9	0	17	45	72	93	424	0	990	3110	3516	5354	14226	0.15
	4	2,4,6,8	0	22	53	82	111	659	0	839	2407	2786	3948	10546	0.08
	4	3,4,6,7	0	13	45	79	104	584	0	719	2458	2887	4586	10187	0.17
	5	1,3,5,7,9	1	18	47	73	91	406	0	1148	3512	3596	5369	12191	0.29
	5	3–7	0	16	48	81	99	584	0	924	2622	3056	4778	10194	0.11
	6	1,2,4,6,8,9	0	26	59	82	95	437	0	1185	3119	3246	4574	10971	0.22
	6	2–4,6–8	0	20	53	84	103	584	0	1106	2815	3097	4416	11634	0.09
	7	2–8	0	25	61	85	103	493	0	1453	3088	3213	4804	9581	0.08
	8	1–4,6–9	0	27	61	82	96	424	0	1481	3414	3458	4813	11634	0.20
	9	1–9	1	29	61	80	96	362	0	1703	3448	3510	4867	11012	0.22
Total volume, m^3	4	1,3,7,9	2.7	123	203	276	400	1039	25	456	1061	1165	1675	3831	0.64
	4	2,4,6,8	11	144	260	279	384	821	3	481	977	1178	1724	3573	0.79
	4	3,4,6,7	4.2	140	263	299	396	990	21	491	1061	1243	1851	3622	0.69
	5	1,3,5,7,9	2.7	137	254	321	450	1073	25	481	1256	1376	1994	4845	0.72
	5	3–7	6.4	172	289	350	493	1203	21	557	1325	1509	2311	4310	0.73
	6	1,2,4,6,8,9	11	168	347	376	530	1288	3	617	1531	1620	2374	5332	0.77
	6	2-4,6-8	6.4	204	344	396	510	1207	8	685	1502	1699	2546	5025	0.80
	7	2–8	6.4	232	387	443	592	1428	8	840	1653	1943	3141	5713	0.82
	8	1–4,6–9	10	206	428	488	662	1556	18	818	1855	2081	3128	7005	0.81
	9	1–9	10	249	460	535	715	1894	18	866	1996	2323	3586	7708	0.84
Total pine volume, m^3	4	1,3,7,9	2.7	72	145	173	250	661	2	229	588	635	910	2037	0.64
	4	2,4,6,8	5.6	84	159	176	262	525	0	254	571	648	955	1908	0.63
	4	3,4,6,7	4.2	83	183	199	275	650	12	218	646	682	1019	2056	0.65
	5	1,3,5,7,9	2.7	74	166	198	283	721	12	236	647	742	1119	2680	0.71
	5	3–7	5.6	98	210	228	309	686	12	318	740	818	1201	2525	0.70
	6	1,2,4,6,8,9	10	100	212	234	345	672	0	333	795	884	1372	2695	0.69
	6	2-4,6-8	5.6	100	239	250	345	808	8	321	854	936	1382	2774	0.74
	7	2–8	5.6	121	264	279	389	844	8	357	953	1061	1600	3242	0.77
	8	1-4,6-9	10	127	274	301	439	910	12	343	1045	1139	1779	3458	0.78
	9	1–9	10	130	297	327	487	946	12	433	1163	1264	2066	3926	0.81
Total spruce volume, m^3	4	1,3,7,9	2.6	23	50	86	95	574	0	40	168	243	354	983	0.33
	4	2,4,6,8	2.8	19	48	76	112	351	0	54	154	240	351	1009	0.57
	4	3,4,6,7	2.6	19	56	89	117	734	0	50	180	248	379	1070	0.47
	5	1,3,5,7,9	2.6	24	52	91	91	769	0	56	209	278	414	1162	0.39
	5	3–7	2.6	18	57	95	124	742	0	73	204	293	461	1141	0.49
	6	1,2,4,6,8,9	2.8	31	64	97	139	672	0	70	256	337	518	1675	0.52
	6	2–4,6–8	2.8	23	60	103	157	737	0	71	255	343	472	1368	0.58
	7	2–8	2.8	23	60	110	160	745	0	101	286	385	579	1419	0.60
	8	1–4,6–9	2.8	37	85	127	174	770	0	77	312	429	658	1992	0.55
	9	1–9	2.8	37	90	134	176	974	0	96	336	471	723	2085	0.58
Total broadleaf volume, m^3	4	1,3,7,9	0	0	13	43	63	443	0	67	223	287	422	1770	0.47
	4	2,4,6,8	0	0	5	41	58	297	0	67	206	291	443	1174	0.56
	4	3,4,6,7	0	0	4	39	55	418	0	66	223	314	455	1512	0.54
	5	1,3,5,7,9	0	0	19	53	80	443	0	88	272	356	523	2327	0.54
	5	3–7	0	0	10	52	89	429	0	95	295	398	559	2069	0.59
	6	1,2,4,6,8,9	0	0	21	61	96	476	0	93	284	399	630	2022	0.64
	6	2–4,6–8	0	0	13	57	76	519	0	114	285	419	632	2096	0.62
	7	2–8	0	0	22	69	107	519	0	140	357	497	740	2653	0.64
	8	1–4,6–9	0	0	33	76	116	645	0	115	396	513	789	2944	0.65
	9	1–9	0	0	40	87	128	645	0	140	457	588	910	3501	0.68

Appendix Table 2. Population parameters estimated for the cluster configurations.

Cluster configuration				Mean volumes, m^3/ha			
Size	Sample plots	Forested land, %	Total volume, m^3	Pi	Spr	BL	All
4	1, 3, 7, 9	64.6	10.1E + 06	75.8	27.9	24.3	127.9
4	2, 4, 6, 8	60.7	9.0E + 06	71.7	24.5	24.8	121.0
4	3, 4, 6, 7	61.6	9.4E + 06	76.1	26.0	22.0	124.1
5	1, 3, 5, 7, 9	64.8	9.9E + 06	72.6	26.2	25.3	124.1
5	3–7	62.4	9.3E + 06	72.7	24.7	23.5	120.9
6	1, 2, 4, 6, 8, 9	62.6	9.5E + 06	72.8	25.1	25.7	123.6
6	2–7	62.6	9.4E + 06	72.2	25.6	24.4	122.3
6	2–4, 6–8	61.4	9.3E + 06	73.4	26.2	23.6	123.2
6	3–8	61.7	9.1E + 06	71.6	24.4	23.7	119.7
7	2–8	62.0	9.2E + 06	71.3	25.2	24.4	121.0
8	1-4, 6–9	62.6	9.6E + 06	73.8	26.2	24.5	124.6
9	1–9	62.9	9.5E + 06	72.2	25.5	25.1	122.7