



Estimating the timber value of a forest property using geographically balanced samples and unoccupied aerial vehicle data

Janne Rätty^{1,*}, Juha Heikkinen², Mikko Kukkonen³, Lauri Mehtätalo ¹, Annika Kangas¹, Petteri Packalen ⁴

¹Natural Resources Institute Finland, Bioeconomy and Environment, Forest Inventory and Planning, Yliopistokatu 6, 80100 Joensuu, Finland

²Natural Resources Institute Finland, Natural Resources, Applied Statistical Methods, Latokartanonkaari 9, 00790 Helsinki, Finland

³Natural Resources Institute Finland, Bioeconomy and Environment, Forest Inventory and Planning, Paavo Havaksen tie 3, 90570 Oulu, Finland

⁴Natural Resources Institute Finland, Bioeconomy and Environment, Forest Inventory and Planning, Latokartanonkaari 9, 00790 Helsinki, Finland

*Corresponding author. Natural Resources Institute Finland, Yliopistokatu 6, 80100 Joensuu, Finland. E-mail: janne.ratty@luke.fi

Abstract

A common task in forestry is to determine the value of a forest property, and timber is the most valuable component of that property. Remotely sensed data collected by an unoccupied aerial vehicle (UAV) are suited for this purpose as most forest properties are of a size that permits the efficient collection of UAV data. These UAV data, when linked to a probability sample of field plots, enable the model-assisted (MA) estimation of the timber value and its associated uncertainty. Our objective was to estimate the value of timber (€/ha) in a 40-ha forest property in Finland. We used a systematic sample of field plots ($n = 160$) and 3D image point cloud data collected by an UAV. First, we studied the effects of spatial autocorrelation on the variance estimates associated with the timber value estimates produced using a field data-based simple expansion (EXP) estimator. The variance estimators compared were simple random sampling, Matérn, and a variant of the Grafström–Schelin estimator. Second, we compared the efficiencies of the EXP and MA estimators under different sampling intensities. The sampling intensity was varied by subsampling the systematic sample of 160 field plots. In the case of the EXP estimator, the simple random sampling variance estimator produced the largest variance estimates, whereas the Matérn estimator produced smaller variance estimates than the Grafström–Schelin estimator. The MA estimator was more efficient than the EXP estimator, which suggested that the reduction of sampling intensity from 160 to 60 plots is possible without deterioration in precision. The results suggest that the use of UAV data improves the precision of timber value estimates compared to the use of field data only. In practice, the proposed application improves the cost-efficiency of the design-based appraisal of a forest property because expensive field workload can be reduced by means of UAV data.

Keywords: stereo matching; local difference estimator; aerial imagery; remotely piloted aerial system; design-based inference; 3D point cloud

Introduction

A value assessment of a forest property is needed, e.g., when the ownership of the forest property is changed. The value assessment should provide unbiased estimates of the forest resources in a transparent and objective way. The commercial value of a forest property consists of several components with the value of the standing timber and the value of the underlying land that produces the timber (Harris et al., 2018). The common approach is to consider the discounted value of the future net harvest revenues, although the true market value of immature timber stands may be less. While the value of timber is the main driver of the value formation associated with forest properties that are harnessed for timber production, the value of the forest property may also be affected by various non-timber factors, such as carbon sequestration and conservation values (Pearce 2001). Nevertheless, the value of timber is essential and, e.g., serves as the starting point in the most common valuation method used to appraise forest properties in Finland (Eerikäinen and Venho 2018).

The timber value of a tree depends on the timber assortment products that can be derived from the stem. The standard timber assortment products in the Nordic forestry are logwood, pulpwood, and energy wood. Given the importance of the timber assortment volumes in the value formation of forests, their estimation has been of major interest in stand-level forest inventories (Karjalainen et al. 2019; Korhonen et al. 2008; Peuhkurinen et al. 2008). In Finland, estimates of the volume of timber assortment by tree species are available in public stand register data (Maltamo and Packalen 2014), although the errors associated with the estimates are too large for the purpose of timber valuation (Vähä-Konka et al. 2020). In addition, the temporal resolution of operational large-area inventories is typically several years, which impedes the availability of up-to-date forest information in the absence of growth models.

Owing to the errors associated with forest attribute maps and the outdated forest information contained in the public stand registers, the valuation of timber in a forest property has traditionally relied on stand-level field visits. Field visits are, however, time

Handling editor: Prof. Teja Kattenborn

Received: November 10, 2023. Revised: February 29, 2024. Accepted: March 3, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Institute of Chartered Foresters.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

consuming and relatively expensive because each forest stand in a forest property must be assessed. Some of the stand-level field inventories also involve subjective components. For example, Finnish stand-level assessments rely on angle-count sampling points that are subjectively spread across forest stands. At each sampling point, the basal area median trees to be calipered are selected by stratum using visual judgment. These subjective components introduce errors into the forest attribute estimate that may vary considerably across operators (Kangas et al. 2004). Alternatively, the timber valuation can rely on a probability sample of field plots that represents the target population and enables the design-based inference of the population parameters using a design-unbiased estimator of mean (or total) timber value. However, the collection of a probability sample for a field data-based estimation of timber value in a forest property is too laborious in most operational settings. The use of remotely sensed data as auxiliary information in the estimation procedure may enable the reduction of the field workload (Puliti et al. 2020).

Unoccupied aerial vehicles (UAVs) have become established platforms for the collection of remotely sensed data in small-scale forest inventories (Hao et al. 2022; Kotivuori et al. 2020; Kukkonen et al. 2021; Puliti et al. 2020). The current UAV technology restricts the geographical extent of UAV data to areas that typically cover a maximum of hundreds of hectares. This explains why the forestry applications of UAV typically focus on inventories of forest stands or forest properties. In practice, the price per hectare is very high in inventories that focus on target populations of less than a few hectares unless the use of UAV data can considerably reduce the workload associated with the collection of field data.

In forestry applications, UAV is usually equipped with a camera that captures aerial images with spectral recordings of the targets of interest. The spectral images can be used to construct image-based (i.e. photogrammetric) three-dimensional (3D) point clouds (IPC) via photogrammetric processing (Bolles et al. 1987). The use of UAV platforms enables the collection of up-to-date 3D data for small areas whenever required, whereas airplane-based campaigns, e.g. country-level acquisition programs for airborne lidar (light detection and ranging) data or aerial images, typically run with temporal resolutions of several years. The temporal mismatches between field and remotely sensed data are undesirable in forest inventories. For example, natural disturbances or silvicultural operations may cause that remotely sensed data collected 1 year ago do not represent the current characteristics of a forest. For this reason, the acquisition of UAV data may be preferable to airplane-based acquisitions at some scales of operation.

The estimation of forest attributes for a finite target population can be carried out using (1) a probability sample of field data with a simple expansion (EXP) estimator to expand the sampled units to the level of target population, (2) a model-assisted (MA) estimator that combines wall-to-wall auxiliary data and a design-based inference of a field-measured probability sample, or (3) a model-based estimator that utilizes auxiliary data, as efficiently as a MA estimator, but without the statistical properties assured by design-based inference. The EXP and MA estimators are design-based estimators that provide design-unbiased estimates for mean and total values of a target population, whereas the properties of the model-based estimator depend on model specification, which means that there is no guarantee of unbiasedness of estimates in a target population. The MA estimator is more efficient than the EXP estimator, given that the variation of the residuals in the sampling units is less than the variation

of the corresponding observed values (Breidt and Opsomer 2017). The use of the MA estimator may result in cost-savings because the number of field plots required for sufficient precision may be reduced due to improved efficiency compared to the EXP estimator.

The estimation of population means (and total values) from a systematic sample can be more efficient, i.e. it enables smaller variance estimates for population means or total values than the estimation based on a simple random sampling (SRS) design (Dunn and Harrison 2018; Mostafa and Ahmad 2018). An inferential challenge associated with the systematic sampling design is that a single systematic sample does not provide a design-unbiased variance estimator to assess the uncertainty of population estimates. For this reason, the efficiency of systematic sampling compared with SRS, e.g., is difficult to assess in practical applications. Variance of an estimator of the population mean (or total values) based on a systematic sampling design is often estimated under the SRS assumption (Magnussen et al. 2020). When systematic sampling is more efficient than SRS, this leads to an overestimation of the true variance (Rätty et al. 2020), although this is not considered as a severe disadvantage in most real-life applications (Magnussen et al. 2020). However, the overestimation of variance may be especially suboptimal in target populations that cover tens of hectares because the collection of too many sampling units, in terms of desired sampling error, rapidly increases the inventory costs per hectare.

There exist several variance estimators that are capable of mitigating the overestimation of variance that results from the false SRS assumption under systematic sampling (Magnussen and Fehrmann 2019; Rätty et al. 2020). A straightforward solution to reduce the upward bias of the SRS variance estimator is to apply an estimator that accounts for spatial autocorrelation among the field plots of a probability sample. So-called local difference estimators are easy to adapt into systematic designs because the model is not required to characterize the spatial autocorrelation in the target population. A simple local difference estimator was adapted to the line survey designs of the first Nordic national forest inventories (NFIs) (Langsæter 1932; Lindeberg 1923). Later, Matérn (1947) elaborated the idea of a local difference estimator for a spatially systematic sampling design. The Matérn (MT) variance estimator considers rectangular-shaped neighborhoods of four sampling units and has been in use in the Finnish NFI since the 5th inventory cycle (Salminen 1973). More recently, Grafström and Schelin (2014) proposed a local difference estimator for spatially balanced sampling, which follows the ideas of the local mean variance estimator introduced in Stevens and Olsen (2004). In contrast to the MT estimator, the Grafström–Schelin (GS) variance estimator does not require a regular grid of sample plots, and the estimator can also deal with other spatially balanced designs. There is a well-established consensus in the literature that the MT and GS variance estimators provide more realistic variance estimates than the SRS estimator under a systematic sampling design (Magnussen et al. 2020).

Here, we estimate the mean timber value (€) per hectare for a 40-ha forest property using geographically balanced sampling and UAV data. The UAV data are used as auxiliary information in the MA estimator. We consider solely the current value of standing timber without accounting for future net harvest revenues because there is no reason to assume that the real harvest revenues will differ between the studied estimation alternatives. We also ignore the value of bare land because it is constant. The study objectives are

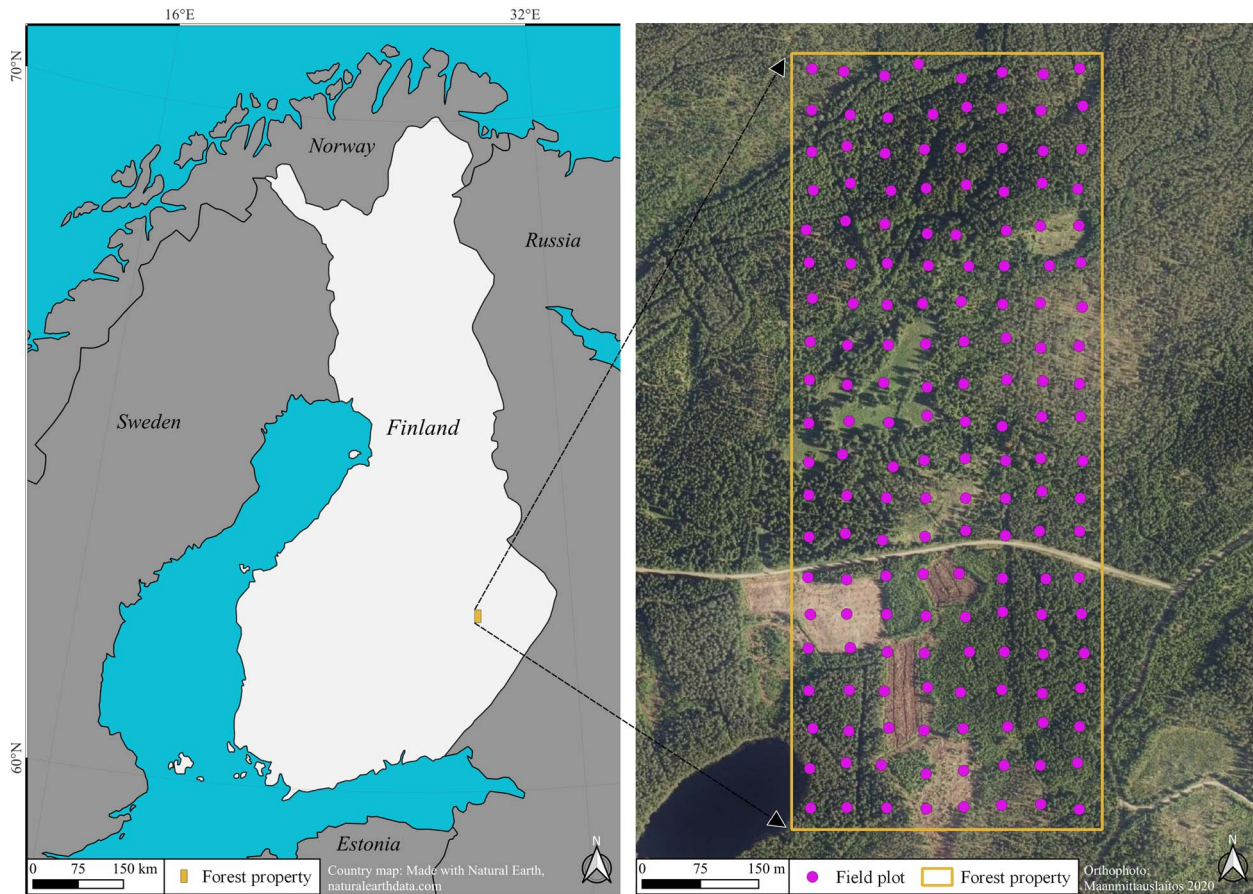


Figure 1. Location of the study site in eastern Finland (left). Systematic sample grid of 160 field plots on the 40-ha forest property (right).

- (1) To compare the SRS and local difference variance estimators for the estimation of timber value using different sampling intensities in the EXP and MA estimation frameworks. Our specific interest is to examine how comparable the results produced using the SRS variance estimator are with those of the local difference estimators under the presence of spatial autocorrelation that is evident when operating at the level of the forest property.
- (2) To investigate the estimated variances associated with the MA and field data-based EXP estimators to assess the efficiency gains that result from the use of auxiliary UAV data in the estimation of timber value.

Materials and Methods

Forest property

The 40-ha forest property of interest (62°36' N 29°2' E) is located in eastern Finland (Fig. 1). The location of the forest property was arbitrarily selected from an inventory area that was originally established for other research and, therefore, the property does not follow the boundaries of a real forest estate. The property represents typical Finnish coniferous-dominated forests that are mainly used to produce timber following the principles of stand-level forest management. The property contains forest stands that cover a range of development stages.

The most abundant tree species are Norway spruce (*Picea abies* [L.] Karst.) and Scots pine (*Pinus sylvestris* [L.]). Deciduous species, mostly birch species (*Betula pendula* Roth. and *Betula pubescens* Ehrh.), are commonly found as mixtures, but they may also grow

as dominant species on fertile soil types and bogs. Other species, such as aspen (*Populus tremula* [L.]) and alder (*Alnus* spp.), occur as minor species on fertile soil types.

Field data

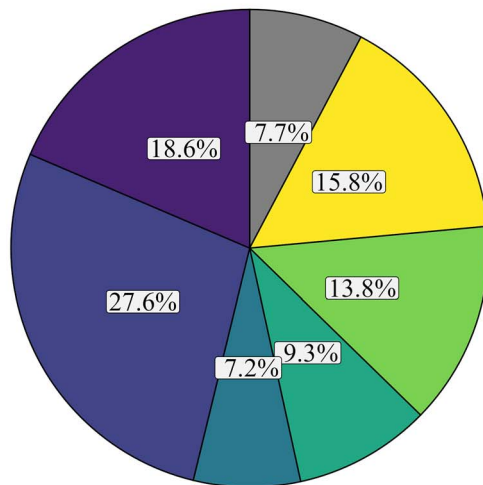
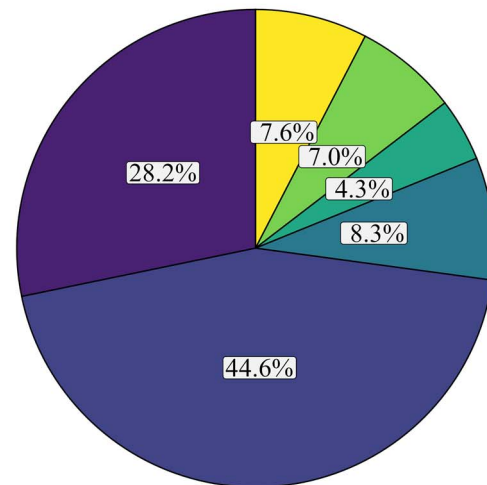
A sample size of 160 field plots was systematically measured within the forest property. The location of the upper-left plot was selected randomly, and the other plots were placed using a systematic 50 m × 50 m grid. The fieldwork crew navigated to these initial plot locations using a low-grade differential global navigation satellite system (GNSS) device. The final plot location was determined by selection of a random direction (0–359°) and distance (0–5 m) from the initial plot location. Therefore, the plot network was not entirely systematic, and the distance between plots was variable (Fig. 1). The centers of the field plots were positioned with submeter accuracy using a Trimble GeoXH DGNS device. The field campaign was implemented between 7 September and 1 October 2020.

The field plots were concentric circular plots of radii 5.64 m and 7.98 m. Diameter at breast height (DBH) thresholds were used to speed up the measurement of the field plots. Trees with a DBH >10.0 cm were measured in the 7.98-m-radius plots, while trees with a DBH between 5.0 and 10.0 cm were measured from the 5.64-m-radius plots. Height and DBH were measured, and tree species was determined for each tree within the plot. Because of the small number of deciduous species present in the plots (other than birch), all deciduous species were grouped in the same species class *deciduous* in the subsequent analyses.

Table 1. Bucking dimensions and the price (€/m³) of pulpwood and logwood used in the calculation of timber values.

	Logwood			Pulpwood		
	Norway spruce	Scots pine	Deciduous	Norway spruce	Scots pine	Deciduous
D_{\min} (cm)	16.0	15.0	18.0	7.0	7.0	7.0
L_{\min} (m)	3.7	3.7	3.7	2.0	2.0	2.0
L_{\max} (m)	6.1	6.1	6.1	–	–	–
Price (€/m ³)	71.9	67.7	51.0	22.6	20.5	21.4

D_{\min} , minimum top diameter of a log; L_{\min} , minimum log length; L_{\max} , maximum log length

A) Proportion of timber volume**B) Proportion of timber value**

Pine logwood
 Pine pulpwood
 Non-commercial

Spruce logwood
 Spruce pulpwood

Deciduous logwood
 Deciduous pulpwood

Figure 2. (A) Proportions of timber assortment volumes in the 160 field plots. (B) Proportions of timber assortment values in terms of total timber value. For timber assortment prices, please refer to Table 1.

The species-specific taper curves described in Laasasenaho (1982) were employed to buck stem volumes into logwood and pulpwood assortment volumes. The bucking procedure maximized the length of the logwood proportion using 30-cm intervals subject to the dimensional criteria of commercial logwood. The pulpwood proportion was calculated as the remainder of the stem until the minimum diameter of the pulpwood was reached after the logwood portion had been bucked. Stem volume that did not fulfill the dimensional criteria of commercial logwood or pulpwood were allocated to the non-commercial category. The bucking dimensions are shown in Table 1. The proportions of timber assortment volumes in terms of total volume in the field plots are shown in Fig. 2A.

The timber assortment volume and their commercial prices define the value of the standing timber. We used the average timber prices in November 2022 (Luonnonvarakeskus 2022). The timber prices by assortment and tree species are shown in Table 1. The tree-level values (€) were further aggregated at the plot level and scaled up to the hectare level to compute the timber value (€/ha). The hectare-level timber value was used as a response variable in the estimation of timber value at the level of the forest property. The proportions of timber assortment values in terms of the total timber value in the field plots are shown in Fig. 2B. A summary of statistics associated with

the field data collected from the forest property is shown in Table 2.

We also used an exogenous set of field plots for the selection of predictor variables for the linear regression model used in the MA estimation (see section “Estimation of timber value”). We refer to the exogenous set of field plots as variable selection data. The variable selection data comprised 156 square plots of 225 m², which were located nearby in the same region but did not overlap with the forest property. The measurement protocol of the variable selection data was similar to the protocol used in the field plots in the forest property with the exception that trees with a DBH >5.0 cm were measured. Note that the exogenous set of field plots is not necessary, albeit recommended, from the viewpoint of operational application, but the model can be formulated based on expert knowledge and the avoidance of complex models. The use of too complex model fitted using endogenous data may lead to the underestimation of variance (Kangas et al. 2016).

UAV data

The UAV images from the 40-ha study area were captured on 27 September 2020, and images from the variable selection data area were captured during summer 2020. The UAV images were captured using a DJI M210 quadcopter (SZ DJI Technology Co., Ltd., Shenzhen, China), which was equipped with two cameras:

Table 2. Mean, standard deviation (SD), and minimum and maximum values of the forest attributes associated with the field data ($n = 160$), collected from the forest property.

	Mean	SD	Min	Max
Volume (m ³ /ha)	152.4	123.9	0	517.7
Logwood volume (m ³ /ha)	81.4	101.0	0	461.5
Pulpwood volume (m ³ /ha)	59.2	49.2	0	243.2
Basal area weighted mean diameter (cm)	18.7	7.1	5.0	39.7
Basal area weighted mean height (m)	15.7	5.2	3.7	27.3
Timber value (€/ha)	6784.9	7199.9	0	30 994.6

a high-resolution camera with conventional red-green-blue (RGB) bands (DJI X5S) and a multispectral (MS) camera that, in addition to RGB, included red-edge (RE) and near-infrared (NIR) bands (AgEagle, Wichita, KS, USA; MicaSense series, model Altum). The pixel values of the MS data were converted to reflectance values (0–1) via reflectance panels using the procedure recommended by MicaSense (MicaSense 2020). Flight parameters were chosen based on the MS camera characteristics for 80% lateral and forward overlap, resulting in an 85% overlap for the DJI camera. The operation altitude of 100 m produced a nominal ground sample distance of 4.3 cm and 2.3 cm for the MS and RGB data, respectively.

Agisoft Metashape (version 1.8.0) was utilized to generate IPC from the UAV image data, using “high” quality image alignment and “high” quality dense point cloud generation with “moderate” filtering. The same parameter values were applied to both MS and RGB data. The MS images were not used to create photogrammetric point clouds, but the exterior orientations and the internal orientation of the MicaSense Altum camera were exported from Metashape after image alignment. These orientations were utilized to assign a MS reflectance value to each point in the IPC using collinearity equations, as described in Packalén et al. (2009). The IPC were then normalized to ground level using previously acquired low-density airborne laser scanning (ALS) data, which belong to the open data of the National Land Survey of Finland (National Land Survey of Finland 2022). A triangulated irregular network (TIN) was created from ALS echoes classified as ground (Axelsson 2000), and the TIN model was subtracted from the point heights of the IPC to compute the above-ground heights.

We extracted plot-level metrics from the height and spectral values associated with the IPC. The height metrics included mean, median, height standard deviation (sdh), skewness, kurtosis, quantiles (q5, q10, . . . , q95), and densities at fixed heights of 1.3, 2.5, 5.0, 10.0, 15.0, 20.0, and 25.0 m (d1.3, d2, . . . etc.). The reflectance metrics comprised mean, standard deviation, skewness, and kurtosis by MS camera bands (R, G, B, RE, NIR). In addition, between-band ratios of the mean and sdh metrics were computed.

Subsamples of the systematic sample

Subsampling by rows and columns

Subsampling of the systematic sample must maintain the geographical balance of the field plots in the population. We refer to subsampling by rows and columns when the subsampling of the systematic sample was carried out by the removal of complete rows or/and columns of the systematic sample. We selected samples that exhibited similar between-plot distances for plots that were in the same sample column or row, although the between-plot distance was allowed to be dissimilar between sample columns and rows. This ensured that every subsample relied on a rectangular grid, thereby enabling the use of the MT

variance estimator (Matérn 1947). In total, 17 subsamples, which included 40, 48, 56, 60, or 80 field plots, were created. An example of the realization of 80 plots subsampled by rows and columns from the systematic sample can be seen in Fig. 3A.

Randomized subsampling with a geographical balance

While the subsampling procedure described in the previous section produces subsamples that rely on a rectangular grid, the number of field plots to be selected cannot be freely controlled. Subsampling by rows and columns results in a limited number of rectangular samples with limited possibilities to repeat the estimation. For these reasons, we proposed a repeatable randomized subsampling procedure that still ensured the spread of field plots over the forest property in a geographically balanced manner.

The randomized subsampling was assisted by a subsampling grid that consisted of 200 m × 200 m squares that were overlaid on the forest property. In the rectangular-shaped forest property of 40 ha, the subsampling grid consisted of 10 cells. The cells were used to allocate the desired number of field plots by randomly selecting $n/10$ plots per cell. The use of 10 cells enabled the iteration of the subsampling of systematic sample m times ($m = 100$) using n values of 40, 60, . . . , 140, in order to study the ranges of estimates and their variances under different sampling intensities. An example of a randomly subsampled realization of 80 field plots can be seen in Fig. 3B.

Voronoi tessellation

A Voronoi tessellation was created for each realization of the sample based on the XY locations of the field plots using the `deldir` package (Turner 2021) in the R environment (R Core Team 2023). Motivation for the use of the Voronoi tessellation was 2-fold. First, the cell areas of the Voronoi tessellation were used as weightings to account for the geographical coverage of the subsample that aligns suboptimally in the population. Henceforth, we refer to the plot-level cell areas of the Voronoi tessellation as sampling weightings. For example, subsampling with removal of the easternmost column of plots causes under-representation of field plots on the eastern edge area of the population. In that case, we set larger sampling weightings for the easternmost field plots than in the other plots. Second, the Voronoi tessellation was used to determine the local neighborhoods of the GS estimator. Examples of the created Voronoi tessellations based on the field plots are shown in Fig. 3.

Estimation of timber value

Workflow

The estimation framework of this study is illustrated in Fig. 4. The methodology associated with the estimators is described in detail in sections “Simple expansion estimator” and “Model-assister estimator”. The estimation framework consists of three steps. First, the systematic sample was subsampled as described in section

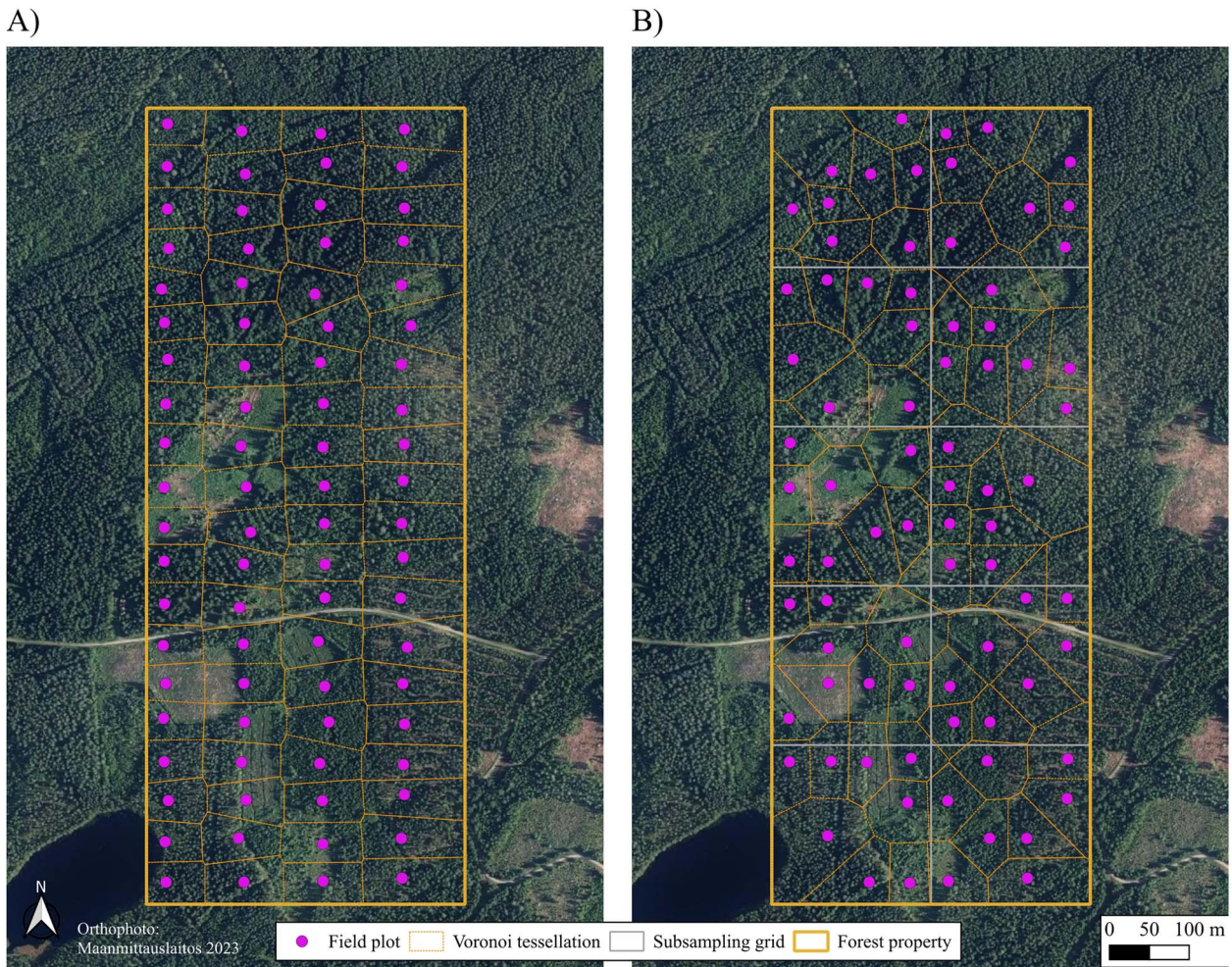


Figure 3. Examples of subsampling procedures used to select 80 plots from the systematic sample of 160 plots. (A) An example of subsampling by rows and columns. (B) An example of randomized subsampling.

“Subsamples of the systematic sample”. Second, the subsamples were used as such in the EXP estimator and in conjunction with the UAV data in the MA estimator to produce estimates of timber value. Third, the variance estimators of SRS, MT, and GS were applied to the subsamples by rows and columns, whereas the SRS and GS variance estimators were applicable to the randomized subsamples.

Simple expansion estimator

We used the EXP estimator to produce mean estimates of timber value $\hat{\mu}_{EXP}$ based on field plots without auxiliary data. The EXP estimates were computed as weighted averages using sampling weightings. The estimated mean timber value (€) per hectare $\hat{\mu}_{EXP}$ is

$$\hat{\mu}_{EXP} = \frac{\sum_{i \in S} w_i y_i}{\sum_{i \in S} w_i} \quad (1)$$

where y_i is the timber value (€/ha) in the field plot of sample i in S , and w_i is the sampling weighting (in hectares) derived from the Voronoi tessellation.

A variance estimator for $\hat{\mu}_{EXP}$ assuming SRS is

$$\hat{v}ar_{SRS}(\hat{\mu}_{EXP}) = \frac{\hat{\sigma}^2}{n} \quad (2)$$

where $\hat{\sigma}^2 = \frac{\sum_{i \in S} w_i (x_i - \hat{\mu}_{EXP})^2}{\sum_{i \in S} w_i (n-1)/n}$ is the sample variance and n refers to the number of field plots in sample S .

Matérn (1947) proposed the MT variance estimator, which is suitable for samples that rely on rectangular grids. The estimator accounts for spatial autocorrelation by inspecting the local differences associated with rectangular-shaped neighborhoods of four plots in the sample. The MT variance estimator is as follows:

$$\hat{v}ar_{MT}(\hat{\mu}_{EXP}) = \frac{1}{4(\sum_{i \in S} w_i)^2} \sum_g (z_{g_1} - z_{g_2} - z_{g_3} + z_{g_4})^2 \quad (3)$$

where $z_{g_{1..4}} = y_{g_{1..4}} w_{g_{1..4}} - \hat{\mu}_{EXP} w_{g_{1..4}}$, g refers to the rectangular-shaped neighborhood of the MT estimator, and subscript $g_{1..4}$ refers to the corner plots of the rectangular-shaped neighborhood. The number of rectangular-shaped neighborhoods $g_{1..4}$ in the sampling design was set equal to n . The $z_{g_{1..4}}$ values of missing field plots associated with the neighborhood, i.e. plots that locate outside the population of interest, were set to zero.

The GS variance estimator is another local difference estimator that relies on the neighborhood of field plots (Grafström and Schelin 2014). The neighborhood is determined based on a distance measure, which is the Euclidean distance computed based on XY coordinates. The GS estimator determines the nearest neighbor plots based on an equal distance value measured from the plot in the center of the neighborhood. This assumption means that the

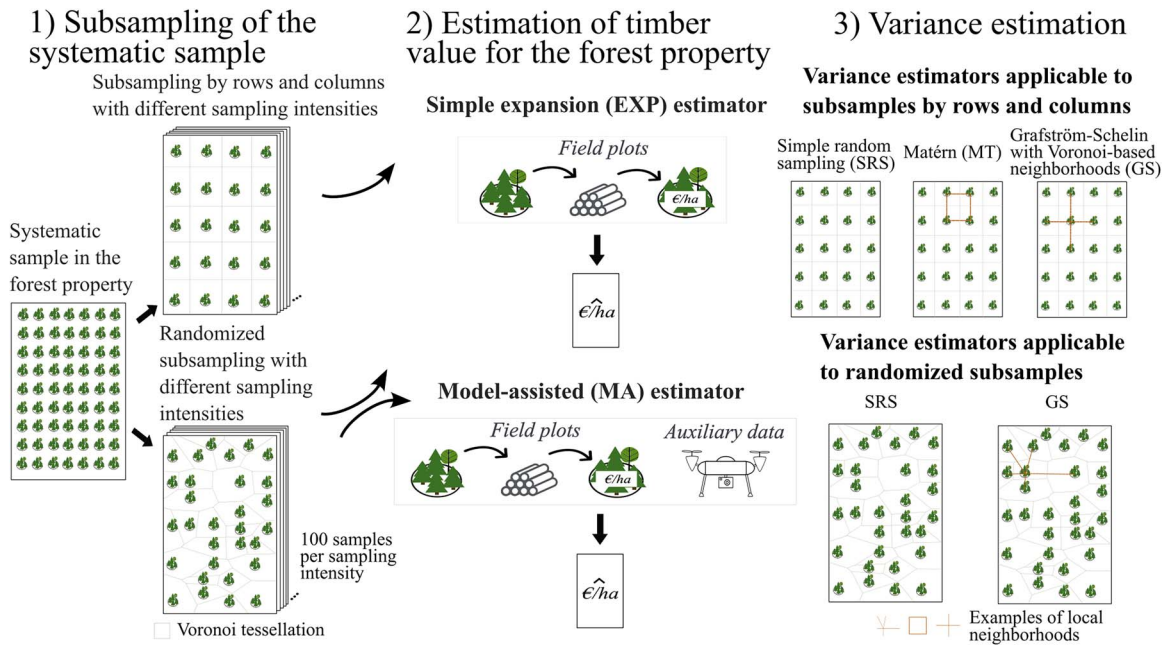


Figure 4. An illustration of the framework associated with the estimation of timber value per hectare.

number of neighbor plots is four for internal plots in a systematic sample. Since the assumption of equal distance would always lead to a single neighbor plot in our sample, we relaxed the assumption of equal distance. The neighborhood, consisting of n_i^* plots, including the center plot i , was determined based on the Voronoi tessellation. The neighborhood was constructed by selecting the center plot of the neighborhood and all the plots that had a Voronoi cell adjacent to the Voronoi cell of the center plot. The neighborhood was further filtered by omitting neighbors whose Voronoi cells had a touching surface length $<5\%$ in terms of the perimeter of the Voronoi cell of the center plot. The GS variance estimator is as follows:

$$\text{var}_{\text{GS}}(\hat{\mu}_{\text{EXP}}) = \sum_{i \in S} \frac{n_i^*}{n_i^* - 1} \left(\frac{y_i}{\pi_i} - \frac{1}{n_i^*} \sum_{j \in S_i^*} \frac{y_j}{\pi_j} \right)^2 / A^2 \quad (4)$$

where π refers the inclusion probability associated with plot i or j , i.e. inversed sampling weight, $A^2 = (\sum_{i \in S} w_i)^2$ is the square of the area covered by the forest property (in hectares), and n_i^* is the number of neighbor plots in neighborhood S_i^* for plot i of sample S .

The standard error (SE) of an estimate is

$$\text{SE}(\hat{\mu}_{\text{EXP}}) = \sqrt{\text{var}(\hat{\mu}_{\text{EXP}})} \quad (5)$$

Model-assisted estimator

We fitted a linear regression model that was used as an assisting model in the MA estimation:

$$\sqrt{y_i} = \beta_0 + \beta_1 x_i^{(1)} + \beta_2 x_i^{(2)} + \beta_3 x_i^{(3)} + e_i, \quad e_i \sim N(0, \sigma^2/w_i) \quad (6)$$

The assisting model explains the relationship between observed timber value y_i in the field plots of the forest property and UAV metrics $x_i^{(1)} \dots x_i^{(3)}$. The model was used to predict timber values of grid cells (\hat{y}_k) and timber values of field plots (\hat{y}_i). The sample weightings (w_i) of the field plots derived from the

Voronoi tessellation were considered in the modeling process by estimating model coefficients using the weighted least squares technique. We square-root-transformed the response variable to avoid negative predictions and to improve the model fit. The square-root-transformed predicted values were transformed back during the prediction phase. The transformation causes bias that was corrected by adding the squared standard error of the residuals to the predicted values (Lappi 1993).

We selected the predictor variables using the independent plot data to minimize the risk of optimizing the model according to the target population when using an automatized variable selection algorithm. The predictor variables of the regression model were selected using a heuristic optimization algorithm, known as simulated annealing (Kirkpatrick et al. 1983). Our implementation is a modified form of the algorithm presented by Packalén et al. (2012), where the algorithm iteratively selects an optimal set of desired number of predictor variables (in our study, 3), while the number of iterations is controlled by an initial temperature, a cooling multiplier, and the number of inner iterations per temperature value. The initial temperature was set to 1, the cooling multiplier was set to 0.99, and the number of iterations per temperature was 50. The algorithm uses mean squared error as a cost function.

The MA estimator employs the probability sample of field plots and the predictive model fitted using the sample field plots and UAV data. The MA estimator (Särndal et al. 1992, section 6.3) for mean timber value is

$$\hat{\mu}_{\text{MA}} = \frac{\sum_{k \in U} \hat{y}_k}{N} + \frac{\sum_{i \in S} w_i e_i}{\sum_{i \in S} w_i} \quad (7)$$

in which the first part of the estimator is an estimate collected from wall-to-wall timber value predictions \hat{y}_k for N grid cells (cell size 204 m²) in target population U , and the second part is an estimate of a correction term in which e_i refers to the residual ($y_i - \hat{y}_i$) in field plot i that belongs to sample S . The correction term corrects the systematic error associated with the model predictions.

The variances associated with the MA estimates were calculated based on the model residuals of the field plots. The use of model residuals from the sample of the target population in the variance estimator may incorporate the risk of underestimating variances, particularly in the case of small samples, because the ideal population residuals are replaced with sample-based residuals (Särndal et al. 1992, section 6.6). Following the idea described by Lundström and Särndal (1999), we adjusted the SRS and GS variance estimators for the loss of degrees of freedom that resulted from the estimation of model parameters, by multiplying sample residuals e_i with a simple adjustment term

$$f_{adj} = \frac{n-1}{n-(M+1)} \quad (8)$$

where $M=3$ corresponds to the number of predictor variables in the regression model.

The SRS and GS variance estimators described in Eqs. 2 and 4 were used for the MA estimates. The sample variance $\hat{\sigma}^2$ of the SRS estimator (Eq. 2) was replaced with $\hat{\sigma}^2 = \frac{\sum_{i \in S} w_i (\hat{e}_{adj,i} - \bar{e}_{adj})^2}{\sum_{i \in S} w_i (n-1)/n}$ where \bar{e} is the weighted mean of adjusted residuals observed in the field plots. In the case of the GS estimator (Eq. 4), y_i and y_j were replaced with the adjusted residual values in plot i and j .

Evaluation of the estimators

We evaluated the precision of the EXP and MA estimates using $1.96 \times SE$ that corresponds to a half-width of the 95% confidence interval (CI). The CIs were reported as percentages with respect to the mean timber value estimate. In addition to the CIs, we studied the efficiency of estimators using relative efficiency (RE), which is a ratio of two variance estimates. The RE values were computed with respect to the variance estimates associated with the EXP estimator and employed the systematic sample of 160 plots. For instance, $RE = \hat{var}(\hat{\mu}_{EXP}^{160}) / \hat{var}(\hat{\mu}_{MA})$ and results in an RE value >1 if the MA estimator is more efficient than the EXP estimator that employs 160 field plots. The RE values under SRS can be interpreted via the number of field plots. For example, an RE value of 2 indicates that a variance estimate equal to $\hat{var}(\hat{\mu}_{MA})$ can be achieved by doubling the number of field plots associated with $\hat{var}(\hat{\mu}_{EXP})$.

Results

Comparison of SRS, GS, and MT variance estimators

The SRS variance estimator produced the largest variance estimates, while the MT estimator resulted in smaller estimates than the GS estimator (Fig. 5). The comparison of variance estimates also indicated that the difference between the MT and GS variance estimates diminished when sampling intensity increased.

The MA estimation

The variable selection algorithm selected the following predictor variables for the assisting model of timber value: the 85th percentile of height values (q_{85}), the ratio of mean reflectance values from the green- and red-edge bands ($mean_G/RE$), and the standard deviation of height values (sdh). We present parameter estimates associated with the systematic sample of 160 field plots in Table 3.

Timber value estimates produced using the EXP estimator and associated variance estimates differed among the 100 iterations when the sampling intensity was low, i.e. when the number of

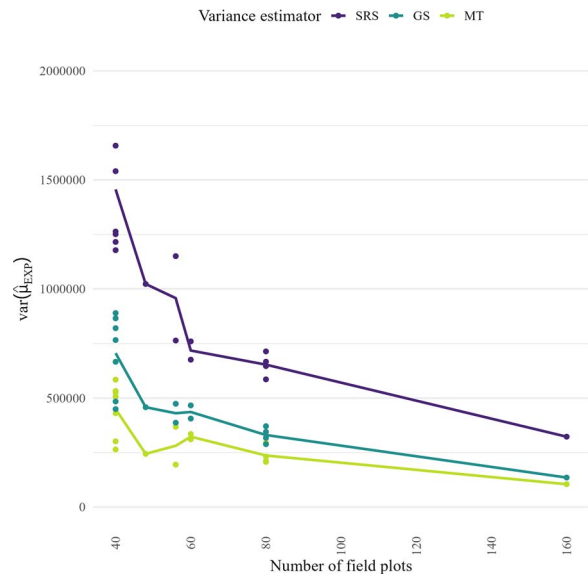


Figure 5. Variance estimates for the estimated mean timber value ($\hat{\mu}_{EXP}$) using simple random sampling (SRS), Matérn (MT), and Grafström–Schelin (GS) variance estimators. Points refer to the individual samples. Lines show the mean trend. EXP—simple expansion estimator.

Table 3. Estimated regression coefficients (weighted least squares) fitted using a systematic sample of 160 plots.

	Predictor variable	Estimate (Std. error)
β_0	(Intercept)	-54.30 (13.47)
β_1	q_{85}	7.59 (0.35)
β_2	$mean_G/RE$	105.28 (31.27)
β_3	sdh	-3.64 (1.13)
σ^2		7.84 ²
R^2		0.88

The response variable (timber value, €/ha) was square root transformed.

field plots <80 . For example, the minimum and maximum estimated timber value with 40 plots were 5372 €/ha and 8851 €/ha, respectively. As a comparison, the corresponding estimates with 100 field plots were 6079 €/ha and 7649 €/ha, respectively. The use of the MA estimator decreased the fluctuation associated with the estimates compared to the EXP estimator. For example, the minimum and maximum estimated timber value with 100 field plots with the MA estimator were 6361 €/ha and 7104 €/ha, respectively. The CIs associated with the EXP and MA estimates at sampling intensities of 40, 60, ..., 140 field plots are shown in Fig. 6. The systematic sample of 160 field plots produced CI values of 10.6% and 16.4% with the GS and SRS variance estimators for the EXP estimates, respectively. Correspondingly, CI values of 5.8% and 6.3% were obtained with the GS and SRS variance estimators for the MA estimates.

The reported CIs associated with the timber value estimates indicated that the MA estimator was more efficient than the EXP estimator, regardless of sampling intensity (Fig. 6). The CIs associated with the EXP estimates were, on average, smaller with the GS variance estimator than with the SRS variance estimator. In the case of the MA estimates, the CIs associated with the GS and SRS variance estimators were not as distinct as with the EXP estimates. At the lowest sampling intensities, the SRS variance estimator produced smaller CIs associated with the MA estimates than the GS estimator.

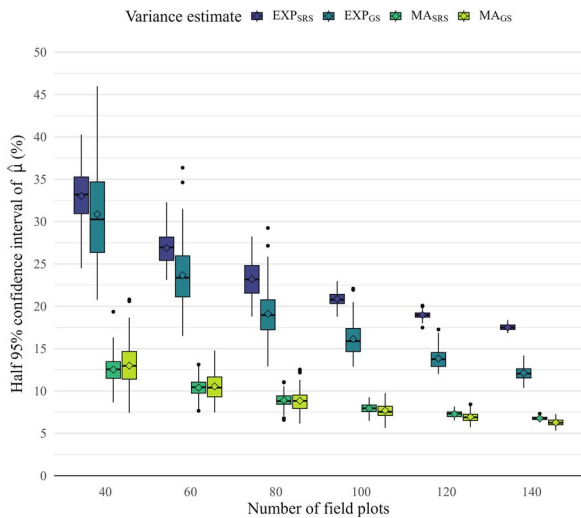


Figure 6. Variation in the estimated half 95% confidence intervals (CIs) of the timber value estimates ($\hat{\mu}$) among 100 iterations of the subsampling procedure per number of field plots (sampling intensity). The variance estimates were produced by simple random sampling (SRS) and Grafström–Schelin (GS) variance estimators. Diamond symbols denote the mean value. MA—model-assisted estimator, EXP—simple expansion estimator.

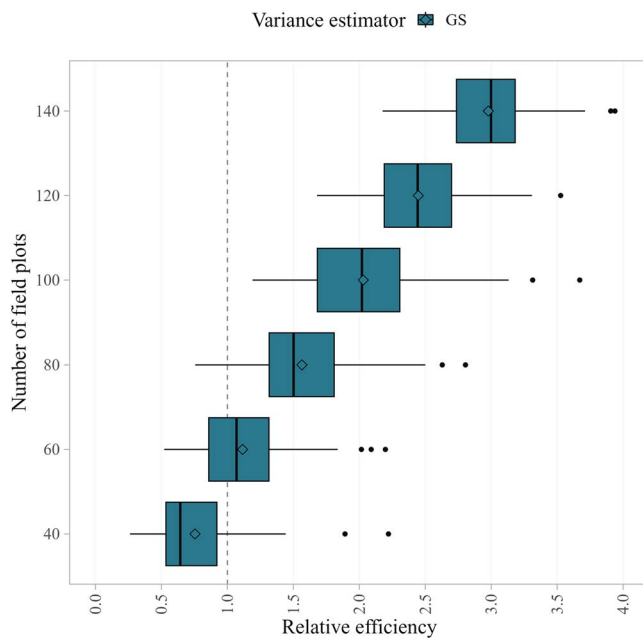


Figure 7. Relative efficiency (RE) showing the efficiency gain achieved by the assisting model when the Grafström–Schelin (GS) variance estimator was used. The RE values were computed in terms of the variance estimates of simple expansion estimator that employed a systematic sample of 160 field plots. The RE values were collected over 100 repetitions of the subsampling procedure. Diamond symbols denote the mean value over 100 repetitions.

The efficiency gain, with respect to the field data-based EXP estimator employing 160 plots, achieved by the assisting model is illustrated in Fig. 7. The MA estimator at a sampling intensity of ≥ 100 field plots was always more efficient than the EXP estimator used with the systematic sample of 160 plots. On average, the MA estimator with 60 plots was roughly as efficient as the EXP estimator with 160 field plots (mean RE: 1.1).

Discussion

Our results showed that the SRS variance estimator produced larger variance estimates than the GS and MT estimators. This finding is in line with the previous studies that reported that the SRS variance estimator produces conservative variance estimates under systematic sampling designs (Heikkinen 2006; Magnussen et al. 2020; Rätty et al. 2020). However, the discrepancy in variance estimates provided by the MT and GS variance estimators associated with the EXP estimates is not well supported by the previous literature. For example, Rätty et al. (2020) reported that the MT and GS estimators produced similar variance estimates under the systematic cluster design of the Finnish NFI, which has a considerably lower sampling intensity than the sampling designs in our study. Here, the MT estimator produced smaller variance estimates than the GS estimator. This can be largely explained by the geographical extent of the local neighborhoods, which are smaller in the MT estimator than in the GS estimator. The latter was modified to use the Voronoi tessellation to determine the neighborhoods. For this reason, the MT estimator was more effective in accounting for spatial autocorrelation than the GS estimator. However, it is safe to assume that the MT variance estimator rarely produces underestimates of actual variance (Magnussen et al. 2020).

The MT estimator requires a sampling design that relies on a rectangular grid, which could hamper its practical applicability. It is worth noting that in this study, we relaxed some of the assumptions associated with the GS estimator to ensure that the GS variance estimator would consider local neighborhoods of more than two field plots in our samples. Against the assumptions proposed in Grafström and Schelin (2014), we allowed the calculation of the local differences in the neighborhoods of unequal Euclidean distances computed from the center of the neighborhood to the neighbor plots. The distance relaxation of the GS variance estimator and the irregularity of our sample plot locations increased the geographical extent of the local neighborhoods. As such, the variance estimates may be more conservative compared with truly systematic samples that rely on a regular grid.

The CI estimates indicated that the use of the assisting model in the MA estimator resulted in efficiency gains compared with the EXP estimator. The CI estimates produced using the SRS and GS variance estimators were largely at the same level regardless of sampling intensity when the MA estimator was applied. This indicates that spatial autocorrelation among the field plots can be accounted for by the assisting model. This finding suggests that little benefit can be achieved using a local difference variance estimator when MA estimation with a well-performing assisting model is used. Similar findings were also reported by Rätty et al. (2021), who compared the GS and SRS variance estimators in the MA estimation of stem frequencies under the systematic Norwegian NFI design, although, in that case, spatial autocorrelation was considerably lower. Our findings also indicated that the decrease in sampling intensity caused SRS variance estimates to approach the GS variance estimates. This can be explained by the fact that heavy subsampling of the initial systematic sample causes the subsamples to become less systematic and to resemble simple random samples, which provides the optimal performance of the SRS variance estimator.

The variance estimation associated with the MA framework is most reliable when the assisting model relies on exogenous auxiliary data, i.e. data collected outside the target population (Särndal et al. 1992). Kangas et al. (2016) reported that the use of samples from the target population in model fitting may result

in an underestimation of variance, especially if the assisting model is not sufficiently smooth. In the context of this study, underestimation of variance is problematic as it may lead to a situation where a suboptimal number of field plots are used in the inventory. In this study, we used exogenous training data for the variable selection in the MA estimation, which reduces the potential risk of variance underestimation that could result from the automatized variable selection in the endogenous data. We also adjusted the variance estimators of the MA estimates for the loss of degrees of freedom that originates from the estimation of the assisting model (Lundström and Särndal 1999). In most practical applications, exogenous data are not available for the purposes of model training. In such cases, the selection of predictor variables for simple regression models of timber-related forest attributes can be carried out based on expert knowledge, and too complex models should be avoided.

The use of UAV data can result in a considerable reduction in the number of field plots: the MA estimator with 60 field plots produced, on average, slightly better precision than the EXP estimator with 160 field plots (mean RE: 1.1). Future studies should investigate the optimal selection of field plot locations, especially with the low sampling intensities. It is worth noting that the systematic sample of 160 plots limited the number of unique combinations of field plots in the randomized subsampling procedure. This accentuates the finite population effect of limited sampling units and produces optimistic reporting of the variation among the 100 iterations especially at the highest sampling intensities. A potential strategy to search for optimal locations of sampling units at each sampling intensity level could include, e.g., the pivotal method that selects spatially balanced samples in the feature space of auxiliary information (Grafström et al. 2012). Another option could be to ensure that the feature space of sampling units covers the feature space of the finite population to be estimated (Queinnec et al. 2021). An important remark on the abovementioned techniques is that they require UAV data or other auxiliary information in the planning stage of a sampling design.

The findings of this study highlight the potential of UAV data in the design-based inference of timber value for forest properties larger than a few hectares. The estimation of timber value differs from forest management inventories because the appraisal of timber value must be carried out at short notice. For these reasons, the only viable remote sensing assisted approach is to utilize UAV sensor data in the estimation of timber value. While the estimation framework described in this study includes the collection of field data, which typically accounts for the largest portion of inventory costs, the use of UAV data can enable a reduction in field work, which makes the inventory considerably less expensive when compared to a field inventory only. As an example, the reduction of 100 field plots corresponds to a saving of 8–10 field workdays for an experienced field crew of two people in our study area.

Conclusions

We evaluated the utility of UAV in the design-based estimation of timber value for a 40-ha forest property. We draw the following conclusions based on our findings:

- (1) The SRS estimator produced the largest variance estimates for the field data-based timber value estimates, while the Grafström–Schelin (GS) estimator, with a Voronoi tessellation-based neighborhood selection, produced larger variance estimates than the Matérn estimator.

- (2) The GS and SRS variance estimators produced similar variance estimates in the MA framework, which would indicate that the use of an assisting model and its residuals could effectively mitigate the appearance of spatial autocorrelation evident among the field plots.
- (3) The replacement of a field data-based estimator with the model-assisted (MA) estimator, which employed both field data and 3D UAV data, enabled a reduction in the sampling intensity of the systematic sample (from 160 plots to 60 plots) without loss of estimator efficiency.

Author contributions

Janne Rätty (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing—original draft, Writing—review & editing), Juha Heikkinen (Conceptualization, Methodology, Software, Writing—review & editing), Mikko Kukkonen (Data curation, Software, Writing—original draft, Writing—review & editing), Lauri Mehtätalo (Conceptualization, Writing—review & editing), Annika Kangas (Conceptualization, Writing—review & editing), and Petteri Packalen (Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing—review & editing)

Conflict of interest

None declared.

Funding

The study was supported by the Research Council of Finland through the Finnish Flagship Programme for the Forest-Human-Machine Interplay—Building Resilience, Redefining Value Networks and Enabling Meaningful Experiences (UNITE) [337655], and the projects Unmanned Aerial Vehicles in Forest Remote Sensing [323484] and Asynchronous Datasets in Large-Area Forest Inventories by Remote Sensing [355267].

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

- Axelsson P. DEM generation from laser scanner data using adaptive TIN models. *Int Arch Photogramm Remote Sens* 2000;**33**:110–117.
- Bolles RC, Baker HH and Marimont DH. Epipolar-plane image analysis: an approach to determining structure from motion. *Int J Comput Vis* 1987;**1**:7–55. <https://doi.org/10.1007/BF00128525>.
- Breidt FJ and Opsomer JD. Model-assisted survey estimation with modern prediction techniques. *Stat Sci* 2017;**32**:190–205. <https://doi.org/10.1214/16-ST5589>.
- Dunn R and Harrison AR. Two-dimensional systematic sampling of land use. *J R Stat Soc Ser C Appl Stat* 2018;**42**:585–601. <https://doi.org/10.2307/2986177>.
- Erikäinen K and Venho M. Deriving market prices for forest-land properties from comparables. *Prop Manag* 2018;**36**:423–445. <https://doi.org/10.1108/PM-07-2017-0043>.
- Grafström A, Lundström NLP and Schelin L. Spatially balanced sampling through the pivotal method. *Biometrics* 2012;**68**:514–520.
- Grafström A and Schelin L. How to select representative samples. *Scand J Stat* 2014;**41**:277–290. <https://doi.org/10.1111/sjso.12016>.

- Hao Y, Widagdo FRA, Liu X et al. Estimation and calibration of stem diameter distribution using UAV laser scanning data: a case study for larch (*Larix olgensis*) forests in Northeast China. *Remote Sens Environ* 2022;**268**:112769. <https://doi.org/10.1016/j.rse.2021.112769>.
- Harris AB, Singleton CN and Straka T. Land value differentials resulting from variability between the sales comparison and income approaches in timberland valuation. *Appraisal J* 2018;**86**:192–205.
- Heikkinen J. Assessment of uncertainty in spatially systematic sampling. In Kangas A, Maltamo M. (eds). *Forest Inventory: Methodology and Applications*. Springer, Netherlands, Dordrecht, 2006, pp. 155–176.
- Kangas A, Heikkinen E and Maltamo M. Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. *Can J For Res* 2004;**34**:916–930. <https://doi.org/10.1139/x03-266>.
- Kangas A, Myllymäki M, Gobakken T et al. Model-assisted forest inventory with parametric, semiparametric, and nonparametric models. *Can J For Res* 2016;**46**:855–868. <https://doi.org/10.1139/cjfr-2015-0504>.
- Karjalainen T, Packalen P, Rätty J et al. Predicting factual sawlog volumes in Scots pine dominated forests using airborne laser scanning data. *Silva Fenn* 2019;**53**:4, 1–17. <https://doi.org/10.14214/sf.10183>.
- Kirkpatrick S, Gelatt CD and Vecchi MP. Optimization by simulated annealing. *Science* 1983;**220**:671–680. <https://doi.org/10.1126/science.220.4598.671>.
- Korhonen L, Peuhkurinen J, Malinen J et al. The use of airborne laser scanning to estimate sawlog volumes. *Forestry (Lond)* 2008;**81**:499–510. <https://doi.org/10.1093/forestry/cpn018>.
- Kotivuori E, Kukkonen M, Mehtätalo L et al. Forest inventories for small areas using drone imagery without in-situ field measurements. *Remote Sens Environ* 2020;**237**:111404. <https://doi.org/10.1016/j.rse.2019.111404>.
- Kukkonen M, Kotivuori E, Maltamo M et al. Volumes by tree species can be predicted using photogrammetric UAS data, Sentinel-2 images and prior field measurements. *Silva Fenn* 2021;**55**:1, 1–15. <https://doi.org/10.14214/sf.10360>.
- Laasasenaho J. Taper curve and volume functions for pine, spruce and birch. *Communicationes Instituti Forestalia Fennica* 1982;**108**:1–74.
- Langsæter A. Nøiaktigheten ved linjetaksering av skog. I. *Meddelelser fra det norske Skogforsøksvesen* 1932;**4**:431–563.
- Lappi J. *Metsäbiometrian menetelmiä*. 12th edn. University of Joensuu, Faculty of Forest Sciences, 1993.
- Lindeberg JW. Calculating the standard error in the strip-survey results. *Acta Forestalia Fennica*, 1923;**25**:5, 1–22.
- Lundström S and Särndal C-E. Calibration as a standard method for treatment of nonresponse. *J Off Stat* 1999;**15**:305.
- Luonnonvarakeskus. Teollisuuspuun kauppa. marraskuu 2022; (in Finnish). Available at <https://www.luke.fi/fi/tilastot/teollisuuspuun-kauppa/teollisuuspuun-kauppa-marraskuu-2022>, Accessed 28 Dec 2022.
- Magnussen S and Fehrmann L. In search of a variance estimator for systematic sampling. *Scand J For Res* 2019;**34**:300–312. <https://doi.org/10.1080/02827581.2019.1599063>.
- Magnussen S, McRoberts RE, Breidenbach J et al. Comparison of estimators of variance for forest inventories with systematic sampling - results from artificial populations. *For Ecosyst* 2020;**7**:1–19. <https://doi.org/10.1186/s40663-020-00223-6>.
- Maltamo M and Packalen P. Species-specific management inventory in Finland. In Maltamo M, Næsset E, Vauhkonen J (eds). *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies, Managing Forest Ecosystems*. Springer, Netherlands, Dordrecht, 2014, pp. 241–252.
- Matérn B. Methods of estimating the accuracy of line and sample plot surveys. *Meddelanden från Statens Skogsforskningsinstitut* 1947;**36**:118–136.
- MicaSense. 2020 Radiometric calibration model for MicaSense sensors. available at <https://support.micasense.com/hc/en-us/articles/115000351194-Radiometric-Calibration-Model-for-MicaSense-Sensors>, Accessed 20 Feb 2020.
- Mostafa SA and Ahmad IA. Recent developments in systematic sampling: a review. *J Stat Theory Pract* 2018;**12**:290–310. <https://doi.org/10.1080/15598608.2017.1353456>.
- National Land Survey of Finland. 2022 Laser scanning data 0,5 p. Available at <https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/laser-scanning-data-05-p>, Accessed Sept 2022.
- Packalén P, Suvanto A and Maltamo M. A two stage method to estimate species-specific growing stock. *Photogramm Eng Remote Sens* 2009;**75**:1451–1460. <https://doi.org/10.14358/pers.75.12.1451>.
- Packalén P, Temesgen H and Maltamo M. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. *Can J Remote Sens* 2012;**38**:557–569. <https://doi.org/10.5589/m12-046>.
- Pearce DW. The economic value of forest ecosystems. *Ecosystem Health* 2001;**7**:284–296. <https://doi.org/10.1046/j.1526-0992.2001.01037.x>.
- Peuhkurinen J, Maltamo M and Malinen J. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: a distribution-based approach. *Silva Fenn* 2008;**42**:4, 625–641. <https://doi.org/10.14214/sf.237>.
- Puliti S, Dash JP, Watt MS et al. A comparison of UAV laser scanning, photogrammetry and airborne laser scanning for precision inventory of small-forest properties. *Forestry* 2020;**93**:150–162. <https://doi.org/10.1093/forestry/cpz057>.
- Queinnec M, Coops NC, White JC et al. Developing a forest inventory approach using airborne single photon lidar data: from ground plot selection to forest attribute prediction. *Forestry* 2021;**95**:347–362. <https://doi.org/10.1093/forestry/cpab051>.
- R Core Team. R: *A Language and Environment for Statistical Computing*. 2023; R Foundation for Statistical Computing, Vienna, Austria.
- Rätty J, Astrup R and Breidenbach J. Prediction and model-assisted estimation of diameter distributions using Norwegian national forest inventory and airborne laser scanning data. *Can J For Res* 2021;**51**:1521–1533. <https://doi.org/10.1139/cjfr-2020-0440>.
- Rätty M, Kuronen M, Myllymäki M et al. Comparison of the local pivotal method and systematic sampling for national forest inventories. *For Ecosyst* 2020;**7**:54. <https://doi.org/10.1186/s40663-020-00266-9>.
- Salminen S. *Tulosten Luotettavuus Ja Karttatulostus Valtakunnan Metsien V Inventoinnissa [Reliability of the Results from the Fifth National Forest Inventory and a Presentation of an Output-Mapping Technique]*.1973; Metsäntutkimuslaitos, Helsinki.
- Särndal C-E, Swensson B and Wretman J. *Model Assisted Survey Sampling*.1992; Springer New York, NY: Springer Science & Business Media.
- Stevens DL and Olsen AR. Spatially balanced sampling of natural resources. *J Am Stat Assoc* 2004;**99**:262–278. <https://doi.org/10.1198/016214504000000250>.

Turner R. Deldir: Delaunay triangulation and Dirichlet (Voronoi) tessellation. R package version 1.0-6. 2021. Available at <https://CRAN.R-project.org/package=deldir>.

Vähä-Konka V, Maltamo M, Pukkala T *et al.* Evaluating the accuracy of ALS-based removal estimates against actual logging data. *Ann For Sci* 2020;**77**:84. <https://doi.org/10.1007/s13595-020-00985-7>.