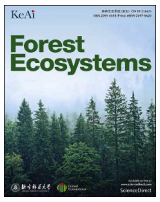
Forest Ecosystems 15 (2026) 100428

Contents lists available at ScienceDirect

Forest Ecosystems

journal homepage: www.keaipublishing.com/cn/journals/forest-ecosystems



New and established regression techniques to address design-bias trends in fixed populations

Magnus Ekström*, Christoffer Axelsson, Göran Ståhl

Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå 901 83, Sweden

ARTICLE INFO

Keywords:
Bias
Design-based inference
Model-based inference
Regression analysis

ABSTRACT

In many studies applying remotely sensed data and regression analysis for assessing ecosystem characteristics, such as biomass or growing stock volume in forests, a trend from over-predicting small true values to under-predicting large true values is observed. The reason for this trend often remains elusive, but it can be shown that it is a direct consequence of, deliberately or by mistake, adopting a design-based inference perspective when evaluating the results from model-based predictions. However, the design-bias trend is problematic in many applications, because the real conditions within the ecosystem studied will not be correctly determined. Instead, predictions tend to be shrunk towards the mean value of the target variable in the sample data used for estimating the parameters of the prediction model. Thus, calibration techniques to mitigate the design-bias trend have been proposed by some authors. In this article, we evaluate various regression techniques with respect to bias. The method of evaluation is founded on design-based inference, and thus, with regard to terminology, the regression techniques are used for estimating fixed quantities at the level of population elements rather than for predicting random quantities, as in the case of model-based inference. With aerial laser scanning data or digital aerial photographs, standard ordinary least squares (OLS) regression combined with classical calibration (CC) and the new MAVGAR method performed best in terms of bias, and produced good or reasonably good root mean square error (RMSE) values. The MAVGAR method aims to minimize the mean of the absolute values of groupwise average residuals, which is the origin of its name. None of the evaluated methods performed well in producing estimates with low bias when optical satellite data were used.

1. Introduction

Regression methods, which in a broad sense include many machine learning techniques, are widely applied as part of procedures for assessing state and change of natural and managed ecosystems based on remotely sensed data. Important examples include habitat conditions for different species (Estes et al., 2010), structural features of forest ecosystems (Dubayah et al., 2022), and yields of agricultural crops (Li et al., 2007). In this article, we focus on continuous variables used for quantifying habitat structures or crops, such as biomass, deadwood, or growing stock volume in forests. Predictions of such variables may be used either for mapping (Nilsson et al., 2017) or for assessing means and totals across large areas (Chen et al., 2016).

However, many analysts that use regression models for assessing ecosystem characteristics based on remotely sensed (or other) data have been surprised to find that, when plotting predicted values versus observed values, a trend from over-predicting small observed values to under-predicting large observed values is found (Persson and Ståhl, 2020). The reason for this trend often remains elusive, especially because traditional regression methods and modern machine-learning methods are known to produce approximately model-unbiased predictions. However, as argued in Ståhl et al. (2024), the trend occurs when analysts adopt a design-based perspective in evaluating the model-based predictions. Whereas model-unbiasedness should theoretically be evaluated across an infinite number of realized populations from a super-population model (Cassel et al., 1977), empirical evaluations of prediction accuracy mostly use data from the only population realisation at hand, i.e., from the real world. In doing so, analysts must carefully distinguish model-bias from design-bias and adjust their evaluations accordingly.

However, if a design-based inference perspective is adopted (Gregoire, 1998), standard model-based techniques such as ordinary least

* Corresponding author.
E-mail address: Magnus.Ekstrom@slu.se (M. Ekström).
Peer review under the responsibility of Editorial Office of Forest Ecosystems.

https://doi.org/10.1016/j.fecs.2026.100428
Received 21 August 2025; Received in revised form 9 January 2026; Accepted 16 January 2026

squares (OLS) regression (Chatterjee and Hadi, 2013) produce design-biased estimates at the level of population elements (Ståhl et al., 2024). Note the distinction between the terms estimation and prediction in this article, following standard use of the terms in the statistical literature. We use the term estimation for assessing fixed quantities, such as values of the variable of interest for individual population elements in design-based inference. We use the term prediction when the quantity is a random variable, such as the value of the variable of interest for population elements in model-based inference.

Design-bias trends are problematic in several applications. With repeated modelling using a certain type of remotely sensed data (e.g., optical data from a satellite-borne sensor) in populations that remain relatively stable across time, such as many forests, a similar magnitude and the same sign of error will be obtained repeatedly for a given population element (Ehlers et al., 2018). In mapping landscapes based on characteristics assessed on a continuous scale, the maps will not display the real variability within the landscape but smaller variability (Kangas et al., 2023). If data are applied in planning or scenario modelling systems, design-bias trends may cause distorted analysis results (Barth et al., 2009). Further, extreme values, like biodiversity hotspots, are not likely to be detected.

Due to the problems encountered in applications, some previous studies have proposed calibration methods for reducing design-bias trends from estimates obtained using model-based techniques based on remotely sensed data. Gilichinsky et al. (2012) used histogram matching (HM), ensuring that estimated values would have approximately the same distribution as a sample from the population. Lindgren et al. (2022a, b) used classical calibration (CC) for modifying their model-based predictions, following methods devised in chemometrics (Shukla, 1972). These methods use standard model-based techniques in a first step and subsequently make modifications.

In design-based inference, the only source of randomness is the sample selection process itself. The population values are considered fixed and non-random, and the randomness comes from how the samples are selected (e.g., simple random sampling, stratified sampling). Therefore, the statistical inferences rely on the probabilities associated with the design of the sample selection procedure. Evaluating estimators in design-based inference using Monte Carlo simulation involves mimicking the sample selection process repeatedly to see how well the estimators perform in terms of bias, variance, mean squared error (MSE), etc. Simulation is especially useful when theoretical derivations are complex or impractical, providing a practical way to approximate the sampling distribution and assess estimator properties.

In contrast, model-based inference uses Monte Carlo simulation to replicate the model's data generation process (not just the sampling) in order to evaluate estimators. When assessing the performance of a procedure that produces a map with estimates of growing stock volume or biomass over an area of interest, it seems more natural to use design-based inference—that is, by evaluating the accuracy of the estimated biomass or growing stock volumes across repeated samples of the population elements. This is because the forest on any given day will look largely the same the next day, and even the next month. A new forest landscape is not repeatedly generated over a few years; therefore, model-based inference is less suitable in this context.

The objective of this study was to evaluate standard methods, methods proposed for mitigating design-bias trends, and two new methods in terms of how effectively they reduce bias trends without causing a large increase in root mean square error (RMSE), compared to OLS regression, which is used as a baseline method. Another objective was to assess how the strength of the association between the target variable and the covariates affects the results. The evaluations were carried out using Monte Carlo simulations under a simple random sampling design.

2. Materials and methods

2.1. Reference and remote sensing data

The data used in the study consist of reference data in the form of growing stock volumes from National Forest Inventory (NFI) plots and three types of remote sensing products: airborne laser scanning (ALS), digital surface model (DSM) derived from aerial imagery, and satellite imagery. We used both permanent and temporary NFI plots inventoried between 2010 and 2023 from a 300 km × 300 km area in northern Sweden (roughly 63°–65.7° N, 17.1°–22.5° E). Using the collected NFI field data, growing stock volumes were calculated with the Heureka software (Wikström et al., 2011).

The temporary plots were only inventoried at a single point in time, and we only matched them to remote sensing data captured in that year. The permanent plots were, however, visited every five years, and we interpolated volumes for intermediate years to get reference volumes for all years. Plot volumes were not interpolated if the volume had declined by more than 5 $m^3 \cdot ha^{-1}$, indicating some type of disturbance during the time between field inventories.

The ALS data were collected by the National Mapping Agency in Sweden. The first campaign (2009–2013) scanned with a point density of 0.5–1 pulses·m^{-2} while the second campaign (2018–2021) had a point density of 1–2 pulses·m^{-2}. The DSM data were produced by the Swedish National Mapping Agency and contain vegetation heights in raster format. These were created using aerial image matching from overlapping imagery. The pixel resolution of the data was 1.0 m for 2018 and 0.4 m for 2022. The satellite imagery was selected with the aim of finding cloud-free scenes from the summer period. All the sentinel imagery had been atmospherically corrected (L2A version). From each of the remote sensing datasets, we extracted features to be evaluated as covariates of growing stock volume. The two features yielding the lowest RMSE were selected using stepwise linear regression with forward selection (leapForward method from the caret package; Kuhn, 2022).

2.2. Design-based assessment of bias and mean squared error of an estimator at the grid-cell level

A sample is selected from a population consisting of N elements. For the elements in the sample, values of the variable of interest as well as one or more covariates are recorded, i.e., we observe $(x_{1i},\ldots,x_{pi},y_i)$, $i = 1,\ldots,n$, where n is the sample size. It is assumed that the values of all covariates $x_1,\ldots,x_p$ are known for the entire population; for example, the remote sensing metrics for all the grid cells in a study area. Any transformations of the basic metrics can also be applied and would then be included as separate covariates. Values of the target variable y (e.g., growing stock volume) are usually only known for the population elements included in the sample.

For a given sampling design $p(\cdot)$, we regard any sample s as the outcome of a set-valued random variable S, whose probability distribution is specified by $p(\cdot)$. That is, $p(s)$ is the probability that $S = s$. Let S denote the set of all possible samples s for which $p(s)$ is strictly positive. Based on S and a specified estimation method, an estimator $\hat{y}_i = \hat{y}_i(S)$ of y_i is a random variable, with expected value (E) and variance (V) defined as Eqs. 1 and 2:

$$E(\hat{y}_i) = \sum_{s\in S} p(s)\hat{y}_i(s) \tag{1}$$

$$V(\hat{y}_i) = E\left((\hat{y}_i - E(\hat{y}_i))^2\right) = \sum_{s\in S} p(s)(\hat{y}_i(s) - E(\hat{y}_i))^2 \tag{2}$$

Two important measures of the quality of the estimator $\hat{y}_i$ are its bias (B) and mean squared error (MSE), defined as Eqs. 3 and 4:

$$B(\hat{y}_i) = E(\hat{y}_i) - y_i \tag{3}$$

$$\mathrm{MSE}(\hat{y}_i) = E\left((\hat{y}_i - y_i)^2\right) = \sum_{s \in S} p(s)(\hat{y}_i(s) - y_i)^2 \tag{4}$$

where it can be verified that $\mathrm{MSE}(\hat{y}_i) = V(\hat{y}_i) + \left[B(\hat{y}_i)\right]^2$ (Särndal et al., 1992).

Note that, by the law of total expectation

$$E(\hat{y}_i) = P(i \in S)E(\hat{y}_i | i \in S) + P(i \notin S)E(\hat{y}_i | i \notin S) \tag{5}$$

and that this expression must equal y_i for $\hat{y}_i$ to be unbiased. For this equality to hold, we would require

$$E(\hat{y}_i | i \notin S) = \frac{y_i - P(i \in S)E(\hat{y}_i | i \in S)}{P(i \notin S)} \tag{6}$$

However, the sampling design provides no information about the value of y_i when $i \notin S$, so there is no way to guarantee that the equality in Eq. 6 is fulfilled. Thus, while population-level targets, such as totals and means, can be estimated in an unbiased manner under the sampling design (Särndal et al., 1992), unbiased estimation at the grid-cell level is generally not attainable. Therefore, this paper examines methods that seek to mitigate systematic design-bias trends, rather than attempting to eliminate them entirely, for grid-cell-level quantities. Moreover, for arbitrary populations, y_i can take any value regardless of the covariates $x_{1i}, \ldots, x_{pi}$; therefore, even when all covariates are known for the entire population, unbiasedness for individual units cannot be guaranteed.

Typically, the number of possible samples and the corresponding possible values of $\hat{y}_i$ are very large, which makes the direct computation of the bias, variance, and MSE of an estimator computationally demanding. A practical alternative is to use a Monte Carlo simulation, in which a large number of independent samples $s_1, \ldots, s_m$ are drawn under the same sampling design $p(\cdot)$ to estimate $E(\hat{y}_i)$ and $V(\hat{y}_i)$ as Eqs. 7 and 8:

$$\hat{E}(\hat{y}_i) = \frac{1}{m} \sum_{j=1}^{m} \hat{y}_i(s_j) \tag{7}$$

$$\hat{V}(\hat{y}_i) = \frac{1}{m-1} \sum_{j=1}^{m} \left(\hat{y}_i(s_j) - \hat{E}(\hat{y}_i)\right)^2 \tag{8}$$

from which estimates of the bias and MSE can be derived.

2.3. Estimation methods

We assume that a simple random sample (without replacement) of n elements is selected from a population consisting of N elements. Based on the sample, the goal is to estimate the y-values in the population, which can be achieved by fitting a hyperplane to the observations $(x_{1i}, \ldots, x_{pi}, y_i), i = 1, \ldots, n$. A fitted hyperplane can be written as Eq. 9:

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_{1i} + \ldots + \hat{b}_p x_{pi}, i = 1, \ldots, N \tag{9}$$

where $\hat{b}_0, \ldots, \hat{b}_p$ are parameters estimated from the sample data, and $\hat{y}_i$ is the estimate of y_i, $i = 1, \ldots, N$.

In OLS regression, the parameter estimates $\hat{b}_0, \ldots, \hat{b}_p$, are obtained by minimizing the sum of squared residuals, where a residual is the distance of an observation to the hyperplane, measured parallel to the y-axis. If desired, a generalization of OLS known as weighted least squares (WLS) can be used, where $\hat{b}_0, \ldots, \hat{b}_p$ are obtained by minimizing a sum of weighted squared residuals (Draper and Smith, 1998). Another approach, known as orthogonal regression or total least squares (TLS) regression, follows a similar principle to OLS but fits a hyperplane by minimizing the sum of squared perpendicular distances from the observations to the hyperplane, rather than the sum of squared residuals (Golub and van Loan, 1980).

Noting the tendency of OLS to underestimate large true values and overestimate small ones, Lindgren et al. (2022a, b) suggested that this bias trend can be reduced using CC (Shukla, 1972). In this approach, a set of estimated values $\hat{y}_i$ of y_i, $i = 1, \ldots, N$, are obtained in a first stage using OLS (or, more generally, WLS). Then, a model is fitted using OLS (Eq. 10):

$$\hat{\hat{y}}_i = \hat{a}_0 + \hat{a}_1 y_i \tag{10}$$

where the double-hat notation indicates that this model provides estimates of the previously estimated values $\hat{y}_i$. Rather than thinking of this as estimating estimates, Lindgren et al. (2022a, b) rewrote this equation as

$$\hat{y}_{i,c} = \frac{\hat{y}_i - \hat{a}_0}{\hat{a}_1} \tag{11}$$

where $\hat{y}_{i,c}$ denotes the suggested calibrated estimate of y_i. Thus, a bias-corrected estimate of the true value of the target variable is provided, based on the initial OLS (or WLS) estimate and the parameters of the calibration model.

Gilichinsky et al. (2012) suggested using HM for adjusting a set of estimated values, $\hat{y}_i$, $i = 1, \ldots, N$. Let F_N denote the empirical cumulative distribution function (ECDF) of $\hat{y}_i$, $i = 1, \ldots, N$, and let G_n denote the ECDF of the sample values y_i, $i = 1, \ldots, n$. HM aims to find a transformed value $\hat{y}_{i,\mathrm{HM}}$ for each $\hat{y}_i$ such that $F_N(\hat{y}_i) = G_n\left(\hat{y}_{i,\mathrm{HM}}\right)$. In practice, HM can only approximate this transformation due to the discreteness of empirical distributions. In our study, the initial estimates $\hat{y}_i$, $i = 1, \ldots, N$, are provided by OLS or WLS.

Various non-parametric methods from statistics and machine learning are also available for regression. In this study, we consider one of the best-known non-parametric methods: K-nearest neighbors (KNN) regression (James et al., 2021). Given a value for K and a population element j with covariate vector $(x_{1j}, \ldots, x_{pj})^{\mathrm{T}}$, KNN regression first identifies the K population elements in the sample whose covariate vectors are closest to $(x_{1j}, \ldots, x_{pj})^{\mathrm{T}}$, denoted by N_j. It then estimates y_j as the average of all responses in N_j. A small value of K results in a flexible fit with low bias and high variance. Since we are interested in estimators with low bias, we chose $K = 1$ for our study, and refer to the method as 1NN.

The tendency for large true values to be underestimated and small true values to be overestimated when applying OLS suggests that the variation in the estimated values is too small compared to the corresponding true values. In an effort to correct this bias, we propose a new adjustment to the OLS estimates.

Let $\bar{x}_1, \ldots, \bar{x}_p$, and $\bar{y}$ denote the sample means of the observations $x_{1i}, \ldots, x_{pi}$, and $y_i, i = 1, \ldots, n$, respectively. By Draper and Smith (1998), the OLS estimates of parameters can be written as Eq. 12:

$$\begin{pmatrix} \hat{b}_1 \\ \vdots \\ \hat{b}_p \end{pmatrix} = \boldsymbol{C}_n^{-1} \boldsymbol{c}_n \quad \text{and} \quad \hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}_1 - \ldots - \hat{b}_p \bar{x}_p \tag{12}$$

where $\boldsymbol{C}_n$ denotes the sample covariance matrix of $(x_{1i}, \ldots, x_{pi}), i = 1, \ldots, n$, and $\boldsymbol{c}_n$ is the vector containing the covariances $\mathrm{cov}_n(y, x_1), \ldots, \mathrm{cov}_n(y, x_p)$, where $\mathrm{cov}_n(y, x_j)$ denotes the sample covariance of $y_i, i = 1, \ldots, n$, and $x_{ji}, i = 1, \ldots, n$.

If the following adjusted parameter estimates are used instead of those obtained by OLS, the variance of the N estimated y-values will equal the variance of the y-values in the sample (Eq. 13):

$$\begin{pmatrix} \widehat{b}_{1,\mathrm{adj}} \\ \vdots \\ \widehat{b}_{p,\mathrm{adj}} \end{pmatrix} = \boldsymbol{C}_N^{-1}\boldsymbol{c}_n \sqrt{\frac{\mathrm{var}_n(y)}{\boldsymbol{c}_n^{\mathrm{T}}\boldsymbol{C}_N^{-1}\boldsymbol{c}_n}} \quad \text{and} \quad \widehat{b}_{0,\mathrm{adj}} = \overline{y} - \widehat{b}_{1,\mathrm{adj}}\overline{x}_1 - \ldots - \widehat{b}_{p,\mathrm{adj}}\overline{x}_p \tag{13}$$

where $\boldsymbol{C}_N$ is the covariance matrix of $(x_{1i},\ldots,x_{pi}), i = 1,\ldots,N$, and $\mathrm{var}_n(y)$ is the sample variance of $y_i, i = 1,\ldots,n$. This proposed estimation procedure is referred to as the variance-adjusted OLS (VAOLS) estimator. Like the OLS hyperplane, the VAOLS hyperplane passes through the point $(\overline{x}_1,\ldots,\overline{x}_p,\overline{y})$, but it is scaled to preserve the variance of the target variable.

Instead of adjusting estimates obtained through OLS or other methods, more direct approaches can be used. For example, we can try to construct estimators such that all population elements with target variable values in a neighborhood of y have at most a small mean bias, ensuring that this holds for all neighborhoods. Based on this idea, we divide the range of the observed y-values in a sample into g disjoint subintervals of equal length. For subinterval i, let m_i denote its midpoint. For a window (neighborhood) of size w, $[m_i - w/2, m_i + w/2)$, let n_i denote the number of y-values falling in the window and denote these values by $y_{ij}, j = 1,\ldots,n_i$. For each window or neighborhood, we would like to find values of $\beta_0,\ldots,\beta_p$ that makes the value of $\frac{1}{n_i}\sum_{j=1}^{n_i}\left(y_{ij} - (\beta_0 + \beta_1 x_{1ij} + \ldots + \beta_p x_{pij})\right)$ as close to zero as possible. Since this is desired for each of the g windows, it makes sense to define $\widehat{b}_0,\ldots,\widehat{b}_p$ to be the values of $\beta_0,\ldots,\beta_p$ that minimize the quantity in Eq. 14:

$$\frac{1}{g}\sum_{i=1}^{g}\left|\frac{1}{n_i}\sum_{j=1}^{n_i}\left(y_{ij} - (\beta_0 + \beta_1 x_{1ij} + \ldots + \beta_p x_{pij})\right)\right| \tag{14}$$

This proposed estimation method is referred to as the MAVGAR estimator, as it minimizes the mean of the absolute values of groupwise average residuals. Although the approach is designed to reduce systematic errors across subpopulations defined by local windows of the target variable, the objective function is averaged over the g windows in Eq. 14, yielding a single global estimator of the parameter vector.

If desired, the window size w can be set equal to the length of a subinterval. Then, for some subintervals, the value of n_i can become very small, and a limitation of MAVGAR is its potential sensitivity to windows containing only a few observations. This situation may arise when such windows include observations that do not follow the general pattern, allowing them to exert an unduly large influence on the estimated parameter vector, particularly if they contain extreme covariate values. For this reason, it may be advantageous to select a value of w larger than the length of a subinterval. Another possibility is to use cross validation for selecting a suitable value of w.

2.4. What is estimated by OLS?

When constructing a map of estimated growing stock volume for an area of interest, it is desirable that the estimates are as close as possible to the true values. As previously mentioned, we consider OLS regression to be our baseline method. To understand what it estimates, we define a general regression superpopulation model with the following properties:

i) $y_1,\ldots,y_N$ are realized values of the random variables $Y_1,\ldots,Y_N$;
ii) $E(Y_i) = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}, i = 1,\ldots,N$, and
iii) $V(Y_i) = \sigma_i^2$ and $\mathrm{Cov}(Y_i, Y_j) = \sigma_{ij}, i,j = 1,\ldots,N$.

Let $\boldsymbol{y}_n$ be the vector of observed y_i values, $i = 1,\ldots,n$, and let $\boldsymbol{X}_n$ be the corresponding design matrix. The first column of $\boldsymbol{X}_n$ consists of ones (to account for the intercept), and column $j+1, j = 1,\ldots,p$, contains the sample values of covariate $x_{ji}, i = 1,\ldots,n$. Assuming, hypothetically, that a complete enumeration of the finite population is available, let $\boldsymbol{y}_N$ and $\boldsymbol{X}_N$ be the corresponding vector and design matrix constructed from $(x_{1i},\ldots,x_{pi},\ y_i), i = 1,\ldots,N$.

Based on a simple random sample $(x_{1i},\ldots,x_{pi},\ y_i), i = 1,\ldots,n$, a more commonly used computational equation for the OLS estimator than Eq. 12 is Eq. 15:

$$\widehat{\boldsymbol{B}} = \left(\boldsymbol{X}_n\boldsymbol{X}_n^{\mathrm{T}}\right)^{-1}\boldsymbol{X}_n\boldsymbol{y}_n \tag{15}$$

If, instead of observing a simple random sample of size n from the N population elements, a complete enumeration of the finite population had been available—i.e., if all the values of $(x_{1i},\ldots,x_{pi},y_i), i = 1,\ldots,N$, were known—then the finite population regression parameter vector $\boldsymbol{B} = \left(\boldsymbol{X}_N\boldsymbol{X}_N^{\mathrm{T}}\right)^{-1}\boldsymbol{X}_N\boldsymbol{y}_N$ could have been computed. Under repeated sampling of the finite population using simple random samples of size n, $\widehat{\boldsymbol{B}}$ is a biased estimator of $\boldsymbol{B}$, but is approximately unbiased for sufficiently large n (Särndal et al., 1992). For n large enough, this implies that $\boldsymbol{x}^{\mathrm{T}}\widehat{\boldsymbol{B}}$ is an approximately unbiased estimator of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}$, where $\boldsymbol{x} = (1, x_1, \ldots, x_p)^{\mathrm{T}}$.

If $\sigma_i^2 = \sigma^2$ for all $i = 1,\ldots,N$, and $\sigma_{ij} = 0$ for all $i \neq j$, then by the Gauss–Markov theorem, $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}$ is the best linear unbiased estimator of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_0,\ldots,\beta_p)^{\mathrm{T}}$ (Greene, 2011). If, in addition, the covariates satisfy the so-called Grenander conditions for well-behaved data, then $\sqrt{N}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta})$ is asymptotically normally distributed with mean zero. Thus, under the superpopulation model and for large populations, $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}$ is very close to $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}$ with high probability, and the OLS estimator $\boldsymbol{x}^{\mathrm{T}}\widehat{\boldsymbol{B}}$ can therefore be viewed as an estimator not only of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}$ but also of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}$ (Heeringa et al., 2010).

The discussion above assumed a constant variance σ^2 and zero covariances for all Y_i and Y_j, where $i \neq j$. However, the conclusion remains valid in more general cases. For example, as noted in Greene (2011), if the covariates are sufficiently well behaved and the covariance σ_{ij} diminishes sufficiently rapidly as the distances between the corresponding grid cells increase, then $\sqrt{N}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta})$ is asymptotically normally distributed with mean zero. Thus, in this more general case as well, the OLS estimator $\boldsymbol{x}^{\mathrm{T}}\widehat{\boldsymbol{B}}$ can be regarded as an estimator of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}$.

A potentially more efficient estimator can be obtained using WLS, where one attempts to estimate the non-constant variances σ_i^2 and subsequently incorporates these estimates when defining the estimator of the parameter vector. WLS is, in turn, a special case of generalized least squares (GLS), which can handle situations where the covariances σ_{ij} are non-zero for $i \neq j$. Typically, when applying GLS, both the variances σ_i^2 and the covariances σ_{ij} must be estimated. Provided these estimates are consistent, the resulting parameter estimator is asymptotically more efficient than the OLS estimator (Greene, 2011). However, for small to medium-sized samples, GLS can actually be less efficient than OLS. In this paper, we do not consider GLS.

When constructing a map of estimated growing stock volume for an area of interest, it is desirable that the estimates are as close as possible to the true values. In grid cell i, with vector $\boldsymbol{x}_i = (1, x_{1i}, \ldots, x_{pi})^{\mathrm{T}}$, a quantity being approximately unbiasedly estimated is $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{B}$ from a purely design-based perspective, or $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}$ from a superpopulation perspective. If the association between the y-values and the covariates is strong, this will typically also be a good estimate of the true value. However, if the association is weak, then for many grid cells i, the OLS estimate $\boldsymbol{x}_i^{\mathrm{T}}\widehat{\boldsymbol{B}}$, as well as $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{B}$ and $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta}$, can deviate substantially from the corresponding true value y_i.

2.5. Monte Carlo simulation setup

For each dataset (i.e., pseudo-population), two covariates were selected to simplify the analysis and facilitate comparisons across estimation methods. The covariates were chosen using stepwise linear

regression with forward selection (leapForward method from the caret package; Kuhn, 2022) in an OLS regression setting, based on all data in the pseudo-population. They are presented in Table 1. Fig. 1 shows growing stock volume plotted against the selected covariates for each pseudo-population considered.

For each pseudo population—with known values of growing stock volume y_i, $i = 1,\ldots,N$, and corresponding known covariate data $(x_{1i},\ldots, x_{pi}), i = 1,\ldots,N$, obtained from remote sensing—10,000 simple random samples (without replacement) of n elements were drawn. For each pseudo-population, n was set to 20% of N. For each sample, estimates of $y_i, i = 1,\ldots,N$, were computed using the estimation methods previously described.

To perform OLS regression, we used the standard lm function in R (R Core Team, 2025). The gls function from the nlme package allows for the specification of weights in WLS using various variance functions (Pinheiro et al., 2021). For some models, we used the fitted (i.e., estimated) y-values as the variance covariate with the varPower function. For other models, individual covariates were used as inputs to the varPower function and then combined using the varComb function.

For KNN regression, we used the knnreg function from the caret package (Kuhn, 2022). For TLS regression, we used the odregress function from the pracma package (Borchers, 2023). For HM, we used the histMatch function from the RStoolbox package (Richards, 2022).

For our MAVGAR method, we set $g = 100$. The window size w was either fixed at 100 or selected from the set $\{20, 40, \ldots, 200\}$ as the value that minimized the 5-fold cross-validated (CV) MSE. Details on the algorithm used to compute k-fold CV MSEs can be found in James et al. (2021).

For each sample and estimation method, estimates of $y_i, i = 1,\ldots,N$, were computed. Then, for each method and population element, estimates of bias and MSE were calculated. To evaluate the estimates of growing stock volume, the population elements were divided into five groups: Group 1 consisted of the 20% of elements with the lowest growing stock volumes; Group 2 included the next lowest 20%, and so on, with Group 5 consisting of the 20% of elements with the highest volumes. Within each group, the average bias and average MSE were computed over the population elements in that group. The bias for a group was defined as this average bias, and the RMSE was defined as the square root of the average MSE. As a reference, average growing stock volumes within groups are provided in Appendix A.

For estimation methods such as OLS, the estimated y-values may be negative. Therefore, we considered two subcases: One in which negative estimated values were accepted, and another in which such values were set to zero.

3. Results

In this section, we mainly focus on one ALS dataset (from 2019), the DSM data, and one satellite dataset (Sentinel-2 from 2018), considering the case where negative volume estimates were set to zero. This choice reflects standard practice. Tables with results for the other datasets under this condition are provided in Appendix B. The corresponding tables for the case where negative volume estimates were retained are presented in Appendix C.

For the ALS data from 2019, MAVGAR, both with and without CV, generally performed best in terms of bias, although OLS with CC was not far behind (Table 2). In terms of RMSE, OLS regression was the overall winner, followed by OLS with CC and VAOLS. MAVGAR also performed quite well. Similar conclusions hold for the case where negative volume estimates were retained, although OLS with CC exhibited large bias in Group 1 in that scenario (Appendix C), especially relative to the average growing stock volume in that group (Appendix A). The results for the ALS data from 2010 were also broadly consistent (Appendices B and C). Due to numerical issues, both variants of the method for computing weights in WLS failed for some of the samples; therefore, we did not present any results for WLS in Table 2.

For the DSM data, the results obtained with WLS were very similar to those from OLS. For this reason, we do not provide specific comments on WLS. MAVGAR, both with and without CV, generally performed best in terms of bias, although OLS with CC was not far behind (Table 3 and Appendix C). In terms of RMSE, OLS and VAOLS yielded the best results, followed by OLS with CC and MAVGAR with CV.

For the Sentinel-2 data from 2018, where the association between growing stock volume and the covariates was weaker than for ALS and DSM (Table 1), the biases were often large or quite large for all methods. For example, for OLS, the average bias in Group 5 was -74.7 $m^3 \cdot ha^{-1}$ (Table 4), while the average growing stock volume in this group was 266.2 $m^3 \cdot ha^{-1}$ (Appendix A). Thus, on average, OLS underestimated volume by approximately 28% in that group. The best-performing methods were OLS with CC and MAVGAR with CV (Table 4). For example, in Group 5, they underestimated volume by about 6.9% and 4.8% on average, respectively. In terms of RMSE, OLS and VAOLS performed best. In comparison, OLS with CC showed relatively poor performance in Groups 3 and 4. MAVGAR, especially without CV, showed high RMSEs across all groups, with only TLS performing worse. The results for the other three satellite datasets were also broadly consistent (Appendix B), except that it was difficult to identify a winner in terms of bias for the Landsat 8 data.

Because of numerical instability in computing WLS weights, neither of the two methods succeeded for all samples; therefore, WLS results were not included in Table 4. However, results for WLS were included in the tables for the other three satellite datasets (Appendices B and C), and in all cases, these results were very similar to the corresponding ones obtained using OLS.

In general, TLS and 1NN showed relatively poor performance, both in terms of bias and RMSE. In the latter case, this might be due to the rather small sample sizes. WLS produced results that were very close to those of OLS. Of the three OLS-adjusted methods—OLS with CC, OLS with HM, and VAOLS—all generally showed a decrease in bias but an increase in RMSE, with the first performing best in terms of bias and the third performing best with respect to RMSE. If a reduction in bias relative to OLS is desired, the simulations suggest that MAVGAR (with or without CV) or OLS with CC are preferable, although the former may not be suitable if the association between the target variable and the covariates is too weak. For example, for reducing bias in Group 5, the domain containing the largest true growing stock volumes, MAVGAR without CV seemed to be the most efficient (Tables 2–4, Appendices B and C). However, when the association between the target variable and the covariates was weak, the corresponding cost in terms of increased RMSE was high, not only for the domain containing the largest true growing stock volumes, but also for other domains. If, in addition, there were observations that did not follow the general pattern between the target

Table 1
Features selected as covariates of growing stock volume. R-squared values were obtained by fitting a linear model to the N observations using OLS regression with the selected covariates.

Data	Year	N	Selected covariates	R-squared
ALS	2010	456	MeanSqH[†], vegQuota[††]	0.86
ALS	2019	758	MeanSqH, vegQuotaFirst[‡]	0.87
DSM	2018	632	MeanHeight, $q90^2$	0.78
LANDSAT8	2014	552	$Band6^2$, $Band6^{0.5}$	0.45
SENTINEL2	2015	1339	Band5, $Band5^{0.5}$	0.50
SENTINEL2	2018	1496	log(Band11)/Band11, log (Band12) × Band11	0.56
SENTINEL2	2019	1183	log(Band2) × Band5, log(Band11)/ Band11	0.54

Note: [†]MeanSqH: average of squared heights; [††]vegQuota: ratio of returns above 1.5 m (measures vegetation density); [‡]vegQuotaFirst: ratio of first returns above 1.5 m.

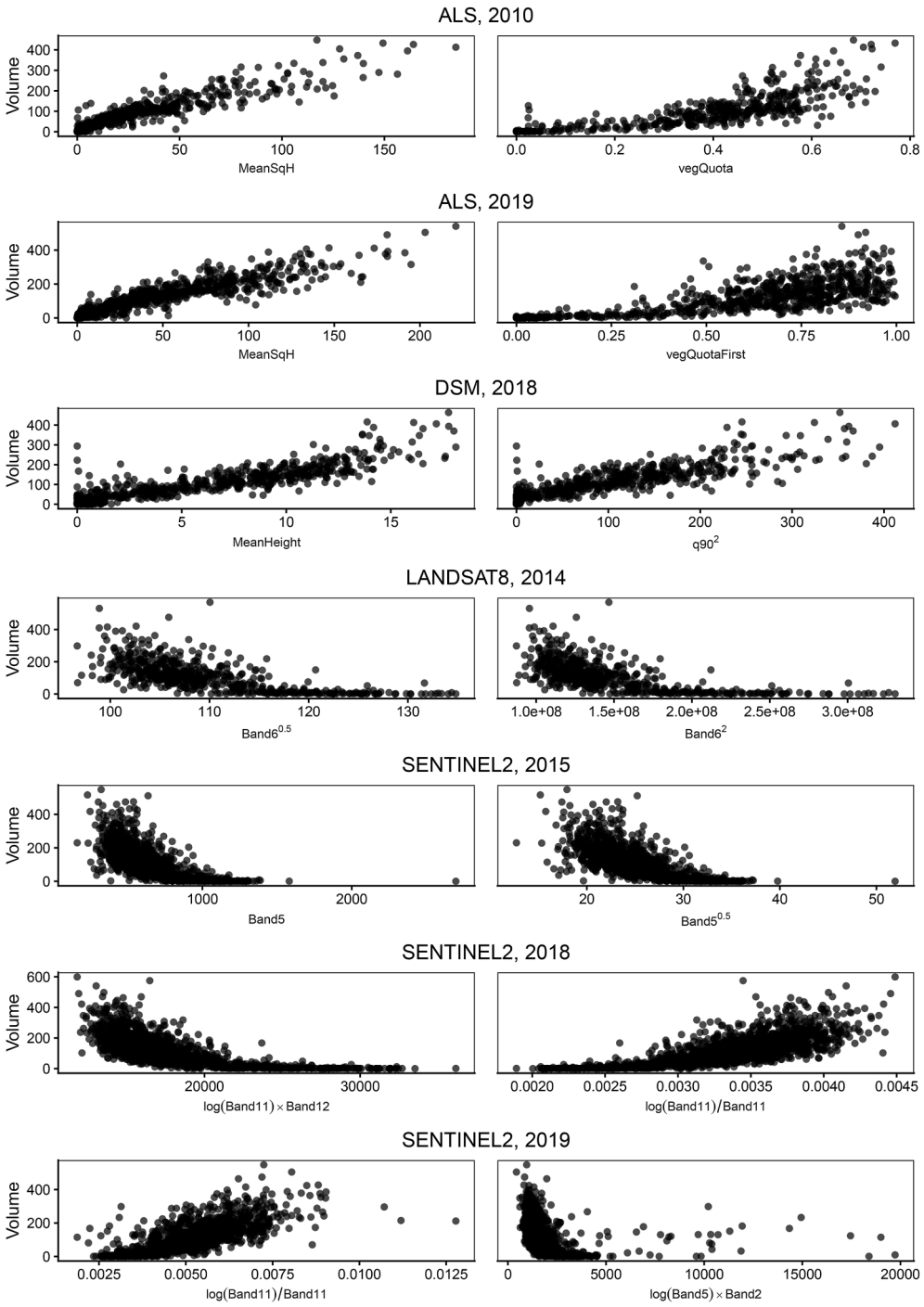


Fig. 1. Growing stock volume plotted against the selected covariates for each pseudo-population.

variable and the covariates, this cost was even higher, which was most notable for the Sentinel-2 data from 2019 (Fig. 1; Appendices B and C).

4. Discussion

The MSE, or its square root, the RMSE, is commonly used to assess the performance of different estimators. However, sometimes it is also necessary to consider bias. When constructing a map of, for example, estimated growing stock volume for an area of interest, it is desirable that the estimates are as close as possible to the true volumes and that the estimation process does not introduce any systematic bias trends. A common issue is that estimates of growing stock volume derived from OLS or other standard estimation methods often exhibit a narrowed range of variation: large values tend to be underestimated, while small values are overestimated (Gilichinsky et al., 2012). Such trends (design-bias trends) in the resulting maps can be problematic, since "extreme" values are often the most important ones in applications. For instance, old forests with high growing stock volumes are of particular interest for both timber harvesting and biodiversity conservation in forestry scenario modelling (Ståhl et al., 2024).

In statistics, it is often desirable to have estimators that are unbiased and have low variance. As noted, unbiased estimation of individual grid cells in a map is generally impossible, but systematic design-bias trends can be mitigated. The MSE of an estimator can be expressed as the sum

Table 2
ALS data from 2019. Bias and RMSE by group ($m^3 \cdot ha^{-1}$), for the case where negative volume estimates were set to zero.

a) Bias

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	4.6	15.0	6.8	2.1	−24.7
OLS + CC	−0.2	7.5	6.4	9.2	−6.0
OLS + HM	9.2	19.4	16.5	9.0	−7.0
1NN	7.2	6.7	5.9	4.4	−22.7
TLS	−1.8	45.1	50.7	27.0	−36.6
VAOLS	2.5	11.6	6.6	5.1	−16.7
MAVGAR	3.4	−1.4	−4.3	6.1	4.5
MAVGAR + CV	1.6	2.3	−0.5	6.2	−1.9

b) RMSE

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	16.3	28.9	28.3	36.3	55.6
OLS + CC	13.7	28.9	32.3	42.4	54.0
OLS + HM	33.0	35.3	33.4	38.6	63.9
1NN	21.4	35.6	40.8	51.4	71.7
TLS	16.3	73.6	75.5	59.6	72.3
VAOLS	15.2	30.1	31.4	39.7	55.5
MAVGAR	16.2	30.3	37.0	51.2	63.4
MAVGAR + CV	15.0	28.5	33.8	46.8	59.0

Table 3
DSM data from 2018. Bias and RMSE by group ($m^3 \cdot ha^{-1}$), for the case where negative volume estimates were set to zero. Fitted *y*-values were used to compute weights in WLS.

a) Bias

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	14.6	4.2	8.7	4.1	−31.6
WLS	14.6	4.2	8.6	4.0	−31.5
OLS + CC	−5.1	−9.2	8.6	15.8	−5.4
WLS + CC	−5.1	−9.2	8.6	15.7	−5.3
OLS + HM	12.8	16.0	16.0	5.2	−17.2
WLS + HM	12.8	15.9	16.0	5.2	−17.2
1NN	9.1	5.7	9.7	2.4	−26.9
TLS	−5.0	2.2	45.8	54.8	−16.4
VAOLS	3.8	−2.5	8.3	8.9	−20.5
MAVGAR	5.3	−6.8	2.9	7.2	1.1
MAVGAR + CV	2.1	−3.8	9.5	13.3	−8.2

b) RMSE

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	18.0	28.1	35.0	34.9	64.6
WLS	17.9	28.0	34.9	34.9	64.6
OLS + CC	11.4	35.8	44.4	47.0	64.0
WLS + CC	11.4	35.8	44.4	46.9	64.0
OLS + HM	23.0	34.9	38.2	35.2	67.4
WLS + HM	23.0	34.9	38.2	35.2	67.5
1NN	31.8	37.7	42.6	44.8	77.3
TLS	15.3	58.8	110.4	120.2	121.2
VAOLS	11.9	31.5	40.1	40.9	63.8
MAVGAR	13.6	40.8	65.5	64.5	90.6
MAVGAR + CV	11.3	34.1	45.5	46.9	67.2

of its variance and the square of its bias (Wackerly et al., 2008), and a reduction in bias often leads to a corresponding increase in variance. For this reason, reducing bias does not necessarily lead to a lower RMSE; on the contrary, reducing bias may even increase the RMSE. Thus, it is often necessary to balance the benefit of low bias against the risk of a larger RMSE.

Compared to OLS regression, our baseline method, Monte Carlo simulations have demonstrated that it is possible to reduce the design-bias trends using alternative methods. Some of these methods can be considered adjustments of OLS, while others, such as MAVGAR, do not incorporate OLS at any stage of the estimation procedure. Each of the three methods derived as adjustments to OLS reduced bias but at the cost

Table 4
Sentinel 2 data from 2018. Bias and RMSE by group ($m^3 \cdot ha^{-1}$), for the case where negative volume estimates were set to zero.

a) Bias

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	15.5	37.1	26.2	−4.2	−74.7
OLS + CC	−2.4	22.1	37.4	21.9	−18.3
OLS + HM	16.1	25.7	22.0	−3.2	−56.1
1NN	12.9	31.7	21.9	−5.2	−61.4
TLS	69.7	−18.5	5.5	6.9	17.5
VAOLS	3.3	29.7	31.0	6.9	−50.8
MAVGAR	44.4	−3.9	19.3	15.7	9.2
MAVGAR + CV	14.4	11.6	29.0	17.0	−12.7

b) RMSE

Method	Group 1	Group 2	Group 3	Group 4	Group 5
OLS	27.8	53.2	51.1	47.4	107.2
OLS + CC	18.5	65.2	85.5	83.5	103.7
OLS + HM	40.6	50.1	61.1	60.5	108.7
1NN	36.3	72.1	74.9	76.4	120.7
TLS	167.0	65.3	113.4	126.2	170.3
VAOLS	22.5	57.8	65.9	61.8	99.9
MAVGAR	122.6	66.2	105.8	113.9	148.2
MAVGAR + CV	54.0	62.3	89.3	91.5	117.1

of an increased RMSE. VAOLS is one of these three methods; it is scaled to preserve the variance of the target variable. Compared to OLS, VAOLS often resulted in a relatively small increase in RMSE and a slight reduction in bias. Among the three OLS-based adjustments, OLS with CC achieved the greatest bias reduction. The only method that rivaled it in terms of bias was MAVGAR, with or without CV. Compared to OLS, these methods (MAVGAR, with or without CV, and OLS with CC) also produced good or reasonably good RMSE values when association between the target variable and the covariates was sufficiently strong.

WLS (including variants with CC and HM) yielded results very similar to their OLS counterparts. TLS did not perform well. This was also true for 1NN, although to a lesser extent. This may be due to the difficulty of finding close neighbors when the sample size is relatively small.

None of the evaluated methods performed well in producing estimates with low bias when covariates derived from satellite data were used. Among the methods assessed in this paper, if a low RMSE is important for such data, OLS (or WLS) should be used. If reducing bias is the priority, OLS with CC (or WLS with CC) appears to be the best option, although this may result in a notable increase in RMSE for some groups.

CRediT authorship contribution statement

Magnus Ekström: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Christoffer Axelsson:** Writing – review & editing, Software, Formal analysis, Data curation. **Göran Ståhl:** Writing – review & editing, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used a publicly accessible version of ChatGPT in order to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Data availability

Data are available upon reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Co-financing was provided by the Swedish Foundation for Strategic Environmental Research through the research program Mistra Digital Forest (DIA 2017/14 #6). We also thank an anonymous referee for constructive comments that helped improve the manuscript.

Appendices. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fecs.2026.100428.

References

Barth, A., Wallerman, J., Ståhl, G., 2009. Spatially consistent nearest neighbor imputation of forest stand data. Remote Sens. Environ. 113 (3), 546–553. https://doi.org/10.1016/j.rse.2008.09.011.

Borchers, H.W., 2023. Pracma: practical numerical math functions. R package version 2.4.4 (accessed 21 July 2025). https://CRAN.R-project.org/package=pracma.

Cassel, C.M., Särndal, C.E., Wretman, J.H., 1977. Foundations of Inference in Survey Sampling. Wiley, New York.

Chatterjee, S., Hadi, A.S., 2013. Regression Analysis by Example, fifth ed. Wiley, Hoboken.

Chen, Q., McRoberts, R.E., Wang, C., Radtke, P.J., 2016. Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. Remote Sens. Environ. 184, 350–360. https://doi.org/10.1016/j.rse.2016.07.023.

Draper, N.R., Smith, H., 1998. In: Applied Regression Analysis, third ed. Wiley, New York. https://doi.org/10.1002/9781118625590.

Dubayah, R., Armston, J., Healey, S.P., Bruening, J.M., Patterson, P.L., Kellner, J.R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J.B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Luthcke, S., 2022. GEDI launches a new era of biomass inference from space. Environ. Res. Lett. 17 (9), 095001. https://doi.org/10.1088/1748-9326/ac8694.

Ehlers, S., Saarela, S., Lindgren, N., Lindberg, E., Nyström, M., Persson, H.J., Olsson, H., Ståhl, G., 2018. Assessing error correlations in remote sensing-based estimates of forest attributes for improved composite estimation. Remote Sens. 10 (5), 667. https://doi.org/10.3390/rs10050667.

Estes, L.D., Reillo, P.R., Mwangi, A.G., Okin, G.S., Shugart, H.H., 2010. Remote sensing of structural complexity indices for habitat and species distribution modeling. Remote Sens. Environ. 114 (4), 792–804. https://doi.org/10.1016/j.rse.2009.11.016.

Golub, H., van Loan, C.F., 1980. An analysis of the total least squares problem. SIAM J. Numer. Anal. 17 (6), 883–893. https://doi.org/10.1137/0717073.

Greene, W.H., 2011. Econometric Analysis, seventh ed. Prentice Hall, Upper Saddle River.

Gilichinsky, M., Heiskanen, J., Barth, A., Wallerman, J., Egberth, M., Nilsson, M., 2012. Histogram matching for the calibration of kNN stem volume estimates. Int. J. Rem. Sens. 33 (22), 7117–7131. https://doi.org/10.1080/01431161.2012.700134.

Gregoire, T.G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Can. J. For. Res. 28 (10), 1429–1447. https://doi.org/10.1139/x98-166.

Heeringa, S.G., West, B.T., Berglund, P.A., 2010. Applied Survey Data Analysis. Chapman & Hall/CRC, Boca Raton. https://doi.org/10.1201/9781420080674.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2021. In: An Introduction to Statistical Learning: with Applications in R, second ed. Springer, Berlin. https://doi.org/10.1007/978-1-0716-1418-1.

Kangas, A., Myllymäki, M., Mehtätalo, L., 2023. Understanding uncertainty in forest resources maps. Silva Fenn. 57 (2). https://doi.org/10.14214/sf.22026.

Kuhn, M., 2022. Caret: classification and regression training. R package version 6.0-93 (accessed 21 July 2025). https://CRAN.R-project.org/package=caret.

Li, A., Liang, S., Wang, A., Qin, J., 2007. Estimating crop yield from multi-temporal satellite data using multivariate regression and neural network techniques. Photogramm. Eng. Rem. Sens. 73 (10), 1149–1157. https://doi.org/10.14358/PERS.73.10.1149.

Lindgren, N., Olsson, H., Nyström, K., Nyström, M., Ståhl, G., 2022a. Data assimilation of growing stock volume using a sequence of remote sensing data from different sensors. Can. J. Rem. Sens. 48, 127–143. https://doi.org/10.1080/07038992.2021.1988542.

Lindgren, N., Nyström, K., Saarela, S., Olsson, H., Ståhl, G., 2022b. Importance of calibration for improving the efficiency of data assimilation for predicting forest characteristics. Remote Sens. 14, 4627. https://doi.org/10.3390/rs14184627.

Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson, S., Nilsson, L., Eriksson, J., Olsson, H., 2017. A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the national forest inventory. Remote Sens. Environ. 194, 447–454. https://doi.org/10.1016/j.rse.2016.10.022.

Persson, H.J., Ståhl, G., 2020. Characterizing uncertainty in forest remote sensing studies. Remote Sens. 12 (3), 505. https://doi.org/10.3390/rs12030505.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team, 2021. Nlme: Linear and nonlinear mixed effects models. R package version 3, pp. 1–152 (accessed 21 July 2025). https://CRAN.R-project.org/package=nlme.

R Core Team, 2025. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, J.A., 2022. Remote Sensing Digital Image Analysis: an Introduction, sixth ed. Springer, Berlin. https://doi.org/10.1007/978-3-030-82327-6.

Särndal, C.E., Swensson, B., Wretman, J., 1992. Model Assisted Survey Sampling. Springer, New York. https://doi.org/10.1007/978-1-4612-4378-6.

Shukla, G.K., 1972. On the problem of calibration. Technometrics 14 (3), 547–553. https://doi.org/10.2307/1267283.

Ståhl, G., Gobakken, T., Saarela, S., Persson, H.J., Ekström, M., Healey, S.P., Yang, Z., Holmgren, J., Lindberg, E., Nyström, K., Papucci, E., Ulvdal, P., Ørka, H.O., Næsset, E., Hou, Z., Olsson, H., McRoberts, R.E., 2024. Why ecosystem characteristics predicted from remotely sensed data are unbiased and biased at the same time–and how this affects applications. For. Ecosyst. 11 (1). https://doi.org/10.1016/j.fecs.2023.100164.

Wackerly, D.D., Mendenhall, W., Scheaffer, R.L., 2008. Mathematical Statistics with Applications, seventh ed. Thomson Learning, Inc., Belmont.

Wikström, P., Edenius, L., Eriksson, L.O., Lämås, T., Sonesson, J., Öhman, K., Waller, C., Klintebäck, F., 2011. The Heureka forestry decision support system: an overview. Math. Comput. For. Nat. Resour. Sci. 3 (2), 87–95.