Taylor & Francis
Taylor & Francis Group

Canadian Journal of Remote Sensing
Journal canadien de télédétection

ISSN: 0703-8992 (Print) 1712-7971 (Online) Journal homepage: www.tandfonline.com/journals/ujrs20

Updating forest maps through data assimilation using remotely sensed data and regression methods designed to avoid bias trends

Christoffer Axelsson, Magnus Ekström, Alex Appiah Mensah, Nils Lindgren & Göran Ståhl

To cite this article: Christoffer Axelsson, Magnus Ekström, Alex Appiah Mensah, Nils Lindgren & Göran Ståhl (2026) Updating forest maps through data assimilation using remotely sensed data and regression methods designed to avoid bias trends, Canadian Journal of Remote Sensing, 52:1, 2623684, DOI: 10.1080/07038992.2026.2623684

To link to this article: https://doi.org/10.1080/07038992.2026.2623684

CANADIAN JOURNAL OF REMOTE SENSING
2026, VOL. 52, NO. 1, 2623684
https://doi.org/10.1080/07038992.2026.2623684

Updating forest maps through data assimilation using remotely sensed data and regression methods designed to avoid bias trends

Mise à jour des cartes forestières par assimilation de données de télédétection et des méthodes de régression conçues pour éviter les biais de tendance

Christoffer Axelsson[a], Magnus Ekström[a], Alex Appiah Mensah[a], Nils Lindgren[b], and Göran Ståhl[a]

[a]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå; [b]The Forest Research Institute of Sweden (Skogforsk), Uppsala

ABSTRACT

We applied data assimilation for updating map information about growing stock volume in forests in a study area in northern Sweden, across a study period ranging from 2010 to 2022. Novel features of the study were that (i) we applied newly developed regression techniques for predicting forest characteristics from remotely sensed data, designed to avoid the bias trends that arise from standard regression methods, (ii) we applied growth models that utilized information about site quality and age, assessed wall-to-wall for the study area, based on data from repeated airborne laser scanning surveys, and (iii) we used a fully empirical approach to computing weights for the DA filter. The results showed that the accuracy obtained for predictions of growing stock volume from an initial laser scanning survey could be improved upon across the study period when a sequence of predictions using optical satellite data, digital aerial photos, and a final laser scanning survey were assimilated. Using only a sequence of optical satellite data and digital aerial photos, the accuracy of the initial laser scanning-based predictions could be maintained across the study period. The new regression methods performed better than standard regression methods in terms of avoiding bias trends, but the best overall results in terms of accuracy were obtained for standard regression combined with classical calibration. The study confirms findings from previous similar studies that data assimilation has a potential to maintain or slightly improve the accuracy of growing stock volume predictions from an initial high-quality laser scanning survey through assimilating a series of predictions from lower-quality remotely sensed data across a relatively long period of time.

RÉSUMÉ

Nous avons appliqué une méthode d'assimilation de données pour mettre à jour l'information cartographique relative au volume de bois sur pied pour les forêts d'une zone d'étude du nord de la Suède, sur la période 2010-2022. Les aspects novateurs de cette étude impliquent que : (i) nous avons utilisé des techniques de régression récemment développées, conçues pour éviter les biais inhérents aux méthodes de régression classiques, pour prédire les caractéristiques forestières à partir de données de télédétection ; (ii) nous avons appliqué des modèles de croissance intégrant la qualité et l'âge des stations, évalués sur l'ensemble de la zone d'étude, à partir d'acquisition multiples de relevés lidar aéroportés ; et (iii) nous avons utilisé une approche entièrement empirique pour le calcul des pondérations du filtre d'assimilation de données. Les résultats ont montré que l'exactitude des prédictions de volume de bois sur pied obtenues à partir d'un relevé lidar initial pouvait être améliorée tout au long de la période d'étude grâce à l'assimilation d'une séquence de prédictions utilisant des données satellitaires optiques, des photographies aériennes numériques et d'un relevé lidar final. L'utilisation d'une séquence de données satellitaires optiques et de photographies aériennes numériques a permis de maintenir l'exactitude des prédictions

ARTICLE HISTORY
Received 30 June 2025
Accepted 13 January 2026

KEYWORDS
Data assimilation; data fusion; forest inventory; mapping; regression analysis

CONTACT Göran Ståhl goran.stahl@slu.se Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07038992.2026.2623684.

initiales basées sur le relevé lidar sur l'ensemble de la période d'étude. Les nouvelles méthodes de régression ont surpassé les méthodes de régression standard pour éviter des biais. Toutefois, les meilleurs résultats globaux en termes d'exactitude ont été obtenus avec la régression standard combinée à un étalonnage classique. Cette étude confirme les résultats d'études similaires antérieures : l'assimilation de données permet de maintenir, voire d'améliorer légèrement, l'exactitude des prévisions de l'accroissement des volumes suite à une première campagne de balayage lidar de haute qualité, grâce à l'assimilation d'une série de prévisions issues de données de télédétection de moindre qualité sur une période relativement longue.

Introduction

Wall-to-wall maps of forest resources are currently increasingly demanded by stakeholders in forestry and nature conservation (e.g., Lechner et al. 2020). In forestry, such maps have been used for a long time for operational decision making (e.g., Bettinger et al. 2016), whereas for nature conservation they are relatively novel features. The quality of this kind of maps has improved considerably during recent years, following the introduction of large-area laser scanning surveys (e.g., Nilsson et al. 2017).

Whereas the accuracy of maps for characteristics such as growing stock volume is relatively high, uncertainties increase over time after a laser scanning survey because the surveys are expensive and can thus only be afforded at sparse intervals. Updating is typically made through applying growth models, but such growth forecasts tend to have poor accuracy due to the limited amount of information available wall-to-wall.

However, data assimilation (DA) is currently finding applications in forest inventories because of increasing amounts of non-expensive remotely sensed data, in addition to data from laser scanning. With DA, a sequence of state observations is combined with model forecasts between the time points of observations (e.g., Evensen et al. 2022). Updates at each time point are made to maximize the accuracy of state estimates. This stands in contrast to updating through model predictions, only, from an initial estimate.

Typical applications of DA involve obtaining best estimates of current or future states of dynamic systems. In the case of forest inventories, the dynamic system is the growing forest. Recent studies of DA in the context of forest inventory can be found in, e.g., Ehlers et al. (2013), Hou et al. (2021) and Lindgren et al. (2022a). Compared to application of DA in other fields, several specific features must be considered in the context of forest inventory.

A first such feature is that model predictions based on remotely sensed data, rather than direct state observations, are commonly applied. Whereas model-based predictions are typically approximately model-unbiased at the level of population elements (e.g., Draper and Smith 1998) they are in the same time design-biased for such elements in single fixed populations (e.g. Ståhl et al. 2024), which is noted through typical over-prediction of small true values and under-prediction of large true values. In this context it is important to distinguish between model-unbiasedness and design-unbiasedness (cf., Cassel et al. 1977). Model-unbiasedness is assessed across hypothetical realizations of populations whereas design-unbiasedness is assessed across hypothetical samples from a single fixed population. For fixed populations, such as real-world forests, model-based prediction may thus lead to problematic fixed-population bias. Such bias cause problems in DA systems (e.g., Lindgren et al. 2022b). In this paper, a new regression technique for mitigating this problem is proposed.

Another problem in forest inventory applications of DA is that prediction errors across time for population elements tend to be strongly correlated, and thus standard DA methods such as the Kalman filter must be modified (e.g., Lindgren et al. 2022a) to take such correlations into account. Further, field reference data for calibrating and evaluating prediction models are relatively sparse, and thus machine learning solutions that require large amounts of data tend to lend themselves poorly to DA systems for forest inventory.

A third problem in forest inventory applications of DA is that information for wall-to-wall growth predictions is limited. For example, information about site quality and stand age is important for accurate growth forecasts, but such information can be difficult to obtain wall-to-wall across large

areas. However, through repeating laser scanning surveys there is a possibility to accurately assess the relative height growth of forests wall-to-wall. Because the height growth at a certain stand height is closely related to site quality and age, new methods for developing wall-to-wall maps of such features have recently emerged (e.g., Noordermeer et al. 2021).

The objective of this study was to assess the performance of DA for updating map information about growing stock volume. Three novel features were introduced compared to previous DA studies, such as Lindgren et al. (2022a). Firstly, we applied new regression methods, designed to reduce the bias trends for predictions in fixed populations that emanate from using standard model-based prediction techniques. Secondly, we applied wall-to-wall information on site quality (Mensah et al. 2023) and stand age, assessed from data from repeated laser scanning, for facilitating accurate wall-to-wall growth updates. Thirdly, we used a fully empirical approach for computing weights for the DA procedure.

The study was conducted in northern Sweden, using field plots from the Swedish National Forest Inventory (NFI) as reference data. From an initial survey based on laser scanning conducted in 2010 we applied DA until 2022 and compared results from different DA schemes, involving assimilating predictions from different sequences of remotely sensed data, or only growth forecasts from an initial prediction of growing stock volume.

Materials and methods

The study area

The study was conducted in northern Sweden (Figure 1), within an approximately 300 × 300 km large study area centered around the Krycklan research site. This area was used for obtaining field

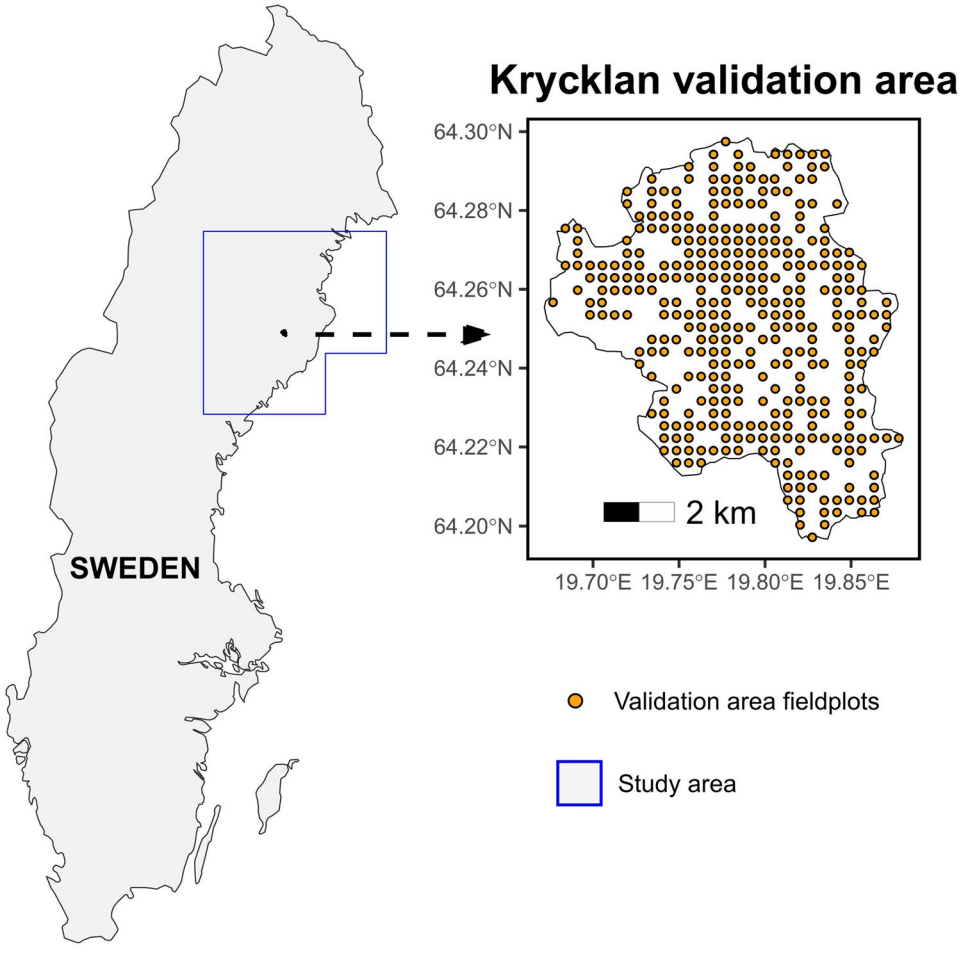

Figure 1. Location of the study area and the Krycklan validation area.

reference data and for general evaluation of the performance of different DA schemes based on NFI data, whereas a dense network of plots within Krycklan (the Krycklan validation area) was used for validating the results from the different DA schemes for the year 2019, based on an independent data set.

Krycklan area validation data

The Krycklan research site is situated approximately 50 km northwest of the city of Umeå in northern Sweden (64°25′ N, 19°80′ E). The forests of the ~7000 hectares large area are dominated by Scots pine and Norway spruce. Several research activities of various kinds are conducted within the area.

For validation, we used approximately 300 field plots distributed in a systematic pattern, inventoried in 2019 across Krycklan. The plots have a radius of 10 meters. During the inventory, stem diameters at 1.3 m above ground (DBH) were measured using a DP II caliper (Haglöf, Sweden). Tree species was recorded, and the heights of subsample trees were measured using a Vertex Laser Geo range finder (Haglöf, Sweden). Based on the collected field data, plot level growing stock volumes were computed using the Heureka software, through species-specific models with diameter and height as input (Wikström et al., 2011).

National forest inventory data from the study area

For model training, and assessment of the performance of DA through cross-validation, we used field plot data from the Swedish NFI (e.g., Fridman et al. 2014). We used data from both permanent (n = 1470) and temporary (n = 3602) NFI plots within the study area (Figure 1). Temporary plots are established every year, but inventoried only at a single point in time, while permanent plots are revisited every five years. Plot volumes of permanent plots were projected to the year of acquiring remotely sensed data from their last recorded measurement using growth models in the Heureka software (Wikström et al. 2011). Because Heureka forecasts growth for five year periods, linear interpolation was applied for intermediate years in case of no signs of disturbance on a plot. Plots were categorized as disturbed if they had a reduction in volume exceeding 5 m^3/ha between measurements. Temporary plots were used for model training only for the year they were measured, i.e., no forecasting was implemented for such plots. The NFI plots were used to train prediction models whenever they had coverage of the remote sensing dataset and had a measured or interpolated volume for the given year.

However, only a subset (n = 587) of the permanent plots were used for assessing the performance of the DA schemes. These plots were spatially distributed across the entire study area. Their distribution across different stand conditions is shown in Table 1. We only included plots with measured or interpolated volumes available for every year during the study period (2010–2022) and thus excluded plots where disturbances had occurred. In addition, a small number of plots were removed because all data for applying the growth models were not available.

As part of the analysis, we categorized the NFI plots into low (≤75 m^3/ha), intermediate (76–150 m^3/ha) and high volume (>150 m^3/ha) plots. The categorization was made according to their volume in 2016 for the purpose of avoiding that plots switched category during the course of the study period.

Table 1. Summary of permanent NFI plots used for evaluating the DA procedure.

Category	Number of plots
Undisturbed plots used for assessing DA performance	587
Pine dominated (≥65% Pine)	248 (42%)
Spruce dominated (≥65% Spruce)	41 (7%)
Mixed conifer (≥65% Conifers)	271 (46%)
Mixed forest (<65% Conifers)	27 (5%)
Low volume (≤75 m^3/ha)	197 (34%)
High volume (>150 m^3/ha)	194 (33%)

Table 2. Remote sensing datasets used for predicting growing stock volume.

Year	ALS	DSM	Sentinel 2	Landsat 8
2010	X			
2011				
2012				
2013				
2014				X
2015			X	
2016				
2017			X	
2018		X	X	
2019	X		X	
2020			X	
2021			X	
2022		X	X	

Remote sensing data for growing stock volume predictions

We gathered remote sensing datasets which were subsequently used for predicting growing stock volumes. Separate prediction models were estimated and applied for each dataset. The datasets fall into three categories: airborne laser scanning (ALS), digital surface models (DSM) derived from aerial photos, and satellite imagery. A summary of the datasets used during the study period (2010–2022) is shown in Table 2. It can be observed that only few datasets were available from the first part of the study period.

ALS data are the best, among the available datasets, for predicting volume. We used data from the National Mapping Agency in Sweden, which carries out wall-to-wall ALS campaigns of most of the territory of Sweden with a rotation of about 7–9 years. The first campaign scanned the study area between 2009 and 2013, and the Krycklan validation area in 2010. The second campaign scanned the study area during 2018 and 2021, and the Krycklan validation area in 2019. The point densities were 0.5-1 and 1–2 pulses/m^2 for the first and second campaign, respectively. During the first campaign, some plots were scanned during leaf on conditions and others during leaf off conditions. Models were developed separately for these two cases and the results were merged.

The DSM is an elevation model of the upper canopy of the vegetation, which is created through aerial image matching. It is produced by the Swedish National Mapping Agency and covers most of Sweden, with varying resolution. We used surface models in raster format from 2018 and 2022, with 1.0 and 0.4 meter resolution, respectively.

We aimed at collecting one scene of satellite imagery per year with the focus of identifying cloud free imagery from the summer period, covering as large parts of the larger study area as possible. Sentinel 2 (L2A version) was available from 2015 and onwards[1], and we added a Landsat 8 scene from 2014 for better temporal coverage of the early part of the study period.

Model predictions of growing stock volume

For each NFI plot, we recorded the value of growing stock volume along with covariates from the remotely sensed data for the plot. Thus, we had data in the form $(x_{1i},\ldots,x_{pi,}\ y_i), i=1,\ldots,n$, where n is the sample size from the NFI within the study area. The values of the covariates $x_1,\ldots,x_p$ were available from the remote sensing data for the entire population, i.e., for all N grid cells (with resolution 12,5 × 12,5 m) in the study area. Note that the NFI plots did not perfectly match the grid cell system and thus for the NFI plots we derived covariate data specifically for the plot areas. The models were set up to predict per-hectare values which were subsequently rescaled.

Transformations of the basic metrics, such as squares, logarithms, and combinations of variables (see Appendix A), were also tested and treated as additional covariates. The values of the target variable, y, were only available for the sample of field plots.

The general objective was to predict the y-values for all elements in the population. This can be accomplished by fitting a regression model to the observed data points $\left(x_{1i},\ldots,x_{pi},y_i\right)$, for $i=1,\ldots,n$. The fitted model takes the form

$$\widehat{y_i}=\widehat{\beta}_0+\widehat{\beta}_1x_{1i}+\ldots+\widehat{\beta}_px_{pi}, i=1,\ldots,N,$$

where $\widehat{\beta}_0,\ldots,\widehat{\beta}_p$, are parameter estimates obtained from the sample data, and $\check{y}_i$ represents the predicted value of y_i, $i=1,\ldots,N$.

If ordinary least squares (OLS) regression is applied, the parameter estimates are given by

$$\begin{pmatrix}\widehat{\beta}_0\\ \vdots\\ \widehat{\beta}_p\end{pmatrix}=\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{y},$$

where $\boldsymbol{y}$ is the vector of observed y_i values, $i=1,\ldots,n$, and $\mathbf{X}$ is the design matrix. The first column of $\mathbf{X}$ consists of unit values (to account for the intercept term), and the columns $j+1$, $j=1,\ldots,p$, contain the sample values of covariate $x_{ji}, i=1,\ldots,n$ (e.g., Draper and Smith 1998).

As noted by Ståhl et al. (2024), OLS predictions exhibit a certain systematic bias for fixed populations; large true values of the target variable tend to be underpredicted, while small values tend to be overpredicted. Lindgren et al. (2022a, 2022b) proposed using classical calibration to reduce this bias. In this two-stage approach, a first set of predictions $\widehat{y}_i$ for y_i, $i=1,\ldots,N$, is initially obtained using OLS. In a second stage, a calibration model is fitted using OLS,

$$\hat{\hat{y}}_i=\widehat{\alpha}_0+\widehat{\alpha}_1 y_i,$$

where the double-hat notation $\hat{\hat{y}}_i$ indicates that these are predictions of the previously predicted values $\widehat{y}_i$. Rather than using this model to predict predictions, Lindgren et al. (2022a, 2022b) reformulated the expression to directly adjust the initial predictions. The calibrated prediction of y_i, denoted $\widehat{y}_{i,c}$, is given by

$$\widehat{y}_{i,c}=\frac{\widehat{y}_i-\widehat{\alpha}_0}{\widehat{\alpha}_1}.$$

This formulation provides a bias-corrected prediction of the target variable based on the initial OLS prediction and the parameters of the calibration model.

As an alternative approach to mitigating bias trends, we applied new methods developed in the context of this study. Instead of determining model parameters by minimizing the sum of squared deviations between model predictions and observations in a standard way (e.g., Draper and Smith 1998), the key idea behind the new regression method was to estimate model parameters by minimizing a measure of systematic prediction bias across the range of the response variable. To ensure that predictions were approximately locally unbiased, the data were partitioned into windows of the response. Within each window, the average residual was computed, and the absolute value of this average residual was used as a measure of local deviation. The model parameters were chosen to minimize the mean of the absolute values of the average residuals, thereby reducing systematic over- or under-prediction in any region of the response space.

Thus, we aimed to make model-based predictions such that, for any value y, the average prediction error would be small among all population elements whose true y-values lie within a neighborhood around y. To ensure that this property would hold across the full range of values, we divided the range of observed y-values in the sample into g disjoint intervals of equal length.

Let m_i denote the midpoint of interval i. For a window (or neighborhood) of width w, defined as the interval $[m_i-w/2, m_i+w/2)$, let n_i represent the number of y-values falling in the window, and denote these values by y_{ij}, $j=1,\ldots,n_i$ (and correspondingly for $x_{1ij},\ldots,x_{pij}$).

Our goal was to find coefficients $\beta_0,\ldots,\beta_p$ such that the average residual in each window,

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

would be as close to zero as possible. To achieve this across all g windows, we defined the "optimal" coefficients $\breve{\beta}_0,\ldots,\breve{\beta}_p$ as those that minimize the quantity

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

If desired, the window size w can be set equal to the length of each interval. However, this may lead to very small values of n_i in some intervals. To address this, it could be beneficial to choose w larger than a single interval length. Alternatively, cross-validation could be used to select an appropriate value for w. In this study, we set $g=100$, and 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 mean squared error (MSE). The algorithm for computing k-fold cross-validated MSEs can be found in, e.g., James et al. (2021).

In the study, OLS regression was applied as a baseline prediction method, which was compared with OLS in combination with classical calibration, and the newly proposed method described above, with and without application of cross-validation to determine window size. Like classical calibration, the new method will not result in unbiased predictions at the level of individual population elements, but it will mitigate the trend from over-predicting small true values to under-predicting large true values, which is otherwise typically a result of model-based prediction in fixed populations.

Model predictions of growth

Growth forecasting is important for two reasons in our proposed DA scheme. Firstly, as previously described, reference data from NFI plots were interpolated between measurements through application of growth models in the Heureka system (e.g., Wikström et al. 2011), specifically tailored to fit NFI data. Secondly, as will be described in the following, growth models were developed to allow for forecasts wall-to-wall, as required in operational DA systems. For this to be possible, the models could only utilize information available wall-to-wall.

Models predicting growth of growing stock volume were developed for plots dominated by pine, spruce, mixed-conifers (pine, spruce, and contorta), and mixed-forests (conifers and broadleaves). We used data from repeated measurements of permanent plots from the Swedish NFI within the study area to estimate the parameters of the growth models. The data covered the period from 2003 to 2022, and comprised observations of 596 five-year growth periods for pine, 207 for spruce, 78 for mixed conifers, and 561 for the mixed-forest group.

Since the permanent plots are remeasured at an interval of five years, the increment data from these plots refer to a five-year increment period. The dynamic algebraic difference form of the Chapman-Richards' generalization of von Bertalanffy's growth model (Mensah et al. 2021) was used to project the five-year growth. In addition, site index was included in the model as an interaction with the initial volume to account for the different carrying capacities of stands of the same age (Mensah et al. 2022). The site-dependent dynamic volume growth model for the coniferous forests (i.e., pine, spruce, and mixed conifers) was formulated as

$$y_2=\left(y_1+\alpha_3\cdot \mathrm{SI}\right)\left[\frac{1-\mathrm{e}^{-\alpha_1 T_2}}{1-\mathrm{e}^{-\alpha_1 T_1}}\right]^{\alpha_2+\alpha_3\cdot \mathrm{SI}}+\varepsilon,$$

where y_1 is the growing stock volume at the initial age (T_1), y_2 is the volume at the end of the growth period, at the remeasurement age (T_2), SI is site index according to site properties (Hägglund and Lundmark, 1977), and the α parameters are to be estimated from the sample data.

Since site index information was unavailable for the mixed-forests group, other easily available data describing site resource potential for tree growth were used to parameterize the volume growth model in this case. The final form of the mixed-forest volume growth prediction model was determined as

$$y_2 = \left(y_1 + \alpha_3 \cdot \text{distC}\right)\left[\frac{1-e^{-\alpha_1 T_2}}{1-e^{-\alpha_1 T_1}}\right]^{\alpha_2 \cdot \text{Alt}^{-\alpha_4}} + \varepsilon,$$

where distC is the distance to the nearest coast (in km) and Alt is altitude (m.a.s.l).

After preliminary tests for autocorrelation and heteroscedasticity (which showed no gain in information compared to their alternative models), the parameters of the final growth models were estimated using nonlinear least squares regression, and the individual random errors (ε) were assumed to be normally and identically distributed, with a constant variance. Independent field inventory data for the period 2015–2019 were available from the Krycklan validation area for pine, spruce, and mixed-conifer forests. These comprised 211 growth observations for pine, 108 for spruce, and 97 for mixed conifers. The fitted models' estimated parameters, residual diagnostics, and validation statistics are given in Appendix B.

Because the DA technique requires annual updating, the growth function was adjusted to make the forecasts compatible. Thus, the volume increment was computed on an annual time scale by computing the first derivative of the growth function with respect to age, and the volume at the endpoint of a period was obtained as the initial volume plus the annual increment.

To implement the growth models for volume forecasting over the entire area, we utilized available remotely-sensed maps of tree species, stand age, and site index. The maps of site index (Mensah et al., 2023) and stand age were derived from nationwide bi-temporal airborne-laser scanning data (2009–2023) and have root mean square prediction errors of 2 m and 8 years (for stands up to 200 years), respectively. The tree species map was created from a combination of ALS and Sentinel 2 satellite data (mainly from the year 2018). The proportions of tree species were modeled with a kNN model using NFI plots as references (Reese et al., 2003).

Forest harvest detection

DA of forest attributes should ideally be integrated with methods to detect sudden disturbances, such as forest harvesting. To identify disturbances within the Krycklan validation area, for which we developed wall-to-wall maps based on DA, we used a map product with detected harvest sites provided by the Swedish Forest Agency (2024). The product is updated every year and relies on change detection in satellite imagery to identify and delineate areas where harvesting has occurred.

Data assimilation

Many different approaches to performing DA exist (e.g., Evensen et al. 2022). The general principle is to combine uncertain model forecasts with repeated uncertain state observations in a fashion so that optimal estimates of past, present, and future states of a dynamic system can be obtained. In many cases the focus is on the model (e.g., Peylin et al. 2016), whereas in other cases, such as in our study, the focus is on the repeated state observations (e.g., Lindgren et al. 2022a). Following Lindgren et al. (2022b), in this study we applied different techniques for making our model-based predictions resemble state observations, mainly by adjusting for bias trends. Further, compared to using a standard Kalman filter for computing the DA predictions (e.g., Kalman 1960), we applied a filter that took into account correlated errors (e.g., Ehlers et al. 2018). Rather than setting up a theoretical system for estimating the relevant variances and covariances across time, based on what data had previously been assimilated (cf., Lindgren et al. 2022a), we used a fully empirical system for estimating the parameters needed for the filter. Thus, at all points in time, the reference NFI plots were used for estimating the variance of errors for the updated state estimate, the variance of errors for the new predictions to be assimilated, and the covariance between these errors.

Because our evaluations utilized the undisturbed NFI plots from the study area, as well as the Krycklan validation area (wall-to-wall), we applied our DA procedures only for those.

The DA procedure comprised the following steps:

1. Initial growing stock volume predictions were made for the 587 undisturbed NFI plots within the study area and wall-to-wall for the Krycklan validation area. For all predictions, we applied 5-fold cross-validation and trained the models using all available temporary and permanent NFI plots.
2. One-year growth was modeled using the growth models developed in the project. Growth was added to the initial volume prediction.
3. We identified if a new remote sensing dataset was available the following year and, if so, it was used to make new predictions of volume. We calculated the weight for the new predictions (e.g., Lindgren et al., 2022a) based on the residual errors from the new predictions and the residual errors from the forecasts from step 2, using the volumes from the NFI plots for that year as reference data:

$$Weight = (Var_1 - Cov) / \left(Var_1 + Var_2 - 2\,Cov\right)$$

4. Here, Var_1 is the variance of residual errors between the forecasts and the references and Var_2 is the variance of residual errors between the new predictions and the references. Cov is the covariance between the two sets of residual errors. These terms were estimated empirically. We updated the estimated volumes by merging the forecasts and the new predictions, using the estimated weights in a composite estimator. The weight computed according to the presented formula was assigned to the new prediction, whereas a weight corresponding to one minus this weight was assigned to the forecast. In the case of several new datasets being available, we chose to only assimilate the best available new dataset for each NFI plot (or map element) and year. Here, ALS was deemed the best, followed by DSM, and satellite imagery.
5. Steps 2 and 3 were repeated until the end of the study period.

For the Krycklan validation area, an additional step was applied between Steps 2 and 3. It involved assessing whether or not a map element had been affected by harvesting. If it had, the forecasted volume was adjusted accordingly.

The weighting principle presented in Step 3 is the foundational principle of inverse variance weighting that is applied both in the Kalman filter and other similar DA techniques (see, e.g., Kalnay 2003, p. 146). However, whereas the basic Kalman filter assumes independent state observations, and independence between state observations and forecasts, the covariance term in Step 3 takes dependencies into account. Such dependencies typically decrease the efficiency of DA (e.g., Ehlers et al. 2018).

For every year, we computed the root mean squared error (RMSE) and bias between the DA-predicted and measured (or interpolated) volumes, based on data from the 587 NFI plots that remained undisturbed across the study period. To evaluate the predictions specifically for the Krycklan validation area, we extracted the predicted growing stock volumes for the year 2019 and compared with independent field plot data from that time point.

Three different DA schemes were evaluated. These were:

Case 1: Prediction of growing stock volume from all available remote sensing datasets. The initial prediction was based on ALS data.

Case 2: Following an initial prediction of growing stock volume based on ALS data, only optical satellite data and DSM data were used for the remainder of the study period.

Case 3: Optical satellite data, only, were used. Contrary to the other two cases, ALS data were *not* used for predicting the initial growing stock volumes in this case.

For all three cases, different methods for mitigating bias were applied, using classical calibration, the new method proposed in this article, and the new method combined with cross-validation. As a first baseline method, growth forecasts, only, were made across the entire study period from an

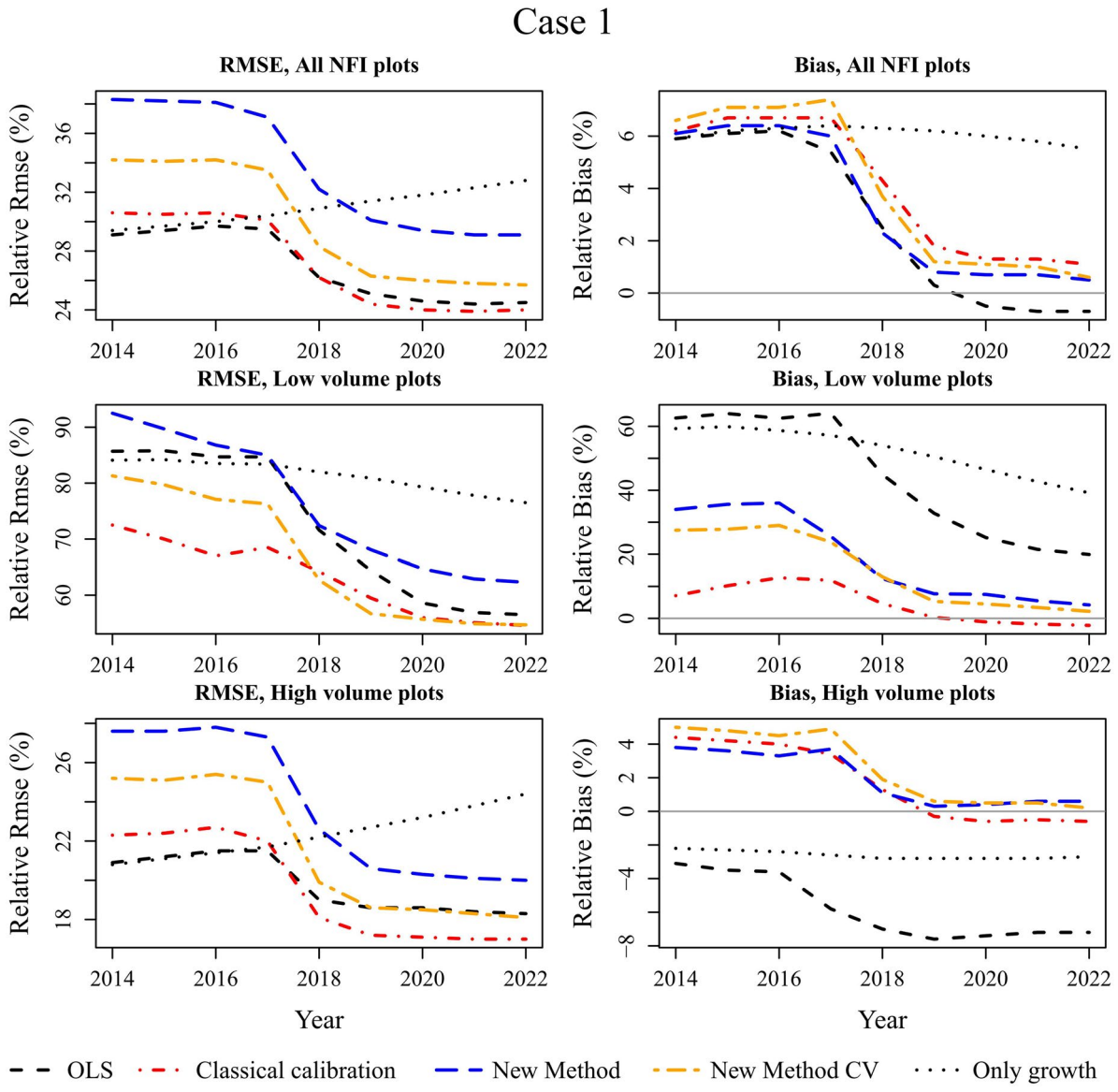

Figure 2. Results for Case 1, i.e., use of all available predictions. Results are reported in terms of relative root mean square error (left) and estimated bias (right), separated on results for all NFI plots (upper panel), plots with low volumes (middle panel), and plots with high volumes (lower panel).

initial growing stock volume prediction (based on OLS regression, without bias adjustment). As a second baseline method, OLS regression was applied together with DA, without bias adjustments.

Results

Below, we present results for the study area based on NFI data and cross-validation, for the three different DA schemes (Cases 1 to 3). Results are reported separately for plots with high and low volumes, to assess what RMSE and bias the different methods led to for such plots.

Note that although our study period formally started in 2010, we present results from 2014 and onwards in the figures, because only in 2014 all plots had obtained a first RS-based volume prediction.

For Case 1 (Figure 2), a general trend of decreasing relative RMSEs across the study period can be observed for all the different prediction methods. OLS regression and OLS regression with classical calibration generally performed best in terms of RMSE, whereas the new regression methods and OLS regression followed by classical calibration reduced the bias considerably

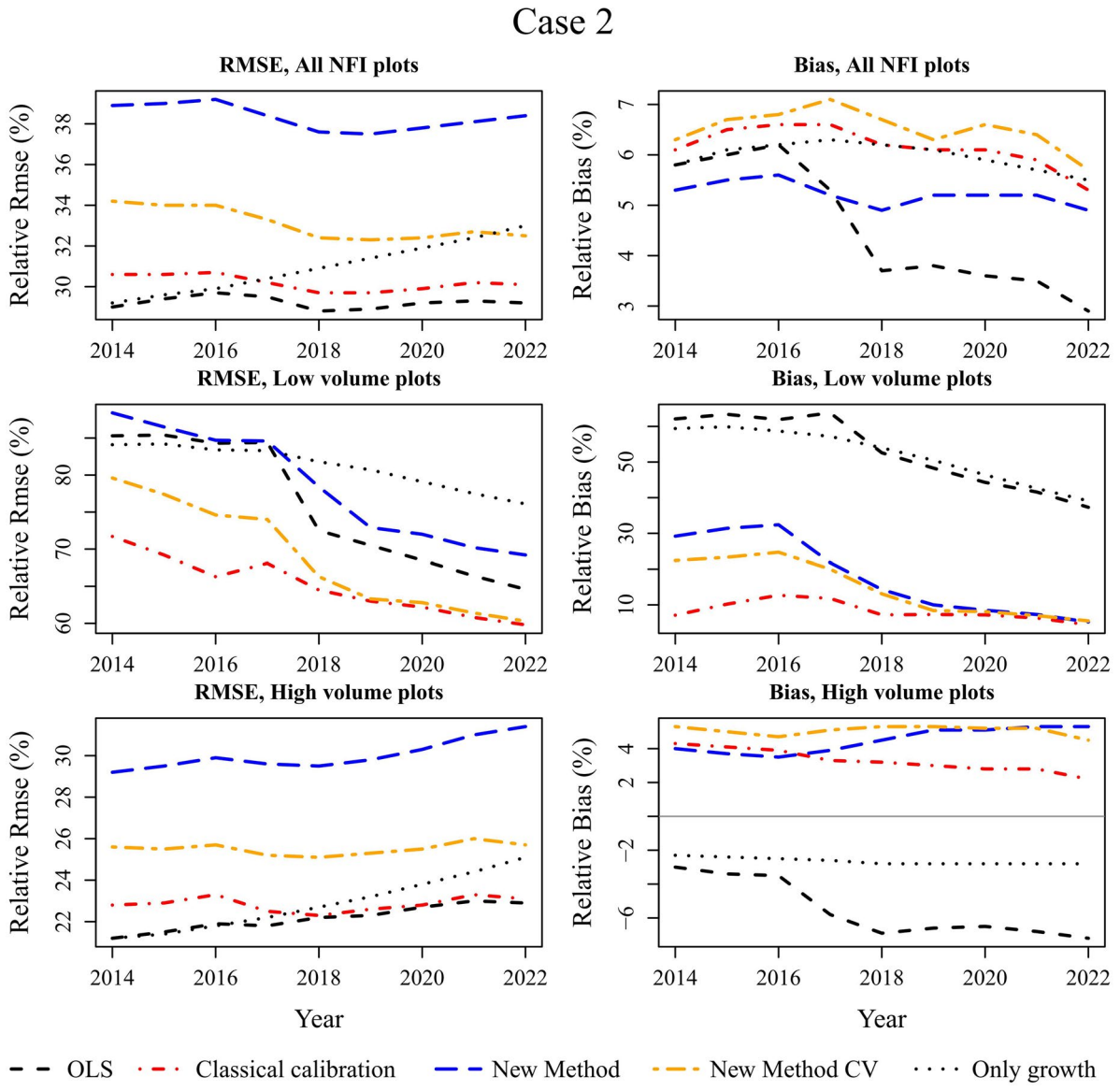

Figure 3. Results for Case 2, i.e., an initial LiDAR-based prediction of growing stock volume was followed by assimilating a sequence of predictions from optical satellite and DSM data. Results are reported in terms of relative root mean square error (left) and estimated bias (right), separated on results for all NFI plots (upper panel), plots with low volumes (middle panel), and plots with high volumes (lower panel).

compared to using only OLS regression (at low and high true volumes). In general, all the DA schemes performed much better in terms of RMSE and bias compared to applying growth forecasts, only, from the initial ALS-based OLS regression estimate (the dotted lines in the figure). Looking at the overall results in terms of RMSE (upper left panel in Figure 2), all DA schemes resulted in decreasing RMSEs across the study period, whereas applying growth forecasts, only, led to increased RMSE.

For Case 2 (Figure 3), it can be observed that the DA schemes implied that the accuracy from the initial assessment, based on ALS data, could be (approximately) maintained across the study period, although results differed slightly between the different study domains. This could be compared to using growth forecasts, only, from the initial ALS-based assessment, which resulted in decreased accuracy across the study period, except in the case of low true volumes (in which case the increased average volumes across time resulted in lower relative RMSEs).

Regarding estimated bias, substantial over-prediction of volume for small true volumes and under-prediction of large true volumes remained a problem when OLS regression was applied. However,

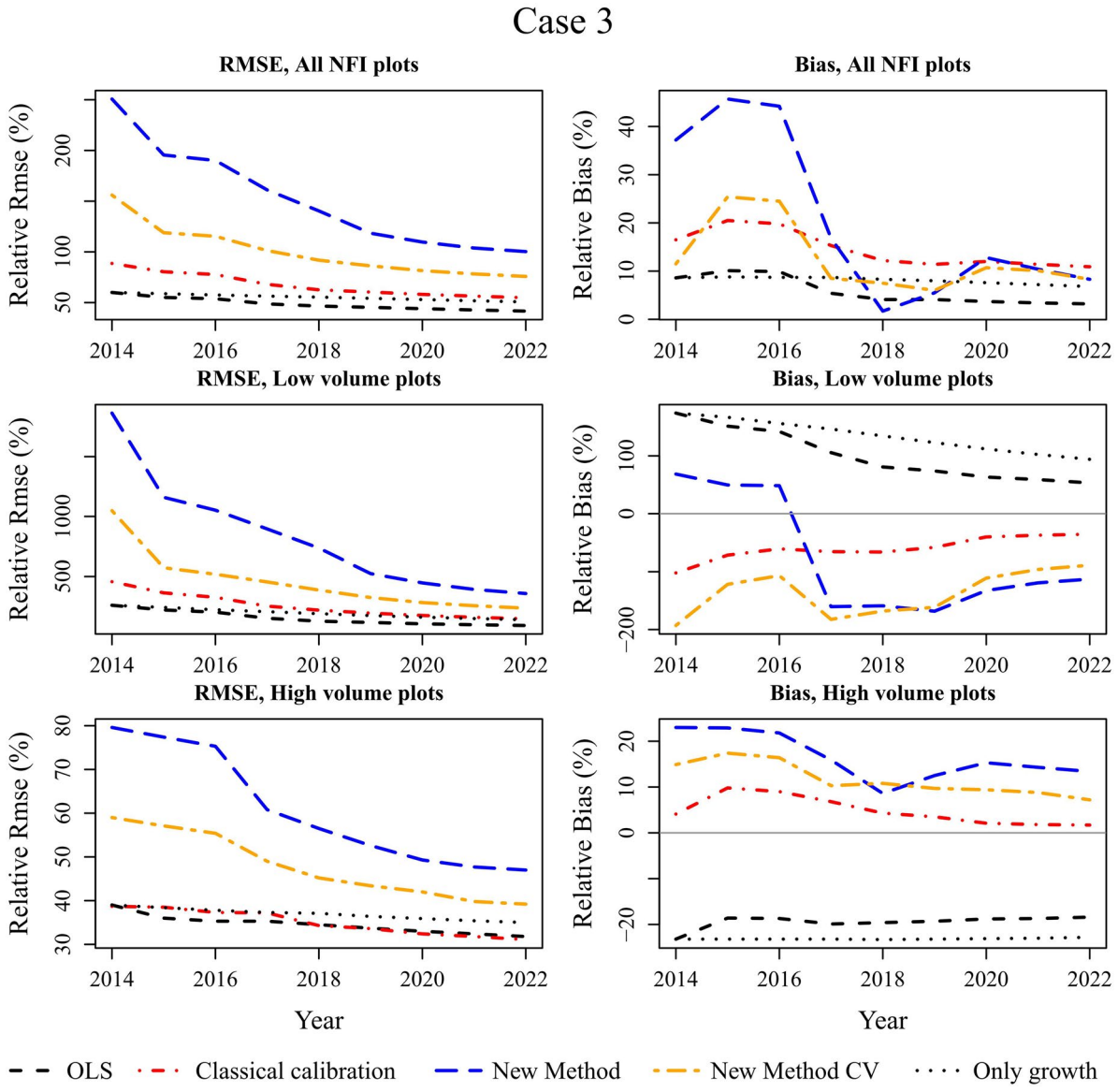

Figure 4. Results for Case 3, i.e., only optical satellite data were applied. Results are reported in terms of relative root mean square error (left) and estimated bias (right), separated on results for all NFI plots (upper panel), plots with low volumes (middle panel), and plots with high volumes (lower panel).

the improvements following application of the alternative approaches were modest, and non-existing when evaluations included all data.

For Case 3 (Figure 4), improvements of accuracy across time could be observed for all the DA methods, but the relative RMSEs remained very large and the improvements compared to using growth forecasts, only, from the initial OLS-based prediction were minor. Bias remained as a substantial problem, which could be mitigated only to some extent through the new methods or through OLS combined with classical calibration, for low and high true volumes.

In Figure 5, we demonstrate the outcome of DA in terms of wall-to-wall maps for the Krycklan validation area. In developing these maps all remote sensing data sources were used, together with the OLS method combined with classical calibration and change detection.

It can be observed that the general distribution of growing stock volume across the Krycklan validation area remained approximately the same between 2010 and 2022. Patterns of change due to growth were similar, with only some small harvested areas.

In Figure 6, we show results based on Krycklan validation data, in terms of state estimates in 2019, based on all three DA schemes (Cases 1 to 3). In total, there were 357 validation plots, with

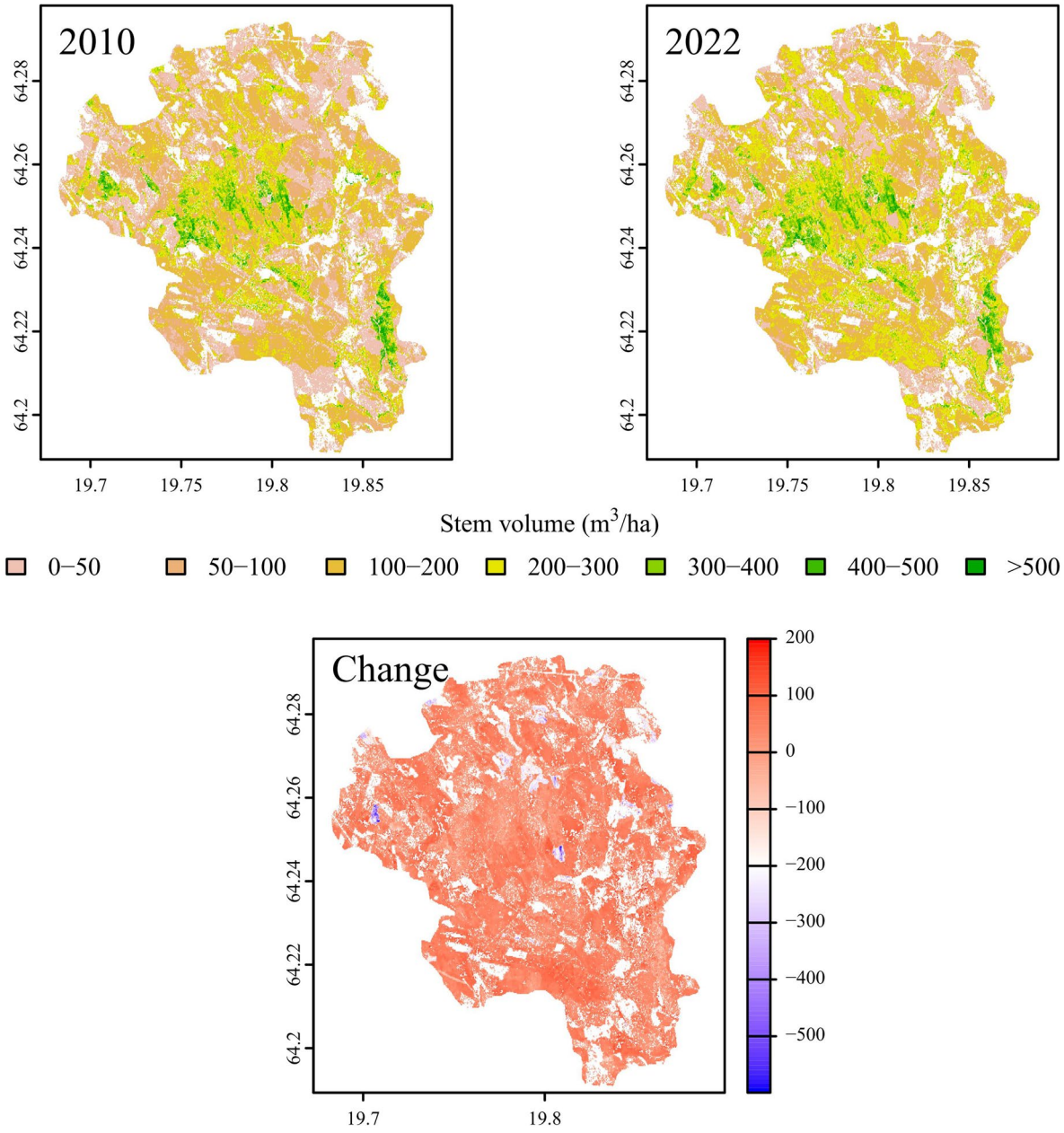

Figure 5. A demonstration of the results in terms of wall-to-wall maps of state, in 2010 and 2022, and change for the Krycklan validation area. The small blue areas are places where harvests had taken place.

97 (27%) being categorized as low volume and 128 (36%) categorized as high volume. Because this area had slightly higher volumes compared to the large study area, we categorized plots below 100 m^3/ha as low volume and above 200 m^3/ha as high volume.

The results from the Krycklan validation area were similar to the results from the large study area, although a difference was that the new regression methods performed slightly better, especially for Cases 1 and 2. The problem with substantial bias following OLS regression at low and high true volumes was apparent also for the Krycklan validation area.

Discussion

Our study confirmed the findings from previous similar studies (e.g., Lindgren et al. 2022a), that DA has a potential to maintain or slightly improve the accuracy of growing stock volume predictions across time. Compared to using only growth forecasts from an initial ALS-based prediction, which

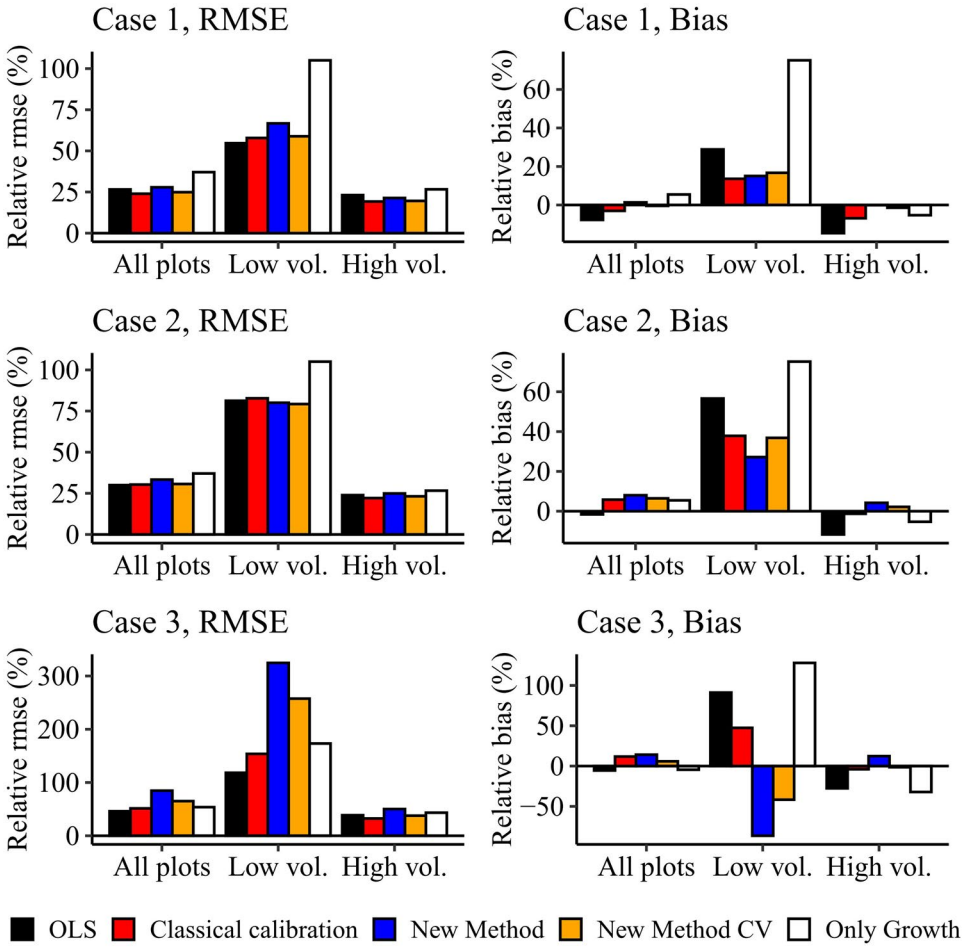

Figure 6. Results in terms of state estimates in 2019 based on independent validation data from Krycklan.

is often applied in practice, DA performed much better in terms of accuracy. The modest improvement of accuracy during the first part of the study period was due to the limited amounts of remotely sensed datasets during this period (see Table 2). Toward the end of the study period, where more remotely sensed datasets were available, the improvements were more pronounced, in line with what could be expected from applying DA. During this part of the study period several individual predictions of growing stock volume, at short intervals, contributed to improving the accuracy of the DA prediction of growing stock volume. Thus, we found that DA could be successfully applied as a method to keep forest maps of growing stock volume up to date, although with differences depending on what sequence of remotely sensed data was applied. With only optical satellite data, the accuracies remained poor.

Using standard prediction methods, i.e., OLS regression in our study, the problem of bias at low and high true volumes was substantial. The newly proposed regression methods in most cases decreased the bias compared to using OLS regression, although in terms of accuracy OLS regression combined with classical calibration was generally a better alternative. This indicates that further study on new prediction techniques could be motivated, e.g. by further exploring how to determine the number and size of windows within which the absolute average differences between observations and predictions are minimized. In the present study, choosing window sizes adaptively through cross validation mostly improved the results compared to using a fixed predetermined window size. Other methods to mitigate design-bias trends should be explored as well.

However, a problem that appears to be difficult to circumvent is that bias mitigation goes hand in hand with increased MSE. Thus, whether or not new regression techniques, or classical calibration, should be adopted depends on the intended use of the DA predictions. If the focus is to obtain best predictions for individual map units, only, bias correction should probably not be

used, whereas if it is important to maintain realistic variability of predicted values within target areas the situation is different. For example, as shown by Ulvdal et al. (2025) for applications in forest planning, biased estimates may cause considerable problems due to incorrect harvest schedulling.

Compared to Lindgren et al. (2022a), another novel feature of our study was that age and site quality were available wall-to-wall based on estimates from repeated acquisitions of ALS data. This is a substantial improvement compared to deriving these features from stand registries of variable quality, as has been the case in previous studies. With the new data at hand, the growth forecasts between recurring predictions based on remotely sensed data could be improved.

Another difference compared to Lindgren et al. (2022a) was that we did not estimate variances and covariances for the DA filter theoretically, but empirically through comparing the predictions with NFI data at all time points. This is a flexible and straightforward approach which, however, relies on sufficiently large sets of field plot data being available at all time points across the DA period. To manage this, we interpolated the state of permanent NFI plots for the years between measurements.

Regarding the methods applied, it may appear odd that we applied simple ordinary least squares regression as a baseline method instead of more advanced methods, such as generalized least squares (e.g., Mehtätalo and Lappi, 2020). However, for our application, using data from plots which were located at long distances from each other (i.e., there were no statistical dependencies between the observations), generalized least squares regression would have led to the same result as ordinary least squares regression (ibid., pp. 67–129).

Our evaluation was made mainly within forests undisturbed by harvesting and damage, although in the Krycklan validation area harvested areas were also identified and mapped. For any future practical DA system it would be important to include routines for identifying harvests and other disturbances and adjust predictions for such changes.

We conclude by suggesting that the proposed DA methodology appears to offer possibilities to maintain or slightly improve the quality of ALS-based predictions of growing stock volume across a relatively long period of time. However, an operational DA system would require fewer manual steps to be user-friendly and would probably need to be automated to a great extent. For example, a user-friendly operational DA system might independently check the availability of new remotely sensed data, perform quality checks, and estimate prediction models. As a consequence, we argue that substantial further development of DA methods for forest inventory remains before user-friendly DA systems for practical forestry might eventually become available.

Note

1. For 2016 we were not able to find a Sentinel 2 image with adequate quality for the study area.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Funding was provided by the Swedish Foundation for Strategic Environmental Research with the research program Mistra Digital Forest (DIA 2017/14 #6).

References

Bettinger, P., Boston, K., Siry, J. P., and Grebner, D. L. 2016. *Forest management and planning*. Academic press, London.

Cassel, C. M., Sarndal, C. E., and Wretman, J. H. 1977. *Foundations of inference in survey sampling*. Wiley, New York.

Draper, N. R., and Smith, H. 1998. *Applied Regression Analysis*. 3rd ed. New York: Wiley. doi:10.1002/9781118625590.
Ehlers, S., Grafström, A., Nyström, K., Olsson, H., and Ståhl, G. 2013. "Data assimilation in stand-level forest inventories." *Canadian Journal of Forest Research*, Vol. 43 (No. 12):pp. 1104–1113. doi:10.1139/cjfr-2013-0250.
Ehlers, S., Saarela, S., Lindgren, N., Lindberg, E., Nyström, M., Persson, H.J., Olsson, H., and Ståhl, G. 2018. "Assessing error correlations in remote sensing-based estimates of forest attributes for improved composite estimation." *Remote Sensing*, Vol. 10 (No. 5):pp. 667. doi:10.3390/rs10050667.
Evensen, G., Vossepoel, F. C., and Van Leeuwen, P. J. 2022. *Data assimilation fundamentals: A unified formulation of the state and parameter estimation problem* (p. 245). Springer Nature, Cham.
Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A.H., and Ståhl, G. 2014. "Adapting National Forest Inventories to changing requirements–the case of the Swedish National Forest Inventory at the turn of the 20th century." *Silva Fennica*, Vol. 48 (No. 3): doi:10.14214/sf.1095.
Hägglund, B., and Lundmark, J.E. 1977. "Site index estimation by means of site properties. Scots pine and Norway spruce in Sweden." *Studia Forestalia Suecica 138: 1-38.*.
Hou, Z., Domke, G.M., Russell, M.B., Coulston, J.W., Nelson, M.D., Xu, Q., and McRoberts, R.E. 2021. "Updating annual state-and county-level forest inventory estimates with data assimilation and FIA data." *Forest Ecology and Management*, Vol. 483 pp. 118777. doi:10.1016/j.foreco.2020.118777.
James, G., Witten, D., Hastie, T., and Tibshirani, R. 2021. *An Introduction to Statistical Learning: With Applications in R*. 2nd ed. Berlin: Springer. doi:10.1007/978-1-0716-1418-1.
Kalman, R.E. 1960. "A new approach to linear filtering and prediction problems."
Kalnay, E. 2003. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, Cambridge.
Lechner, A.M., Foody, G.M., and Boyd, D.S. 2020. "Applications in remote sensing to forest ecology and management." *One Earth*, Vol. 2 (No. 5):pp. 405–412. doi:10.1016/j.oneear.2020.05.001.
Lindgren, N., Olsson, H., Nyström, K., Nyström, M., and Ståhl, G. 2022a. "Data assimilation of growing stock volume using a sequence of remote sensing data from different sensors." *Canadian Journal Remote Sensors*, Vol. 48 (No. 2):pp. 127–143. doi:10.1080/07038992.2021.1988542.
Lindgren, N., Nyström, K., Saarela, S., Olsson, H., and Ståhl, G. 2022b. "Importance of calibration for improving the efficiency of data assimilation for predicting forest characteristics." *Remote Sensing*, Vol. 14 (No. 18): pp. 4627. doi:10.3390/rs14184627.
Mehtätalo, L., and Lappi, J. 2020. *Biometry for Forestry and Environmental Data: With Examples in R*. Chapman and Hall/CRC, New York.
Mensah, A.A., Holmström, E., Petersson, H., Nyström, K., Mason, E.G., and Nilsson, U. 2021. "The millennium shift: Investigating the relationship between environment and growth trends of Norway spruce and Scots pine in northern Europe." *Forest Ecology and Management*, Vol. 481 pp. 118727. doi:10.1016/j.foreco.2020.118727.
Mensah, A.A., Holmström, E., Nyström, K., and Nilsson, U. 2022. "Modelling potential yield capacity in conifers using Swedish long-term experiments." *Forest Ecology and Management*, Vol. 512 pp. 120162. doi:10.1016/j.foreco.2022.120162.
Mensah, A.A., Jonzén, J., Nyström, K., Wallerman, J., and Nilsson, M. 2023. "Mapping site index in coniferous forests using bi-temporal airborne laser scanning data and field data from the Swedish national forest inventory." *Forest Ecology and Management*, Vol. 547 pp. 121395. doi:10.1016/j.foreco.2023.121395.
Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., *et al.* 2017. "A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory." *Remote Sensing of Environment*, Vol. 194 pp. 447–454. doi:10.1016/j.rse.2016.10.022.
Noordermeer, L., Gobakken, T., Næsset, E., and Bollandsås, O.M. 2021. "Economic utility of 3D remote sensing data for estimation of site index in Nordic commercial forest inventories: a comparison of airborne laser scanning, digital aerial photogrammetry and conventional practices." *Scandinavian Journal of Forest Research*, Vol. 36 (No. 1):pp. 55–67. doi:10.1080/02827581.2020.1854340.
Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., *et al.* 2016. "A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle." *Geoscientific Model Development*, Vol. 9 (No. 9):pp. 3321–3346. (doi:10.5194/gmd-9-3321-2016.
Reese, H., Nilsson, M., Pahlén, T.G., Hagner, O., Joyce, S., Tingelöf, U., Egberth, M., and Olsson, H. 2003. "Countrywide estimates of forest variables using satellite data and field data from the national forest inventory." *AMBIO: A Journal of the Human Environment*, Vol. 32 (No. 8):pp. 542–548. doi:10.1639/0044-7447(2003)032[0542:ceofvu]2.0.co;2.
Ståhl, G., Gobakken, T., Saarela, S., Persson, H.J., Ekström, M., Healey, S.P., Yang, Z., *et al.* 2024. "Why ecosystem characteristics predicted from remotely sensed data are unbiased and biased at the same time–and how this affects applications." *Forest Ecosystems*, Vol. 11 (No. 1):pp. 100164. (doi:10.1016/j.fecs.2023.100164.
Swedish Forest Agency 2024. *Detecting harvesting*. http://www.skogsstyrelsen.se/laddanergeodata
Ulvdal, P., Ståhl, G., Sängstuvall, L., Eriksson, L.O., and Öhman, K. 2025. "Long-term strategic forest planning based on biased remote sensing predictions." *Forests Monitor*, Vol. 2 (No. 1):pp. 138–175. doi:10.62320/fm.v2i1.25.
Wikström, P., Edenius, L., Eriksson, L.O., Lämås, T., Sonesson, J., Öhman, K., Waller, C., and Klintebäck, F. 2011. "The Heureka forestry decision support system: an overview." *Mathematical and Computational Forestry & Natural Resource Sciences*, Vol. 3 (No. 2):pp. 87–95. *88).*