Remote Sensing of Environment 278 (2022) 113074

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

# Comparing frameworks for biomass prediction for the Global Ecosystem Dynamics Investigation

Svetlana Saarela [a,b,*], Sören Holm [b], Sean P. Healey [c], Paul L. Patterson [d], Zhiqiang Yang [c], Hans-Erik Andersen [e], Ralph O. Dubayah [f], Wenlu Qi [f], Laura I. Duncanson [f], John D. Armston [f], Terje Gobakken [a], Erik Næsset [a], Magnus Ekström [b,g], Göran Ståhl [b]

[a] *Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway*
[b] *Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden*
[c] *USDA Forest Service, Rocky Mountain Research Station, Ogden, UT, USA*
[d] *USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, USA*
[e] *USDA Forest Service, Pacific Northwest Research Station, Seattle, WA, USA*
[f] *Department of Geographical Sciences, University of Maryland, College Park, MD, USA*
[g] *Department of Statistics, USBE, Umeå University, Umeå, Sweden*

ARTICLE INFO

Edited by Marie Weiss

*Keywords:*
Carbon monitoring
GEDI
Hierarchical model-based inference
Hybrid inference
Mean square error
Model-based inference
Remote sensing
TanDEM-X

ABSTRACT

NASA's Global Ecosystem Dynamics Investigation (GEDI) mission offers data for temperate and pan-tropical estimates of aboveground forest biomass (AGB). The spaceborne, full-waveform LiDAR from GEDI provides sample footprints of canopy structure, expected to cover about 4% of the land area following two years of operation. Several options are available for estimating AGB at different geographical scales. Using GEDI sample data alone, gridded biomass predictions are based on hybrid inference which correctly propagates errors due to the modeling and accounts for sampling variability, but this method requires at least two GEDI tracks in the area of interest. However, there are significant gaps in GEDI coverage and in some areas of interest GEDI data may need to be combined with other wall-to-wall remotely sensed (RS) data, such as those from multispectral or SAR sensors. In these cases, we may employ hierarchical model-based (HMB) inference that correctly considers the additional model errors that result from relating GEDI data to the wall-to-wall data. Where predictions are possible from both hybrid and HMB inference the question arises which framework to choose, and under what circumstances? In this paper, we make progress towards answering these questions by comparing the performance of the two prediction frameworks under conditions relevant for the GEDI mission. Conventional model-based (MB) inference with wall-to-wall TanDEM-X data was applied as a baseline prediction framework, which does not involve GEDI data at all.

An important feature of the study was the comparison of AGB predictors in terms of both standard deviation (SD: the square root of variance) and root mean square error (RMSE: the square root of mean square error – MSE). Since, in model-based inference, the true AGB in an area of interest is a random variable, comparisons of the performance of prediction frameworks should preferably be made in terms of their RMSEs. However, in practice only the SD can be estimated based on empirical survey data, and thus it is important also to study whether or not the difference between the two uncertainty measures is small or large under conditions relevant for the GEDI mission.

Our main findings were that: (i) hybrid and HMB prediction typically resulted in smaller RMSEs than conventional MB prediction although the difference between the three frameworks in terms of SD often was small; (ii) in most cases the difference between hybrid and HMB inference was small in terms of both RMSE and SD; (iii) the RMSEs for all frameworks was substantially larger than the SDs in small study areas whereas the two uncertainty measures were similar in large study areas, and; (iv) spatial autocorrelation of model residual errors had a large effect on the RMSEs of AGB predictors, especially in small study areas. We conclude that hybrid inference is suitable in most GEDI applications for AGB assessment, due to its simplicity compared to HMB inference. However, where GEDI data are sparse HMB inference should be preferred.

* Corresponding author at: P.O. Box 5003, NMBU, 1432, Ås, Norway.
*E-mail address:* svetlana.saarela@nmbu.no (S. Saarela).

https://doi.org/10.1016/j.rse.2022.113074
Received 7 February 2021; Received in revised form 2 February 2022; Accepted 1 May 2022
Available online 24 May 2022

## 1. Introduction

It is well known that LiDAR 3D data provide an excellent basis for assessing aboveground biomass (AGB) in forests (e.g., Solodukhin et al., 1977; Nelson et al., 1988; Hyyppä and Hallikainen, 1996; Hyyppä and Inkinen, 1999; Næsset, 2002). Consequently, many countries nowadays apply airborne LiDAR technology for stand-level forest inventories guiding management decisions (e.g., Næsset, 2004). LiDAR data have also been evaluated for large-area surveys, such as national forest inventories (e.g., Andersen et al., 2009; Asner, 2001; Gobakken et al., 2012; Saarela et al., 2015). A restriction in such cases is that wall-to-wall coverage of airborne laser scanning data is expensive, which has inspired the development of sample-based approaches (e.g.,Nelson et al., 2004, 2012; Gregoire et al., 2011).

Another option to reduce cost per area unit is to launch satellite-borne LiDARs. NASA's ICESat-I mission (Zwally et al., 2002) provided global coverage of LiDAR sample data with large footprints. Although primarily intended for assessment of ice cover, several studies have applied ICESat data for assessing AGB across large areas (e.g., Boudreau et al., 2008; Healey et al., 2012; Neigh et al., 2013; Holm et al., 2017). In 2019, NASA's Global Ecosystem Dynamics Investigation (GEDI; Dubayah et al., 2014, 2020) started to acquire LiDAR waveform data across the tropical, subtropical, and temperate biogeographical zones, with a primary focus to support AGB assessments (Healey et al., 2021; Duncanson et al., 2022; Dubayah et al., 2022). GEDI acquires data as samples with footprints of about 25 m diameter. Following two years of operation, the sample will cover about 1-5% of the land area within the GEDI domain (Fig. 1).

Several preparatory studies have applied simulated GEDI data (e.g., Patterson et al., 2019; Qi et al., 2019) that are derived from airborne laser scanning data within a simulator developed by Hancock et al. (2019). First studies employing on-orbit GEDI data include Healey et al. (2020) and Potapov et al. (2021), who developed global forest height maps by combining Landsat and GEDI data, Rishmawi et al. (2021) who derived forest structure indicators across the USA and Mexico by combining GEDI and NOAA data, and Healey et al. (2021), who developed the GEDI gridded biomass product.

Many remote sensing studies assess the accuracy of predictions through pairwise comparison of predicted values and field references at the level of typical assessment units, such as plots or stands (e.g., Persson, and Ståhl, 2020). Results are often reported as empirical root mean square errors based on the differences between predictions and field references. As pointed out by Gregoire et al. (2016), there is no straightforward way to extend such uncertainty assessments to uncertainties of predicted means and totals for larger study regions, which are required for evaluating whether or not results can be trusted in, e.g., national policy processes or reports according to international agreements. Thus, analysts face several challenges when combining remotely sensed (RS) and field data for large-area surveys of AGB where uncertainties should be reported correctly. This involves selecting the appropriate inferential framework, i.e. design-based or model-based (Gregoire, 1998), and handling all major sources of uncertainty under the framework chosen. Whereas design-based inference requires fairly large probability samples of field plots, in addition to RS data, model-based inference does not and may do well with less costly purposive field sampling (e.g., Ståhl et al., 2016; Magnussen et al., 2016).

In some studies, multiple sources of hierarchically nested RS data have been used in model-based surveys of AGB resources (e.g., Nelson et al., 2009; Neigh et al., 2013). A statistical framework for handling the uncertainty assessment in such cases was developed by Saarela et al. (2016), and later further developed by Saarela et al. (2018, 2020) for application under a wider range of conditions.

The GEDI mission was designed to specifically use these statistical frameworks for providing mean and uncertainty estimates for AGB at the resolution of its 1-km gridded data product (Patterson et al., 2019; Healey et al., 2021) and for arbitrarily sized regions at the national scale and beyond. Using the GEDI sample data alone, the mission approach is to use hybrid inference (e.g., Ståhl et al., 2011; Corona et al., 2014; Ståhl et al., 2014; Ståhl et al., 2016; Patterson et al., 2019), which predicts the average AGB in the area of interest as the mean of model predictions based on GEDI's footprint samples. The corresponding uncertainty is the sum of a sampling variability component and two model error components (Ståhl et al., 2011; Patterson et al., 2019). The properties of hybrid inference in connection with the GEDI mission have been thoroughly investigated by Patterson et al. (2019). Hybrid inference requires at least two tracks in a 1-km grid-cell or any larger area.

A first question arises as to how to estimate mean and uncertainty if there is only one track or no tracks in a grid-cell. GEDI's current coverage has many gaps at the scale of 1 km cells. In this case an alternative to hybrid inference is to combine GEDI data with other RS data available wall-to-wall, such as demonstrated here using data from the TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement)

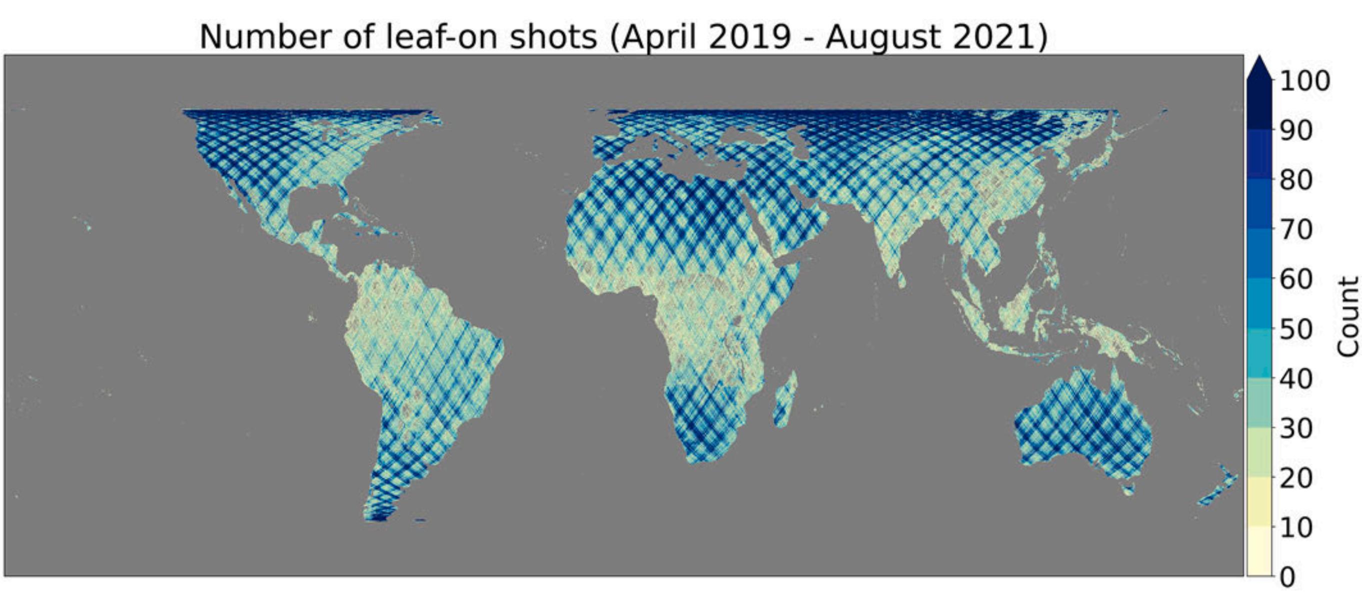


**Fig. 1.** An overview of the available total number of GEDI shots per 1-km tile over land surfaces; for deciduous forests and woodlands, only leaf-on shots are included in the count. The striped pattern and the lower density close to the equator is due to the orbital resonance of the international space station.

twin satellite. This can be achieved through hierarchical model-based inference (HMB: Saarela et al., 2016), where GEDI-based AGB predictions are used for local training of a TanDEM-X model (cf., Healey et al., 2020), which is then subsequently used for predicting AGB density for areas that do not have at least two tracks. In this case, the overall prediction uncertainty is due to four model error components, linked to parameter estimation uncertainty and model residual errors for the two models involved.

A second questions arises concerning the efficacy of hybrid vs. HMB where there may be sufficient GEDI tracks to formally invoke hybrid but the coverage is sparse. In those cases is it better to use a hybrid estimate or an HMB estimate leveraged the wall-to-wall observation of another sensor, and how to decide between the two?

Finally, a third question is whether models that use the wall-to-wall data by themselves, without reference to GEDI observations, are significantly less accurate than using hybrid or HMB with GEDI data, and under what conditions.

Thus, the main goal of this study is to compare hybrid inference and HMB inference under survey conditions relevant for the GEDI mission to evaluate which method should be preferred under different conditions. For this purpose, we applied case study data with simulated GEDI data, TanDEM-X data, and field data from areas in the western USA and Costa Rica. A novel feature of the study is the comparison of AGB predictors in terms of both variance and mean square error (MSE). Variance is a measure of uncertainty where the variability of a predictor is related to its expected value. MSE is a measure where the variability is related to the true AGB value in a study area, and thus it is a more relevant measure of uncertainty than variance in this type of surveys. However, only the variance can be estimated based on empirical survey data in practical applications (cf., Ståhl et al., 2016), since assessing MSE requires large amounts of field data from each study area. Thus, it is important to study whether or not the difference between the two measures is negligible or significant under conditions relevant for the GEDI mission, because the uncertainty measure reported from a survey application will be the variance (or to be precise, its estimate). As a baseline methodological alternative, we also evaluated the variance and MSE of conventional model-based inference (Cassel et al., 1977; McRoberts, 2010a) using wall-to-wall TanDEM-X data alone.

An important technical part of the study is the development of an MSE formula for hybrid inference and a new MSE formula for HMB inference following Saarela et al. (2020), reflecting the contribution from the model residual errors in both models. Further, the MSE and variance formulas are partitioned into components, whereby it is possible to assess the impact of different factors related to study area conditions and survey parameters on variances and MSEs. To support the analysis, an R function `MethComp()` available in the HMB R package (Saarela et al., 2022) was developed and applied; it allows users to perform comparisons using their own data and assumptions.

The remainder of this paper is organized as follows. We first describe the RS and field data used in the analyses. We then present our methodological approaches that allow for comparison between hybrid and HMB, including details of derivation of MSE and variance formulas, computational considerations and key assumptions. Results of our analyses are given next as a function of area size, spatial correlation of errors, the number of field plots GEDI coverage, and correlation among variables. Lastly, we frame our discussion around the central question of when to use hybrid, HMB or conventional model-based inference, and the conditions that may factor into this consideration.

## 2. Material and methods

### 2.1. Data

Empirical data were available from two different study areas: the Teakettle Experimental Forest (TEF) in the Western Sierra Nevada of California and the La Selva Biological Station (LSBS) in the Atlantic lowlands of Costa Rica. Fig. 2 presents an overview of the sites and the data.

For each site, we have field data, simulated GEDI data, and TanDEM-X data. Basic characteristics of the two study sites are given in Table 1. Field data were collected in 2008 at the Californian site (TEF) and in 2009 for the site in Costa Rica (LSBS), as described by Qi et al. (2019).

The simulation of medium footprint (25 m) GEDI full-waveform data was performed by reprocessing small-footprint discrete return airborne LiDAR data collected in September 2008 at TEF and in September – October 2009 at LSBS, using the GEDI waveform simulator (Hancock et al., 2019). Relative height metrics at 2% intervals were subsequently derived from these pseudo-waveforms (Qi et al., 2019). GEDI groundtrack patterns were simulated based on the expected number of times the International Space Station would pass each site for the full two-year mission period. One TanDEM-X dataset was selected for each site based on acquisition time, to minimize the temporal difference compared to the airborne LiDAR data acquisition, and spatial baseline to ensure an appropriate height of ambiguity across the site. Only HH (signal transmitted horizontally and received horizontally) polarization data were used to reflect the greater availability of HH data globally. The spatial resolution of TanDEM-X raster data was 25 m. For more details about the study sites, field data, simulated GEDI data, and TanDEM-X data, see Qi et al. (2019).

### 2.2. Methods

The comparison of methods was founded on empirical data from the two study areas. In addition, we performed sensitivity analyses where we varied important survey and study area features, such as GEDI sampling fraction and study area size. The analyses focused on comparing the MSEs of the different methods for predicting AGB density, and on studying differences between MSEs and variances for each of the methods, since only the variance can normally be estimated in practical survey applications (e.g., Ståhl et al., 2016). We apply the general convention in statistics of *predicting* random variables and *estimating* fixed parameters. In model-based inference, the target AGB density in a study area is a random variable, and thus we apply predictors for predicting it (rather than estimators for estimating it, which would be the case in design-based inference). However, usage of these terms tends to differ slightly between authors (e.g., Cassel et al., 1977).

For a predictor, $\widehat{\mu}$, of AGB density the mean square error is defined as $\text{MSE}(\widehat{\mu}) = E[(\widehat{\mu} - \mu)^2]$, where $\mu$ is the true AGB density, which is a random variable in model-based inference. The variance is defined as $V(\widehat{\mu}) = E[(\widehat{\mu} - E[\widehat{\mu}])^2]$. Thus, whereas the MSE measures the performance of a predictor in relation to the true value, the variance measures its performance in relation to its expected value (e.g., Cassel et al., 1977). This difference has important implications in model-based inference, where a model may fit well in some study areas and poorly in other, implying that the variance may be small although the MSE is large. Thus, the relevant measure to base comparisons on is the MSE rather than the variance, although Ståhl et al. (2016) and Patterson et al. (2019) have argued that for large study areas the difference, in relative terms, between MSE and variance would be small, and McRoberts et al. (2018) empirically demonstrated this. In the present study we evaluated the performance of predictors over small to large areas of interest.

For predictors applied in the framework of conventional model-based inference, the MSE contains two basic components: one reflects the precision whereby the parameters in the model are estimated and the other how well the estimated model fits the particular target area. The first component depends on how the model is specified, the number of observations used for estimating it, and the correlation between AGB and RS data. The latter component involves the size and spatial correlation of model residual errors in the study area (e.g., Cassel et al., 1977). In hybrid inference, an additional term adds to the MSE. This term reflects sampling variability, since RS data in this case are not

 

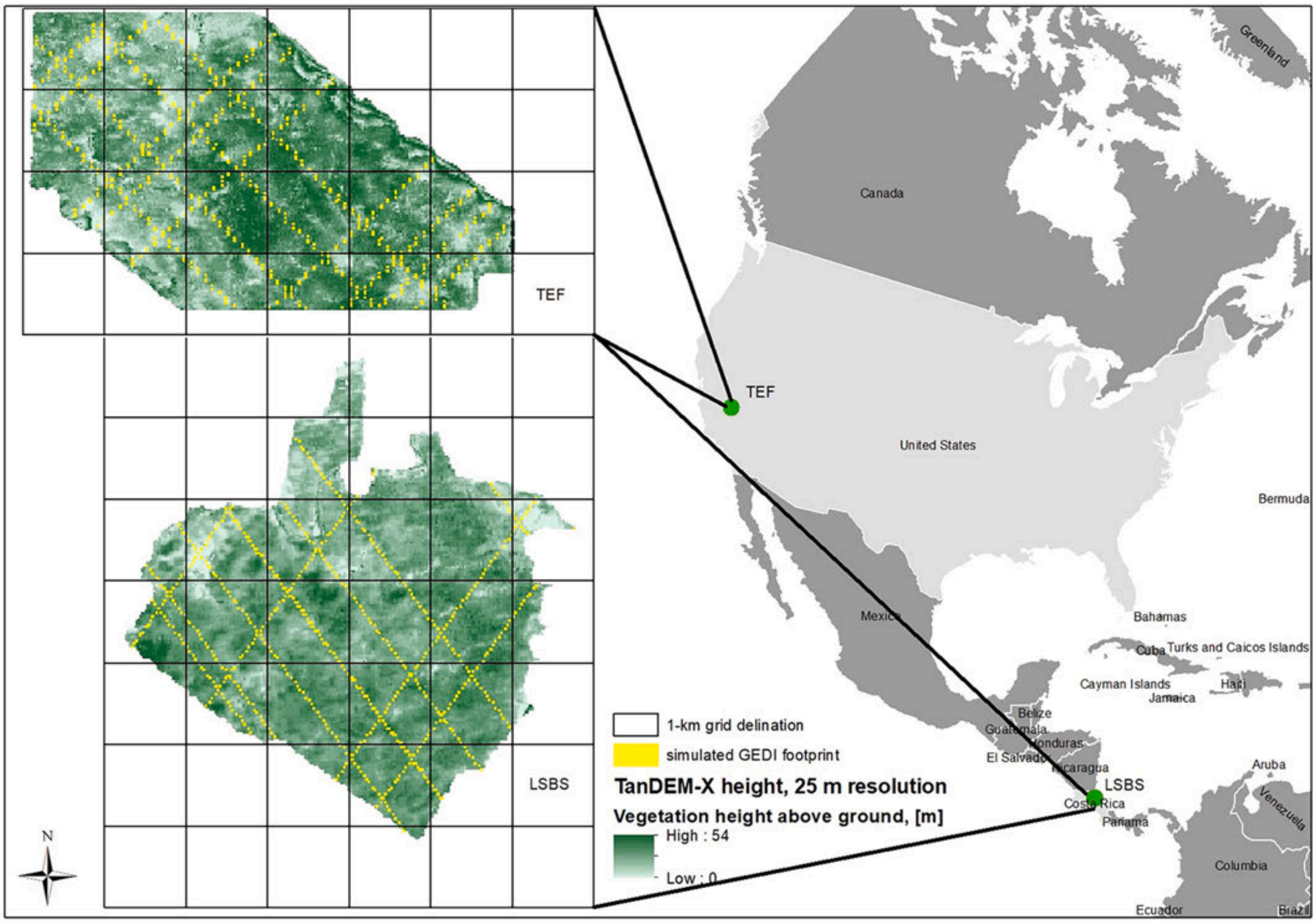


**Fig. 2.** Study area locations.

**Table 1**
Details about the study areas (Average AGB is based on predictions using hybrid inference.).

| Study area | Number of field plots | Number of simulated GEDI footprints | Average AGB, [Mg·ha$^{-1}$] | Area of interest, size | |
|---|---|---|---|---|---|
| | | | | in pixels | in hectares, [ha] |
| TEF (California, USA) | 105 | 715 | 426 | 25383 | 1586 |
| LSBS (Costa Rica) | 170 | 637 | 214 | 26523 | 1658 |

available wall-to-wall. Comparing HMB inference with conventional model-based inference, the MSE of a predictor will involve error components from two models rather than from one. However, since a large set of "pseudo-field" data (the AGB predictions based on GEDI data) is available for estimating the model parameters, the resulting MSE may still be small for HMB inference. For details, see Appendix A.

Formulas for MSE and variance have been derived for conventional model-based inference, hybrid inference, and HMB inference in previous studies (e.g., Cassel et al., 1977; Ståhl et al., 2011; Ståhl et al., 2014; Saarela et al., 2016, 2018, 2020; McRoberts et al., 2018). However, these formulas provide limited insight about how different study area and survey features affect MSEs and variances. Therefore, in this study the formulas were decomposed into terms that could be intuitively understood, and varied to identify under what conditions different methodological approaches should be preferred. Thus, although MSE formulas are available from previous studies, an important technical part of the study was to develop a set of variance and MSE formulas in formats suitable for the purpose of the study. To achieve this, we applied simple linear models in all cases. Whereas this would be a restriction in practical application, it should be sufficient for comparing the general performance of the different methods. Further, we assumed that the GEDI footprints could be treated as a simple random sample although in reality the footprints are acquired in linear-shaped clusters, which in ensemble resemble a spatially systematic sample. This simplified treatment of systematic sampling, i.e. to treat the sampling units as being selected through simple random sampling, is commonly applied in practical surveys (Gregoire, and Valentine, 2008; Magnussen et al., 2020).

The details of the derivation of MSE and variance formulas for the three methods compared in the study are given in Appendix A. Fig. 3 provides a graphical overview of the methods and sources of data applied in the three methodological frameworks compared.

### 2.3. Computational approach

To compare the performance of the three methods under different conditions we based the computations on the empirical conditions in the two study areas, and then performed sensitivity analyses where different survey and study area features were varied while others were held constant. This type of analysis was possible through applying the new formulas for variance and MSE, as described in Appendix A.

The analyses were separated into three modules, focusing on

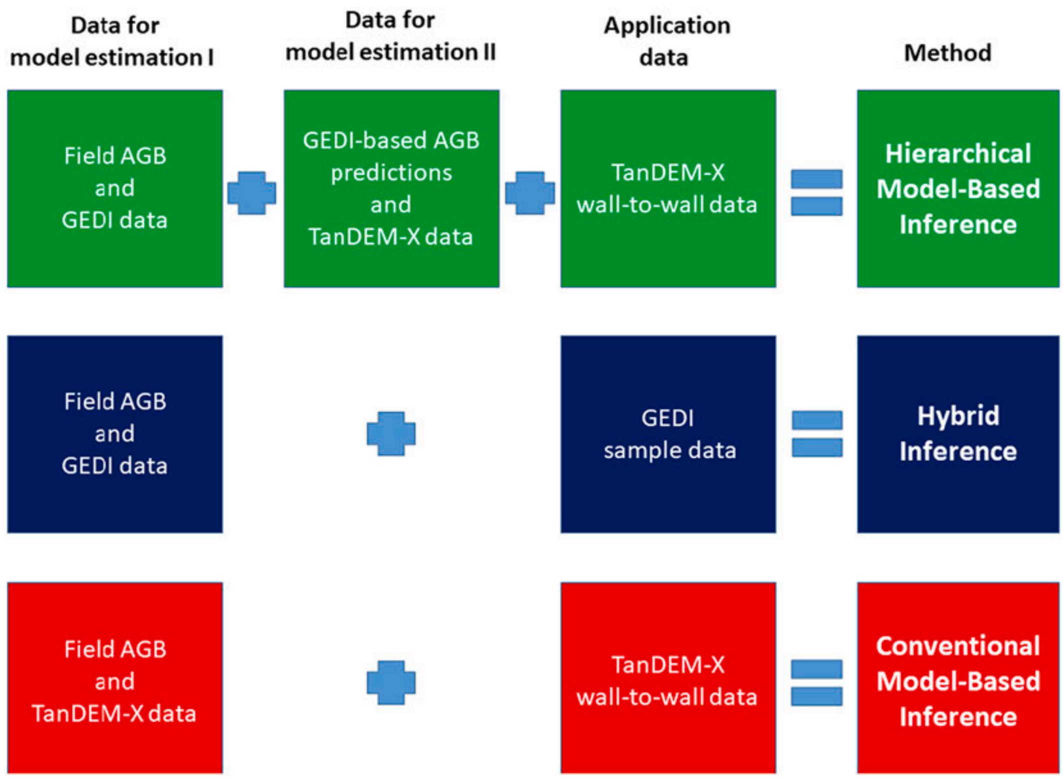


**Fig. 3.** Graphical overview of the methods compared.

different key features regarding the comparison of the three methods. The modules are:

I *Study area size and spatial autocorrelation of model errors*

It is known from previous studies that variance and MSE should be of similar magnitude in large areas of interest (e.g., Ståhl et al., 2016; Patterson et al., 2019). However, for small and intermediately large areas it is important to assess the magnitude of the difference between these two uncertainty measures, since only the variance can be estimated in practical surveys. Further, the accuracy of AGB density predictors depends on the number of GEDI samples in the area of interest, which varies with study area size at a given GEDI sampling fraction. If a survey aims at reporting relative MSEs smaller than a certain threshold, it is important to assess how large the study area needs to be before the threshold is reached.

Another important factor in this type of analysis is to what extent neighboring population elements (pixels) are similar with regard to model errors, i.e. to what extent the model errors are spatially correlated. With strong spatial autocorrelation, the MSEs of AGB predictors will be larger, and thus the areas of interest have to be larger before the magnitudes of variances and MSEs tend to be similar. The analyses under this module thus also sought to assess how large areas would be required under different spatial autocorrelation assumptions, for MSEs and variances to be equally large.

II *Number of field plots and GEDI sampling fraction*

The number of field plots available affects the precision whereby the parameters of the models can be estimated. Thus, the more field plots, the better the accuracy of predictors in model-based inference would be. However, field plots tend to be a sparse resource in practical surveys, and thus a pertinent question is how many plots are needed to obtain "fair" results in terms of MSE of AGB predictors.

The GEDI sampling fraction, i.e. the proportion of the area of interest covered by GEDI footprints, is important for both hybrid and HMB inference. An important issue is whether or not there is a breakpoint, in terms of GEDI sampling fraction, where HMB inference should be preferred over hybrid inference.

III *Correlation between variables*

The strength of the correlation between GEDI and field AGB data, as well as between TanDEM-X and field AGB data, is an important factor in model-based inference. With strong correlation, model-based inference performs very well whereas the opposite holds for weak correlation. In this module, we evaluated MSEs and variances for different correlation assumptions. The analysis also involved the correlation between GEDI and TanDEM-X data, since this is an important factor in HMB inference.

### 2.3.1. *Baseline assumptions*

Baseline settings for the analyses are given in Table 2. The correlations between variables were derived from the empirical study area data, whereas the baselines for other features are motivated in the text, below. During the analyses, one of these factors at the time was varied while the others were kept at their baseline levels. However, as explained in detail in Appendix A, in some cases a factor could not be varied independently from other factors and in such cases the consequential change in other factors, implied by varying the study factor, had to be taken into account.

The baseline study area size (100 ha) corresponds to the size of tiles for which basic GEDI estimates will be presented (e.g., Dubayah et al., 2014, 2020; Patterson et al., 2019). The baseline GEDI sampling fraction (2%) was motivated by the expected coverage of GEDI data, following two years of operation (Patterson et al., 2019; Dubayah et al., 2020).

Spatial autocorrelation of model residual errors was specified as the correlation between two neighboring map units. Few empirical studies are available regarding such correlations in the context of forest inventories based on RS data (e.g., McRoberts et al., 2018). However,

**Table 2**
Baseline settings used in the computations for the two study areas.

| Baseline settings | | TEF (Califorina, USA) | LSBS (Costa Rica) |
|---|---|---|---|
| Dataset sizes: | Number of field plots | 100 | 100 |
| | GEDI sampling fraction | 2% | 2% |
| | Area support factor for HMB | 9 | 9 |
| | Study area size | 100 ha | 100 ha |
| | Population unit (pixel) size | 625 $m^2$ | 625 $m^2$ |
| Residual errors spatial autocorrelation: | Field AGB – GEDI model | 0.7 | 0.7 |
| | Field AGB – TanDEM-X model | 0.9 | 0.9 |
| | GEDI AGB – TanDEM-X model | 0.8 | 0.8 |
| Correlation between variables: | Field AGB and GEDI | 0.715 | 0.613 |
| | Field AGB and TanDEM-X | 0.349 | 0.103 |
| | GEDI and TanDEM-X | 0.748 | 0.637 |

strong correlation between the explanatory variable and the study variable normally implies weak spatial autocorrelation of model errors, although the actual distribution of forest types across the study area is also influential. With spatial autocorrelation 0.8 between neighboring units, the correlation is about 0.01 between units at 20 units distance, i.e. 500 meters in this study. With spatial autocorrelation 0.9 the corresponding distance is about 40 map units, i.e. 1000 meters. We assumed that the spatial autocorrelation of model errors followed an isotropic process, based solely on the geographical distance between population units.

In HMB, prediction models based on TanDEM-X data are estimated from "pseudo-field" AGB data obtained through GEDI predictions. In estimating such models, it is straightforward to use data not only from within the area of interest but also from neighboring similar areas (cf., Saarela et al., 2018). The "area support factor" determines from what geographical neighborhood data are collected for estimating the model linking pseudo-field AGB with TanDEM-X data. In this study it was set to 9, i.e. HMB prediction models were developed based on GEDI and TanDEM-X data from the actual target area, and in addition from 8 neighboring (square) units equally large as the target area.

### 2.4. Presentation of results

In the results chapter, we present RMSEs (the square root of MSE) and standard deviations (SDs: the square root of the variance) for the AGB predictors in relative terms as

$$\text{rRMSE} = 100\% \times \frac{\sqrt{\text{MSE}(\hat{\mu})}}{\text{mean AGB}} \tag{1}$$

and

$$\text{rSD} = 100\% \times \frac{\sqrt{V(\hat{\mu})}}{\text{mean AGB}}, \tag{2}$$

where $\hat{\mu}$ is the predicted AGB density and *mean AGB* is the mean AGB density from the field plots in the study area.

## 3. Results

The structure of the results section follows the three study modules, i.e. first we address effects of study area size and spatial correlation of model errors, secondly effects of number of field plots and GEDI sampling fraction, and thirdly effects of different correlations between the variables involved. Regarding notation in the figures, "HMB" denotes hierarchical model-based inference, "HY" hybrid inference, and "MB" conventional model-based inference.

### 3.1. Study area size and spatial correlation of model errors

As seen in Fig. 4, the size of the study area of interest has substantial impact on the relative RMSEs and the relative SDs. A main result is that the rRMSEs for hybrid and HMB prediction are almost the same across the entire range of study area sizes, for both study areas. The rRMSEs decrease from more than 30% for small areas to about 5% for large areas. For hybrid inference, the difference between rRMSE and rSD is always relatively small; for HMB it is large for small study areas, and for conventional model-based inference the difference is very large, especially for small study areas.

The effect of spatial autocorrelation of model errors is shown in Fig. 5. It can be observed that for hybrid and HMB inference, the difference between rRMSE and rSD begins to be substantial at spatial correlations larger than about 0.7 for the model linking GEDI and AGB data. For conventional model-based inference a similar result was obtained in the LSBS study area whereas for the TEF study area a large difference between rRMSE and rSD can be observed already when the spatial autocorrelation is about 0.4. For the model linking GEDI and TanDEM-X data, affecting only HMB inference, the impact of spatial autocorrelation of model errors was less pronounced compared to the previously described cases.

### 3.2. Number of field plots and GEDI sampling fraction

In Fig. 6, results are presented for different numbers of field plots used for estimating the parameters of the models linking GEDI data and field AGB, and TanDEM-X data and field AGB. As expected, the uncertainties are reduced when more field plots are available, especially for the rSD uncertainty measure. However, the reduction for the rRMSE measure is less pronounced, and major differences in uncertainties in this case are seen only at relatively small numbers of field plots.

The performance of hybrid and HMB prediction depends on the GEDI sampling fraction (Fig. 7). It can be observed that the rSD and the rRMSE of hybrid inference are closely linked to the GEDI sampling fraction, as expected. However, since HMB utilizes GEDI data also from the neighborhood of the study area, it is less affected by limited availability of GEDI footprints, and for sparse GEDI samples HMB is superior to hybrid inference in terms of rRMSE. The breakpoint was about 2% in the TEF study area and about 1.5% in the LSBS study area

### 3.3. Correlation between variables

Three different correlations affect the performance of the methods studied: the correlation between field AGB and the GEDI variable, the correlation between field AGB and the TanDEM-X variable, and the correlation between the GEDI and TanDEM-X variables. Fig. 8 shows that the stronger the correlation is between field AGB and GEDI data, the smaller rSDs and rRMSEs are obtained for hybrid and HMB predictors. A 10% rRMSE level was reached when this correlation was about 0.75 in

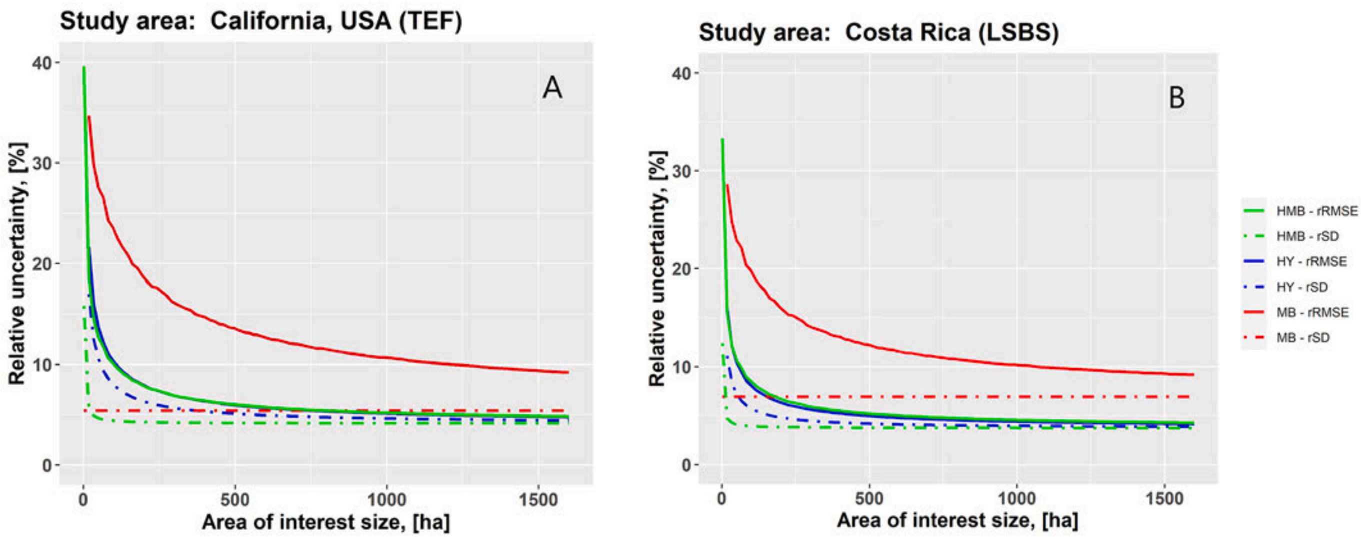


**Fig. 4.** rRMSE and rSD across study area size; the left panel shows results for the California (TEF) study area, the right for the Costa Rica (LSBS) study area. The baseline settings are given in Table 2.

the TEF study area and about 0.5 in the LSBS study area. The correlation between field AGB and TanDEM-X data affected the performance of conventional model-based prediction differently in the two study areas. A 10% rRMSE level was reached when the correlation was about 0.6 in the TEF study area, and 0.25 in the LSBS study area. Further, the correlation between GEDI and TanDEM-X data greatly affects the performance of HMB prediction. A 10% rRMSE was reached when this correlation was about 0.7 in the TEF study area and about 0.5 in the LSBS study area.

The results in terms of rRMSE and rSD were found to be very similar for the hybrid and HMB methods when varying the strength of the correlation between GEDI and field AGB. For conventional model-based inference there was a large difference in results between the two study areas. Whereas in the TEF study area the correlation between TanDEM-X and AGB data would have to be stronger than about 0.6 for the method to be superior to the other two in terms of rRMSE, in the LSBS study area the threshold correlation was 0.25. A strong correlation between GEDI and TanDEM-X data also substantially increases the accuracy of HMB estimation.

## 4. Discussion

### 4.1. Hybrid or HMB inference?

The comparison of hybrid and HMB inference under conditions relevant for the GEDI mission (Dubayah et al., 2020) revealed many interesting properties of the methods. Firstly, it was observed that none of the methods could be singled out as generally better than the other; the results were case dependent. However, at a 2% GEDI sampling fraction (the baseline assumption) the two methods performed about equally well. With sparser GEDI data, HMB should be preferred whereas the opposite holds for denser GEDI samples. Although not specifically evaluated in the study, the GEDI sampling fraction for which the two methods could be expected to perform equally well depends on the size of the case study area. For large study areas, the threshold should be smaller and for small areas it should be larger. Thus, for very large study areas hybrid inference would always be the preferred method. However, for small and intermediately large study areas, and sparse GEDI samples, the advantage of HMB is that it can utilize GEDI data from neighboring areas for estimating the TanDEM-X model that is applied across the study area for predicting AGB density. Thus, this model can be adapted to local conditions which is an obvious advantage compared to developing and applying a global model, which would be the case in conventional model-based inference (Healey et al., 2020).

A practical advantage of hybrid inference compared to HMB inference is its simplicity, since it involves only one type of RS data and one model, whereas HMB requires two. As a rule of thumb from this study, for small and moderately large study areas hybrid inference should be preferred over HMB inference when the GEDI sampling fraction exceeds 2%.

### 4.2. Variance or MSE?

MSE (or rRMSE) is the relevant measure to base comparisons of predictors on (e.g., McRoberts et al., 2018), since it is a measures of uncertainty in relation to the actual AGB density in a study area. Variance (or rSD) on the other hand, is a measure of uncertainty in relation to the expected value of a predictor; in model-based inference, the variance does not take into account the random variability of the true AGB densities (cf., Cassel et al., 1977). Thus, if the model fit in a study area is poor this will not be reflected in the variance. However, as noted in the introduction, only the variance and not the MSE can be estimated and reported in practical applications (e.g., Ståhl et al., 2016; Saarela et al., 2016; McRoberts et al., 2018).

Large differences between rRMSEs and rSDs were obtained in several cases, especially for small study areas and for cases with strong spatial autocorrelation of model residual errors. Strong autocorrelations of this type typically emanate from models based on data with weak correlation between the explanatory variable and the response variable. Using such models in AGB density predictors, the rSD may be very small whereas the rRMSE is large, especially if the model parameters are estimated from a large sample of field plots. In extreme cases, a model might always predict a mean value, regardless of the input RS data for the map units. Obviously, the usefulness of such models is limited (e.g., Healey et al., 2020). In the study, large differences between rRMSEs and rSDs

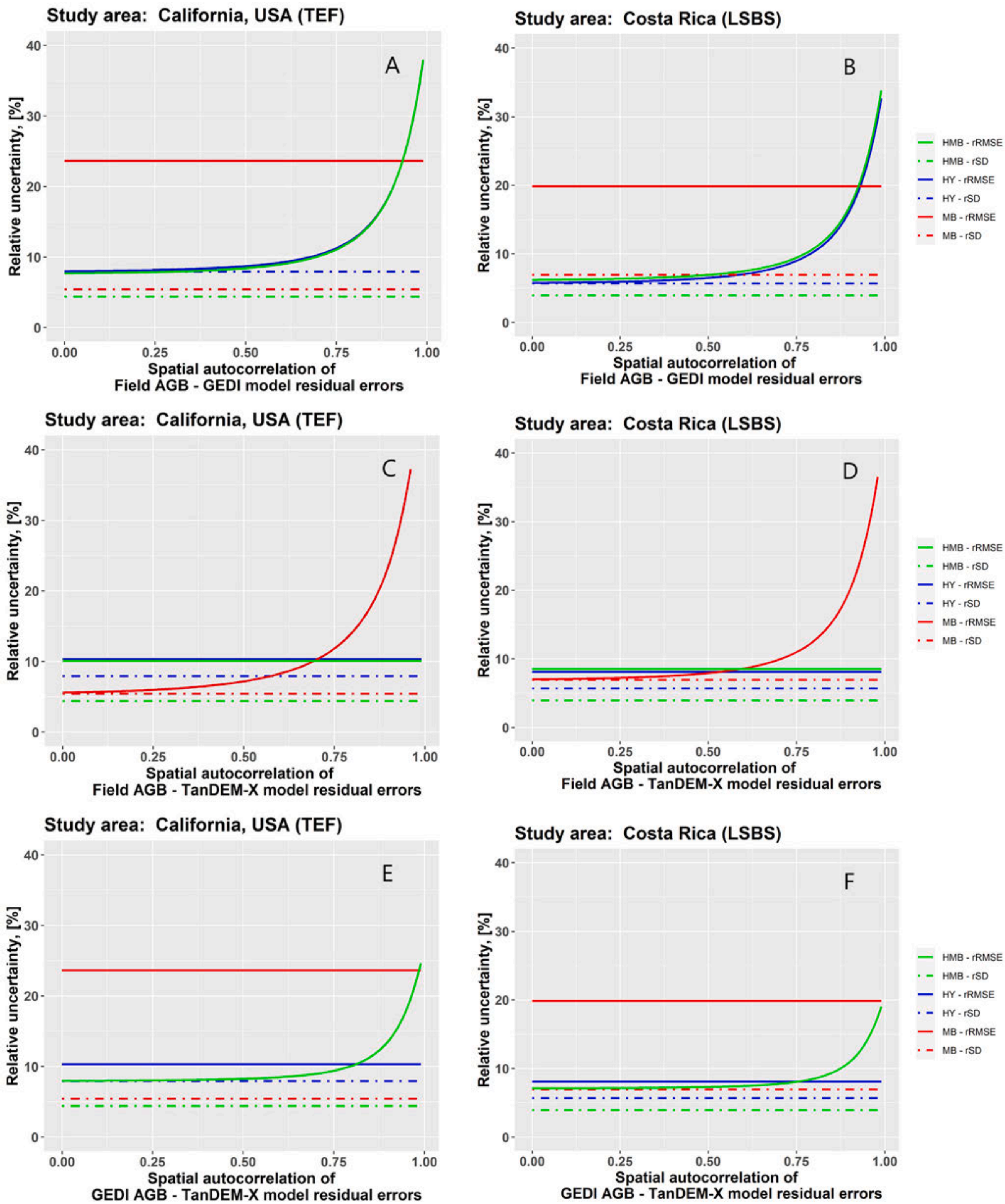


**Fig. 5.** rRMSE and rSD across different spatial autocorrelations of model residual errors; the left panel shows results for the California (TEF) study are, the right for the Costa Rica (LSBS) study area. The baseline settings are given in Table 2.

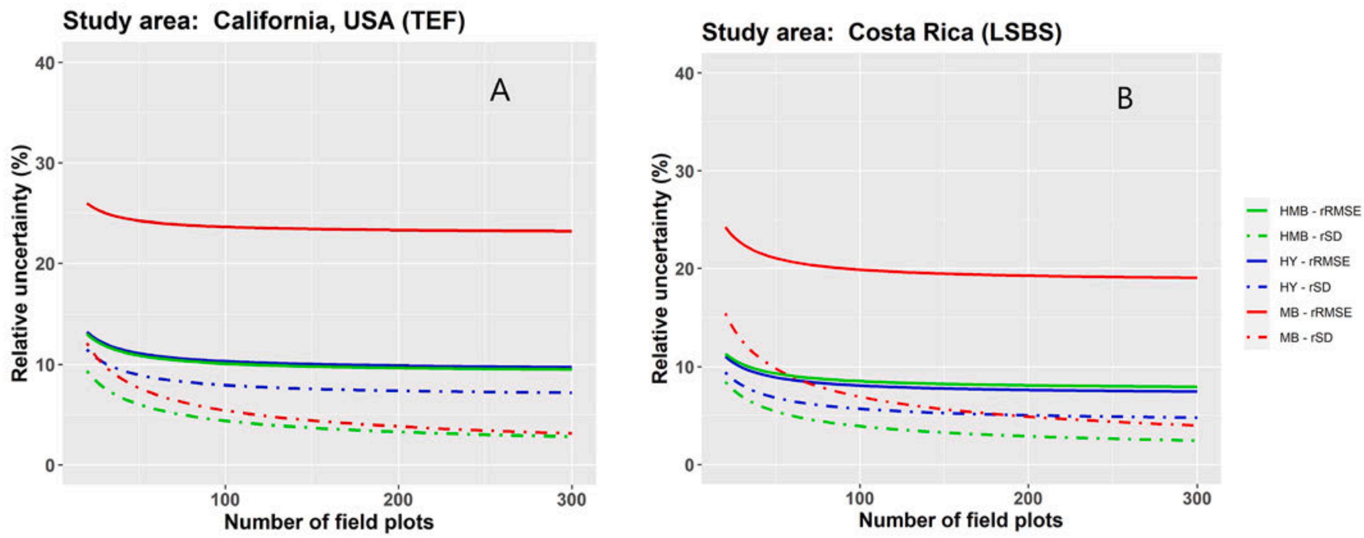


**Fig. 6.** rRMSE and rSD for different numbers of field plots; the left panel shows results for the California (TEF) study area, the right for the Costa Rica (LSBS) study area. The baseline settings are given in Table 2.

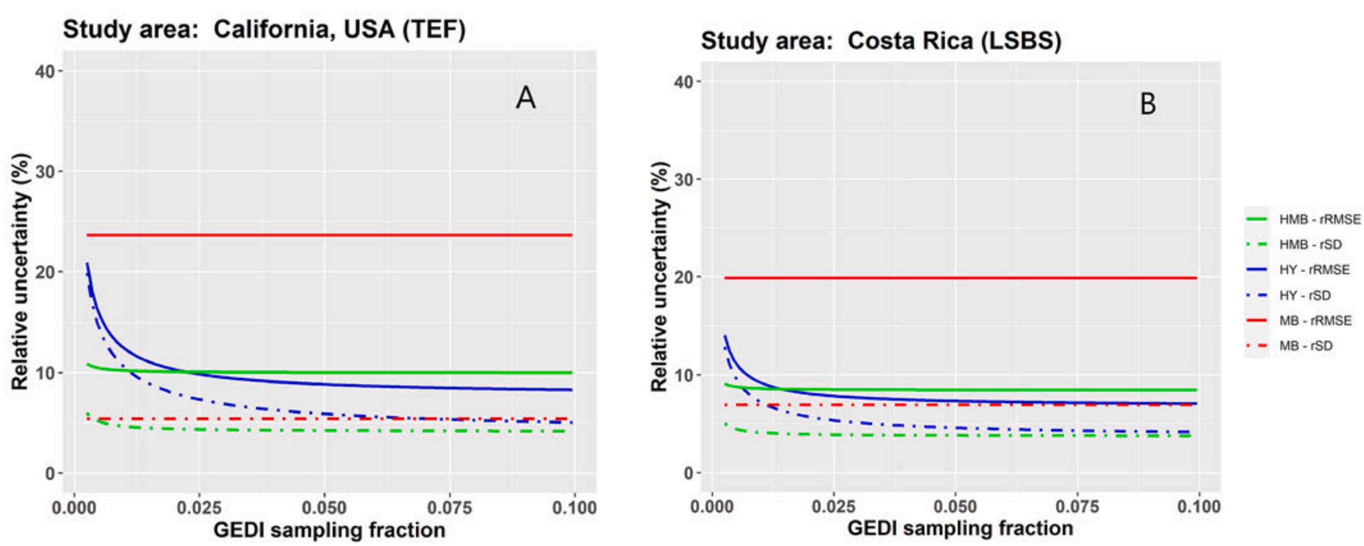


**Fig. 7.** rRMSE and rSD over different GEDI sampling fractions; the left panel shows results for the California (TEF) study area, the right for the Costa Rica (LSBS) study area. The baseline settings are given in Table 2.

were often obtained for conventional model-based inference, especially for small study areas. Further, this difference was larger for HMB than for hybrid inference, since HMB involves model residual errors from two different models. However, for very large study areas it can be theoretically argued and empirically shown (e.g., Ståhl et al., 2016; McRoberts et al., 2018; Patterson et al., 2019) that the difference between MSE and variance will be small and thus the variance is a good approximation of the MSE. For the study area size magnitudes investigated in this study, this could be observed for hybrid and HMB inference, but the difference remained large for conventional model-based inference at all area sizes investigated.

It is sometimes of general interest to assess the contribution of different components to the total MSE in predicting AGB density. For hybrid and HMB inference, the contribution of field AGB – GEDI model errors to the total MSE was evaluated. For hybrid inference, the portion of the total MSE due to model errors was found to be largely depending on the GEDI sampling fraction. With small sampling fractions the sampling error portion was large, whereas for large sampling fractions the MSE was almost entirely due to model errors. For HMB inference two models are involved. Typically, about 50-75% of the total MSE was found to be due to AGB - GEDI model errors. The large error portion emanating from this model is partly due to the propagation of errors from the first model through the second model (e.g., Saarela et al., 2018).

### 4.3. Strength of the correlation between field AGB and RS data

In the study, the baseline correlations between field AGB and RS data were taken from the empirical data in the study areas. However, when

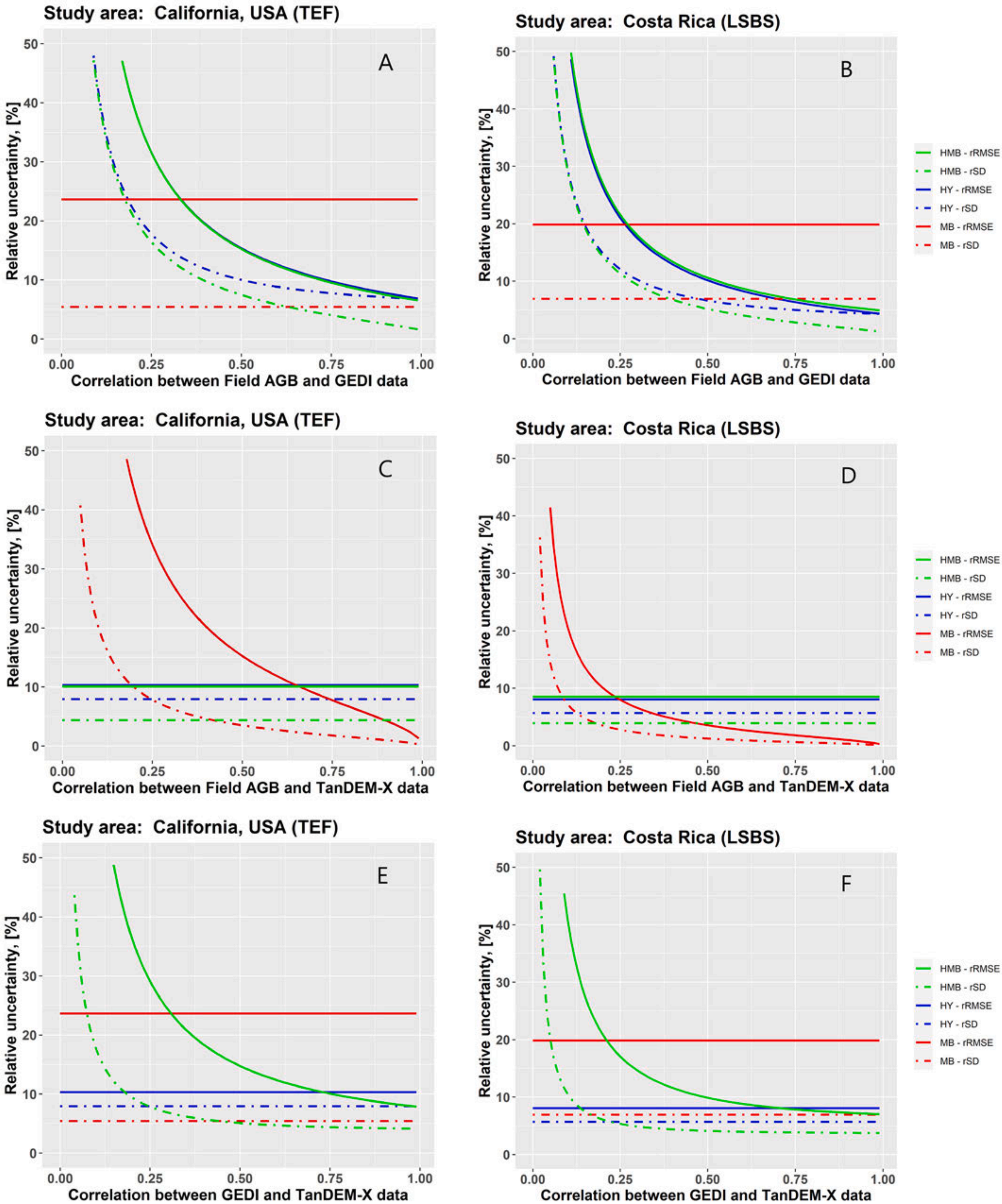


**Fig. 8.** rRMSE and rSD across different correlations between variables; the left panel shows results for the California (TEF) study area, the right for the Costa Rica (LSBS) study area. The baseline settings are given in Table 2.

performing the sensitivity analysis where the assumption about these correlations was varied, substantial changes in results were observed (Fig. 8). Obviously, with very strong correlations all methods perform very well, not least conventional model-based inference. However, the logic behind the evaluations made was that GEDI data correlate more strongly with field AGB than other types of RS data.

In the study areas, the correlations between field AGB and RS data were found to be rather weak and in the LSBS study area it was surprising to note a rather strong correlation between GEDI and TanDEM-X data, whereas the correlation between field AGB and TanDEM-X data was weak. The reason for this is unclear, but may partly be due to the accuracy by which field plots were geo-located (McRoberts, 2010b; Milenković et al., 2017).

Evaluations under conditions different from the ones used in this study can be undertaken using the function `MethComp()` in the R package HMB (Saarela et al., 2022), which produces results similar to those presented in the results section of this study, for survey and study area conditions specified by the user.

### 4.4. Conclusions for the GEDI mission

The study shows that both hybrid and HMB inference are strong options for predicting AGB density in connection with the GEDI mission. Hybrid inference should be preferred over HMB inference when the study areas are large. For small and intermediately large study areas (order of magnitude 100-1000 ha) HMB should be preferred over hybrid inference when the GEDI sampling fraction is small; the study suggests a threshold of about 2%. However, a user should be aware that the variance may be substantially smaller than the MSE when HMB is applied in small study areas, and thus reported confidence intervals (based on variance estimates) may be too optimistic.

## CRediT authorship contribution statement

**Svetlana Saarela:** Conceptualization, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing, Software. **Sören Holm:** Investigation, Methodology, Writing – review & editing. **Sean P. Healey:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Writing – review & editing. **Paul L. Patterson:** Data curation, Investigation, Writing – review & editing. **Zhiqiang Yang:** Data curation, Investigation, Writing – review & editing, Software. **Hans-Erik Andersen:** Data curation, Investigation, Writing – review & editing. **Ralph O. Dubayah:** Data curation, Investigation, Funding acquisition, Writing – review & editing, Project administration. **Wenlu Qi:** Data curation, Investigation, Writing – review & editing. **Laura I. Duncanson:** Data curation, Investigation, Writing – review & editing. **John D. Armston:** Data curation, Investigation, Writing – review & editing. **Terje Gobakken:** Investigation, Writing – review & editing. **Erik Næsset:**Conceptualization, Investigation, Writing – review & editing. **Magnus Ekström:** Investigation, Writing – review & editing. **Göran Ståhl:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing, Software.

## 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.


## Acknowledgment

Funding was provided by grants from NASA's Carbon Monitoring System (grant 80HQTR18T0016, grant 80HQTR21T0013), NASA GEDI Inter-Agency Agreement (RPO201523), the Swedish Kempe Foundation (Kempestiftelserna: SMK-1847), the Swedish Research Council for Sustainable Development (FORMAS: FR-2019/0007) and the Swedish National Space Agency (SNSA-171/19). The simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the High Performance Computing Center North (HPC2N). The authors are thankful to the Editor and two anonymous Reviewers, whose comments helped us to improve the clarity of the article.


## Appendix A. Derivation of MSEs and variances for hybrid, model-based and hierarchical model-based inference

### A.1. Hybrid inference

We denote an area of interest (AOI) as $U$, with $N$ population units (grid-cells). A sample of $M$ GEDI footprints available in the AOI is denoted as $Sa$. A dataset with field AGB and GEDI variables is denoted $S$ and has $n$ units.

A regression model 'F' is given:

$$\mathrm{F}:\ \mathbf{y} = \beta_0 + \beta_1\mathbf{x} + \boldsymbol{\varepsilon};\ ,\ \boldsymbol{\varepsilon};\ \sim N(\mathbf{0}, \omega^2\mathbf{I}), \tag{A.1}$$

where $\mathbf{y}$ is a vector of field AGB values, $\beta_0$ and $\beta_1$ are model parameters to be estimated, $\mathbf{x}$ is a vector of the GEDI variable, and $\boldsymbol{\varepsilon};$ is a vector of residual errors with a constant variance $\omega^2$ (cf., Mehtatalo and Lappi, 2020).

The AOI population mean $\mu$ under simple random sampling without replacement is predicted as

$$\begin{aligned}\widehat{\mu}_{\mathrm{HY}} &= \widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\overline{\mathrm{x}}_{\mathrm{Sa}} \\ &= \overline{\mathbf{x}}_{\mathrm{Sa}}\widehat{\boldsymbol{\beta}}_S,\end{aligned} \tag{A.2}$$

where the subscript $HY$ denotes "hybrid", $\widehat{\beta}_{0_S}$ and $\widehat{\beta}_{1_S}$ are estimated model parameters based on the dataset $S$, $\overline{\mathbf{x}}_{\mathrm{Sa}} = \begin{pmatrix} 1 & \overline{x}_{\mathrm{Sa}} \end{pmatrix}$ is the mean of the GEDI variable (including a vector of units) over the sample $Sa$.

The true mean value can be expressed through the model F as

$$\mu = \beta_0 + \beta_1 \bar{x}_U + \bar{\varepsilon}_{;U}, \tag{A.3}$$

where $\bar{x}_U$ is the mean of the GEDI variable over the AOI and $\bar{\varepsilon}_{;U}$ is the mean of random residual errors over the AOI.

The expected value of the hybrid predictor $\widehat{\mu}_{\mathrm{HY}}$ can be derived using the law of total expectation, i.e.

$$\mathrm{E}[\widehat{\mu}_{\mathrm{HY}}] = \mathrm{E}_D\ [\mathrm{E}_F\ [\widehat{\mu}_{\mathrm{HY}}]] \tag{A.4}$$

where $\mathrm{E}_D\ [\widehat{\mu}_{\mathrm{HY}}]$ is an expectation of the hybrid population mean predictor due to the sampling design, and $\mathrm{E}_F\ [\widehat{\mu}_{\mathrm{HY}}]$ is an expectation due to the model F.

We express the true population mean using the hybrid predictor's expectation as

$$\begin{aligned}\mu\ &= \beta_0 + \beta_1 \bar{x}_U + \bar{\varepsilon}_{;U}\\ &= \mathrm{E}_F\ \left[\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right] + \bar{\varepsilon}_{;U}\\ &= \mathrm{E}_D\ \left[\mathrm{E}_F\ \left[\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_{\mathrm{Sa}}\right]\right] + \mathrm{B}_D[\widehat{\mu}_{\mathrm{HY}}] + \bar{\varepsilon}_{;U}\\ &= \mathrm{E}_D\ [\mathrm{E}_F\ [\widehat{\mu}_{\mathrm{HY}}]] + \mathrm{B}_D[\widehat{\mu}_{\mathrm{HY}}] + \bar{\varepsilon}_{;U}\\ &= \mathrm{E}[\widehat{\mu}_{\mathrm{HY}}] + \mathrm{B}_D[\widehat{\mu}_{\mathrm{HY}}] + \bar{\varepsilon}_{;U},\end{aligned} \tag{A.5}$$

where $\mathrm{B}_D[\widehat{\mu}_{\mathrm{HY}}]$ is a potential bias of the hybrid population mean predictor due to the sampling design.

Thus, we can derive an expression of the MSE of the hybrid predictor $\widehat{\mu}_{\mathrm{HY}}$ as

$$\begin{aligned}\mathrm{MSE}(\widehat{\mu}_{\mathrm{HY}})\ &= \mathrm{E}\left[(\widehat{\mu}_{\mathrm{HY}} - \mu)^2\right]\\ &= \mathrm{E}\left[(\widehat{\mu}_{\mathrm{HY}} - \mathrm{E}[\widehat{\mu}_{\mathrm{HY}}] - \mathrm{B}_D[\widehat{\mu}_{\mathrm{HY}}] - \bar{\varepsilon}_{;U})^2\right]\\ &= \mathrm{V}(\widehat{\mu}_{\mathrm{HY}}) + \mathrm{B}_D^2[\widehat{\mu}_{\mathrm{HY}}] + \mathrm{V}(\bar{\varepsilon}_{;U}).\end{aligned} \tag{A.6}$$

The formula is valid under the assumption of independence between $\widehat{\mu}_{\mathrm{HY}}$ and $\bar{\varepsilon}_{;U}$. If the hybrid predictor $\widehat{\mu}_{\mathrm{HY}}$ is design-unbiased, as in our study, then the MSE expression Eq. (A.6) reduces to

$$\mathrm{MSE}(\widehat{\mu}_{\mathrm{HY}}) = \mathrm{V}(\widehat{\mu}_{\mathrm{HY}}) + \mathrm{V}(\bar{\varepsilon}_{;U}). \tag{A.7}$$

Using the law of total variances we can decompose $\mathrm{V}(\widehat{\mu}_{\mathrm{HY}})$ further

$$\begin{aligned}\mathrm{V}(\widehat{\mu}_{\mathrm{HY}})\ &= \mathrm{E}_D\ [\mathrm{V}_F\ (\widehat{\mu}_{\mathrm{HY}})] + \mathrm{V}_D\ (\mathrm{E}_F\ [\widehat{\mu}_{\mathrm{HY}}])\\ &= \mathrm{E}_D\ \left[\mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_{\mathrm{Sa}}\right)\right] + \mathrm{V}_D\ \left(\mathrm{E}_F\ \left[\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_{\mathrm{Sa}}\right]\right)\\ &= \mathrm{E}_D\ \left[\mathrm{V}_F\ \left(\widehat{\beta}_{0_S}\right) + \bar{x}_{\mathrm{Sa}}^2\mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right) + 2\bar{x}_{\mathrm{Sa}}\mathrm{Cov}\left(\widehat{\beta}_{0_S}, \widehat{\beta}_{1_S}\right)\right] + \mathrm{V}_D\ (\beta_0 + \beta_1\bar{x}_{\mathrm{Sa}})\\ &= \mathrm{V}_F\ \left(\widehat{\beta}_{0_S}\right) + \mathrm{E}_D\ [\bar{x}_{\mathrm{Sa}}^2]\mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right) + 2\bar{x}_U\mathrm{Cov}\left(\widehat{\beta}_{0_S}, \widehat{\beta}_{1_S}\right) + \beta_1^2\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}})\\ &= \mathrm{V}_F\ \left(\widehat{\beta}_{0_S}\right) + (\bar{x}_U^2 + \mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}}))\mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right) + 2\bar{x}_U\mathrm{Cov}\left(\widehat{\beta}_{0_S}, \widehat{\beta}_{1_S}\right) + \beta_1^2\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}})\\ &= \mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right) + (\beta_1^2 + \mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right))\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}})\\ &= \mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right) + \mathrm{E}_F\ \left[\widehat{\beta}_{1_S}^2\right]\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}}),\end{aligned} \tag{A.8}$$

where $\mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right)$ is due to the estimated model parameters using the model F; $\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}})$ is due to the probability design to sample the GEDI variable, and $\mathrm{E}_F\ \left[\widehat{\beta}_{1_S}^2\right] = (\beta_1^2 + \mathrm{V}_F\ (\widehat{\beta}_{1_S}))$.

Thus, the MSE of the AOI population mean predictor $\widehat{\mu}_{\mathrm{HY}}$ can be expressed as

$$\mathrm{MSE}(\widehat{\mu}_{\mathrm{HY}}) = \mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right) + (\beta_1^2 + \mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right))\mathrm{V}_D\ (\bar{x}_{\mathrm{Sa}}) + \mathrm{V}(\bar{\varepsilon}_{;U}). \tag{A.9}$$

The first term of the expression is the variance due to the estimated model parameters $\widehat{\boldsymbol{\beta}}_S$, the second term is the variance due to the sampled *Sa* data from the target AOI under the simple random sampling without replacement design multiplied by $\mathrm{E}_F\ \left[\widehat{\beta}_{1_S}^2\right]$, and the third term is the variance due to the residual errors in model F.

We now elaborate the first term $\mathrm{V}_F\ \left(\widehat{\beta}_{0_S} + \widehat{\beta}_{1_S}\bar{x}_U\right)$ to the generic expression

$$
\begin{aligned}
\mathrm{V}_F\left(\widehat{\beta}_{0_S}+\widehat{\beta}_{1_S}\overline{x}_U\right) &= \overline{\mathrm{x}}_U\mathrm{Cov}(\widehat{\boldsymbol{\beta}}_S)\overline{\mathrm{x}}_U^{\top} \\
&= \begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}\mathrm{Cov}(\widehat{\boldsymbol{\beta}}_S)\begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}^{\top} \\
&= \omega^2\begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}(\mathrm{X}_S^{\top}\mathrm{X}_S)^{-1}\begin{pmatrix} 1 \\ \overline{x}_U \end{pmatrix} \\
&= \omega^2\begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}\begin{pmatrix} n & \sum_{i=1}^{n}x_i \\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^2 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ \overline{x}_U \end{pmatrix} \\
&= \frac{\omega^2}{n\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}x_i}\begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}\begin{pmatrix} \sum_{i=1}^{n}x_i^2 & -\sum_{i=1}^{n}x_i \\ -\sum_{i=1}^{n}x_i & n \end{pmatrix}\begin{pmatrix} 1 \\ \overline{x}_U \end{pmatrix} \\
&= \frac{\omega^2}{n\sum_{i=1}^{n}(x_i-\overline{x}_S)^2}\begin{pmatrix} 1 & \overline{x}_U \end{pmatrix}\begin{pmatrix} n\overline{x^2}_S & -n\overline{x}_S \\ -n\overline{x}_S & n \end{pmatrix}\begin{pmatrix} 1 \\ \overline{x}_U \end{pmatrix} \\
&= \frac{\omega^2}{n}\frac{\sum_{i=1}^{n}(x_i-\overline{x}_U)^2}{\sum_{i=1}^{n}(x_i-\overline{x}_S)^2} \\
&= \frac{\omega^2}{n}\left(1+\frac{(\overline{x}_S-\overline{x}_U)^2}{\sigma_{x_S}^2}\right),
\end{aligned}
\tag{A.10}
$$

where $\overline{x}_S$ is the mean value of the GEDI variable over dataset $S$, and $\sigma_{x_S}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_S)^2$ is the variance of the GEDI variable over $S$.

In the second term of the MSE($\widehat{\mu}_{\mathrm{HY}}$) expression Eq. (A.9), the variance due to the probability sample of GEDI data under a simple random sampling without replacement design can be rewritten as

$$
\mathrm{V}_D\left(\overline{x}_{\mathrm{Sa}}\right)=\left(1-\frac{M}{N}\right)\frac{\sigma_{x_U}^2}{M}, \tag{A.11}
$$

where $\sigma_{x_U}^2=\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x}_U)^2$ is the population variance of the GEDI variable.

The third term in the MSE($\widehat{\mu}_{\mathrm{HY}}$) expression Eq. (A.9) is the variance of the mean of the residual errors, which can be rewritten as

$$
\begin{aligned}
\mathrm{V}(\overline{\varepsilon_{;}}_U) &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Cov}\left(\varepsilon_{;i},\varepsilon_{;j}\right) \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Corr}\left(\varepsilon_{;i},\varepsilon_{;j}\right)\sqrt{\mathrm{V}(\varepsilon_{;i})\mathrm{V}\left(\varepsilon_{;j}\right)}.
\end{aligned}
\tag{A.12}
$$

With homoscedastic and independent residual errors, the variance $\mathrm{V}(\overline{\varepsilon_{;}}_U)$ is

$$
\mathrm{V}(\overline{\varepsilon_{;}}_U) = \frac{\omega^2}{N}. \tag{A.13}
$$

However, in case of spatial autocorrelation between residual errors, the variance is

$$
\mathrm{V}(\overline{\varepsilon_{;}}_U)=\frac{\omega^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Corr}\left(\varepsilon_{;i},\varepsilon_{;j}\right). \tag{A.14}
$$

In module I, we analyzed the predictors' sensitivity to the magnitude of spatial autocorrelation. We modeled the autocorrelation through the power function,

$$
\mathrm{V}(\overline{\varepsilon_{;}}_U)=\frac{\omega^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_F^{\mathrm{dist}_{i,j}}, \tag{A.15}
$$

where $\rho_F$ is a spatial autocorrelation between two neighboring units and given in Table 2 for Field AGB – GEDI model, and $dist_{i,j}$ is a distance between the *ith* and *jth* units.

In module III, we showed the effect of different correlations between variables on the uncertainty of the prediction methods. To accommodate this, we expressed the residual error variance $\omega^2$ through the correlation between the field AGB and GEDI variables, $\mathrm{Corr}(y,x)$, (cf., Eq. 7 Helland, 1987):

$$\omega^2 = \beta_1^2\sigma_{x_S}^2\frac{1-\mathrm{Corr}^2\ (y,x)}{\mathrm{Corr}^2\ (y,x)}. \tag{A.16}$$

And thus, the entire $\mathrm{MSE}(\widehat{\mu}_{\mathrm{HY}})$ expression decomposed into elements is

$$\begin{aligned}\mathrm{MSE}(\widehat{\mu}_{\mathrm{HY}}) &= \frac{(\beta_1^2\sigma_{x_S}^2\frac{1-\mathrm{Corr}^2\ (y,x)}{\mathrm{Corr}^2\ (y,x)})}{n}(1+\mathrm{DF}_{X_{\mathrm{HY}}}^2)\\ &+(\beta_1^2+\frac{\beta_1^2\sigma_{x_S}^2\frac{1-\mathrm{Corr}^2\ (y,x)}{\mathrm{Corr}^2\ (y,x)}}{n\sigma_{x_S}^2})(1-\frac{M}{N})\frac{\sigma_{x_U}^2}{M}\\ &+\frac{\beta_1^2\sigma_{x_S}^2\frac{1-\mathrm{Corr}^2\ (y,x)}{\mathrm{Corr}^2\ (y,x)}}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_F^{\mathrm{dist}_{i,j}},\end{aligned} \tag{A.17}$$

knowing that $\mathrm{V}_F\ \left(\widehat{\beta}_{1_S}\right) = \frac{\omega^2}{n\sigma_{x_S}^2} = \frac{\beta_1^2\sigma_{x_S}^2\frac{1-\mathrm{Corr}^2\ (y,x)}{\mathrm{Corr}^2\ (y,x)}}{n\sigma_{x_S}^2}$ and $\mathrm{DF}_{X_{\mathrm{HY}}} = \frac{(\overline{x}_U-\overline{x}_S)}{\sigma_{x_S}}$, where DF is a disparity factor, which will be explained in Appendix A.4.

### A.2. Conventional model-based inference

Like for hybrid inference, we denote an area of interest (AOI) as $U$ with $N$ population units (grid-cells). A dataset with field AGB and TanDEM-X variables is denoted as $S$ and has $n$ units.

A regression model 'G' is introduced:

$$\mathrm{G}:\ \mathrm{y} = \alpha_0+\alpha_1\mathrm{z}+\boldsymbol{\upsilon},\ \boldsymbol{\upsilon}\sim N(\mathbf{0},\delta^2\mathrm{I}), \tag{A.18}$$

where $\mathbf{y}$ is a vector of field AGB values, $\alpha_0$ and $\alpha_1$ are model parameters to be estimated, $\mathbf{z}$ is a vector of the TanDEM-X variable, and $\boldsymbol{\upsilon}$ is a vector of residual errors with a constant variance $\delta^2$ (cf., Mehtatalo and Lappi, 2020).

The AOI population mean $\mu$ is predicted by

$$\begin{aligned}\widehat{\mu}_{\mathrm{MB}} &= \widehat{\alpha}_{0_S}+\widehat{\alpha}_{1_S}\overline{z}_U\\ &= \overline{\mathrm{z}}_U\widehat{\boldsymbol{\alpha}}_S,\end{aligned} \tag{A.19}$$

where the subscript *MB* denotes "model-based", $\widehat{\alpha}_{0_S}$ and $\widehat{\alpha}_{1_S}$ are estimated model parameters based on the dataset $S$, and $\overline{\mathrm{z}}_U = \begin{pmatrix}1 & \overline{z}_U\end{pmatrix}$ is a mean of the TanDEM-X variable (including a vector of units) over the AOI $U$.

The true AOI population mean $\mu$ can be expressed as

$$\mu = \alpha_0+\alpha_1\overline{z}_U+\overline{\upsilon}_U, \tag{A.20}$$

where $\overline{\upsilon}_U$ is the mean of residual errors over $U$.

The MSE of $\widehat{\mu}_{\mathrm{MB}}$ is expressed as (e.g., p. 127 Cassel et al., 1977)

$$\mathrm{MSE}(\widehat{\mu}_{\mathrm{MB}}) = \overline{\mathrm{z}}_U\mathrm{Cov}(\widehat{\boldsymbol{\alpha}}_S)\overline{\mathrm{z}}_U^{\top}+\mathrm{V}(\overline{\upsilon}_U). \tag{A.21}$$

As in the case of the F model, in this case we also assume that $(\widehat{\alpha}_{0_S}+\widehat{\alpha}_{1_S}\overline{z}_U)$ and $\overline{\upsilon}_U$ are independent.

The first term of the expression is the variance due to the estimated model parameters $\widehat{\boldsymbol{\alpha}}_S$, and the second term is the variance due to the residual errors.

Similarly to case of hybrid inference, the first term can be elaborated further

$$\begin{aligned}\bar{z}_U \mathrm{Cov}(\widehat{\boldsymbol{\alpha}}_S)\bar{z}_U^\top &= \begin{pmatrix} 1 & \bar{z}_U \end{pmatrix} \mathrm{Cov}(\widehat{\boldsymbol{\alpha}}_S) \begin{pmatrix} 1 & \bar{z}_U \end{pmatrix}^\top \\ &= \delta^2 \begin{pmatrix} 1 & \bar{z}_U \end{pmatrix} (\mathrm{Z}_S^\top \mathrm{Z}_S)^{-1} \begin{pmatrix} 1 \\ \bar{z}_U \end{pmatrix} \\ &= \frac{\delta^2}{n} \frac{\sum_{i=1}^n (z_i - \bar{z}_U)^2}{\sum_{i=1}^n (z_i - \bar{z}_S)^2} \\ &= \frac{\delta^2}{n} (1 + \frac{(\bar{z}_S - \bar{z}_U)^2}{\sigma_{z_S}^2}),\end{aligned} \tag{A.22}$$

where $\sigma_{z_S}^2 = \frac{1}{n}\sum_{i=1}^n (z_i - \bar{z}_S)^2$.

To perform analyses in module I, we expressed the variance of the mean residual errors $\mathrm{V}(\bar{v}_U)$ as

$$\mathrm{V}(\bar{v}_U) = \frac{\delta^2}{N^2} \sum_{i=1}^N \sum_{j=1}^N \rho_G^{\mathrm{dist}_{i,j}}, \tag{A.23}$$

where $\rho_G$ is a spatial autocorrelation between two neighboring units and given in the Table 2 for Field AGB – TanDEM-X model.

To accommodate analyses in module III, we expressed the residual error variance $\delta^2$ through the correlation between field AGB and TanDEM-X variables (cf. Eq. 7 Helland, 1987):

$$\delta^2 = \alpha_1^2 \sigma_{z_S}^2 \frac{1 - \mathrm{Corr}^2\ (\mathrm{y}, \mathrm{z})}{\mathrm{Corr}^2\ (\mathrm{y}, \mathrm{z})}. \tag{A.24}$$

And thus, the entire expression of $\mathrm{MSE}(\widehat{\mu}_{\mathrm{MB}})$ is

$$\mathrm{MSE}(\widehat{\mu}_{\mathrm{MB}}) = \frac{\alpha_1^2 \sigma_{z_S}^2 \frac{1-\mathrm{Corr}^2\ (\mathrm{y,z})}{\mathrm{Corr}^2\ (\mathrm{y,z})}}{n} (1 + \mathrm{DF}_{\mathrm{Z_{MB}}}^2) + \frac{\alpha_1^2 \sigma_{z_S}^2 \frac{1-\mathrm{Corr}^2\ (\mathrm{y,z})}{\mathrm{Corr}^2\ (\mathrm{y,z})}}{N^2} \sum_{i=1}^N \sum_{j=1}^N \rho_G^{\mathrm{dist}_{i,j}}, \tag{A.25}$$

where a disparity factor (DF) in TanDEM-X data for conventional MB model-based inference is expressed as $\mathrm{DF}_{\mathrm{Z_{MB}}} = \frac{(\bar{z}_U - \bar{z}_S)}{\sigma_{z_S}}$ (see Appendix A.4 for more details).

### A.3. Hierarchical model-based inference

For the hierarchical model-based (HMB) inference we use similar notation as before, i.e. an area of interest (AOI) is denoted $U$, with $N$ population units (grid-cells). A sample of $M$ GEDI footprints available in the AOI is denoted as $Sa$. A dataset with field AGB and GEDI variables is denoted as $S$ and has $n$ units.

In HMB inference, two models are needed. The model 'Q' is defined as:

$$\mathrm{Q}: \mathrm{X} = \mathrm{Z}\boldsymbol{\Gamma} + \mathrm{D},\ \mathrm{D} \sim N(\mathbf{0}, \Theta \otimes \mathrm{I}), \tag{A.26}$$

where **X** has a column of units as the first column, and **x** as the second; **Z** has a column of units as the first column, and **z** is the second; $\boldsymbol{\Gamma}$ is a matrix with the first column containing 1 and 0 (to address the column of units in **X**) and the second column containing model parameters $\gamma_0$ and $\gamma_1$ (to predict **x**), and **D** is a matrix, with a first column of zeros, and a second column with **d**; $\otimes$ denotes the Kronecker product of the matrices $\Theta$ and **I**, where $\Theta$ is a matrix of size 2×2, with 0 and $\theta^2$ on the diagonal and zeros elsewhere in our case.

The second model, denoted 'G*', is:

$$\mathrm{G}^*: \mathrm{X}\boldsymbol{\beta} = \mathrm{Z}\boldsymbol{\alpha}^* + \mathrm{u},\ \mathrm{u} \sim N(\mathbf{0}, \lambda^2\mathrm{I}). \tag{A.27}$$

The model G* is obtained by multiplying both sides of the model Q with $\boldsymbol{\beta}$ and denoting $\boldsymbol{\Gamma}\boldsymbol{\beta} = \boldsymbol{\alpha}^*$, $\mathbf{D}\boldsymbol{\beta} = \mathbf{u}$ and $\boldsymbol{\beta}^\top\Theta\boldsymbol{\beta} = \lambda^2$, i.e.:

$$\begin{aligned}\mathrm{G}^*: \mathrm{X}\boldsymbol{\beta} &= \mathrm{Z}\boldsymbol{\Gamma}\boldsymbol{\beta} + \mathrm{D}\boldsymbol{\beta},\ \mathrm{D}\boldsymbol{\beta} \sim N(\mathbf{0}, \boldsymbol{\beta}^\top\Theta\boldsymbol{\beta} \otimes \mathrm{I}) \\ &= \mathrm{Z}\boldsymbol{\alpha}^* + \mathrm{u},\ \mathrm{u} \sim N(\mathbf{0}, \lambda^2\mathrm{I}),\end{aligned}$$

where $\mathrm{X}\boldsymbol{\beta} = \mathrm{E}_F\ [\mathrm{X}\widehat{\boldsymbol{\beta}}_S]$, i.e. it is an expectation of the predicted AGB over the GEDI footprints using model F.

Using the expression Eq. (A.27) and for given estimated $\boldsymbol{\beta}$, the model parameters $\boldsymbol{\alpha}^*$ are estimated as

$$\widehat{\boldsymbol{\alpha}}_{\mathrm{Sa}}^* = (\mathrm{Z}_{\mathrm{Sa}}^\top \mathrm{Z}_{\mathrm{Sa}})^{-1} \mathrm{Z}_{\mathrm{Sa}}^\top \mathrm{X}_{\mathrm{Sa}} \widehat{\boldsymbol{\beta}}_S, \tag{A.28}$$

which is the HMB predictor of model parameters $\boldsymbol{\alpha}^*$. In the expression Eq. (A.28) we can see that $(\mathrm{Z}_{\mathrm{Sa}}^\top \mathrm{Z}_{\mathrm{Sa}})^{-1} \mathrm{Z}_{\mathrm{Sa}}^\top \mathrm{X}_{\mathrm{Sa}}$ in fact is an OLS estimator of model parameters $\boldsymbol{\Gamma}$ in the multivariate model Q. The estimator $\widehat{\boldsymbol{\alpha}}_{\mathrm{Sa}}^*$ can be rewritten further as

$$\widehat{\boldsymbol{\alpha}}_{\mathrm{Sa}}^* = \widehat{\boldsymbol{\Gamma}}_{\mathrm{Sa}} \widehat{\boldsymbol{\beta}}_S, \tag{A.29}$$

where $\widehat{\boldsymbol{\Gamma}}_{\text{Sa}} = (\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1}\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{X}_{\text{Sa}}$.

Thus, the AOI population mean $\mu$ is predicted by

$$\begin{aligned}\widehat{\mu}_{\text{HMB}} &= \widehat{\alpha}_{0_{\text{Sa}}}^{*} + \widehat{\alpha}_{1_{\text{Sa}}}^{*}\overline{z}_{U} \\ &= \overline{\mathrm{z}}_{U}\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}.\end{aligned} \tag{A.30}$$

The expectation of true AOI population mean due to model F, $\overline{\mathrm{x}}_{U}\boldsymbol{\beta}$, can be expressed through the model G*, i.e.

$$\overline{\mathrm{x}}_{U}\boldsymbol{\beta} = \alpha_{0}^{*} + \alpha_{1}^{*}\overline{z}_{U} + \overline{u}_{U}, \tag{A.31}$$

where $\overline{u}_{U}$ is the mean of residual errors over $U$.

The MSE of $\widehat{\mu}_{\text{HMB}}$ is expressed as

$$\text{MSE}(\widehat{\mu}_{\text{HMB}}) = \mathrm{E}\left[(\widehat{\mu}_{\text{HMB}} - \mu)^{2}\right] \tag{A.32}$$

Expressing $\mu$ using model F through $\mu = \overline{\mathrm{x}}_{U}\boldsymbol{\beta} + \overline{\varepsilon}_{;U}$, and $\widehat{\mu}_{\text{HMB}}$ using model G* through $\widehat{\mu}_{\text{HMB}} = \widehat{\alpha}_{0_{\text{Sa}}}^{*} + \widehat{\alpha}_{1_{\text{Sa}}}^{*}\overline{z}_{U}$, we have

$$\begin{aligned}\text{MSE}(\widehat{\mu}_{\text{HMB}}) &= \mathrm{E}\left[(\widehat{\alpha}_{0_{\text{Sa}}}^{*} + \widehat{\alpha}_{1_{\text{Sa}}}^{*}\overline{z}_{U} - (\overline{\mathrm{x}}_{U}\boldsymbol{\beta} + \overline{\varepsilon}_{;U}))^{2}\right] \\ &= \mathrm{E}\left[(\widehat{\alpha}_{0_{\text{Sa}}}^{*} + \widehat{\alpha}_{1_{\text{Sa}}}^{*}\overline{z}_{U} - (\alpha_{0}^{*} + \alpha_{1}^{*}\overline{z}_{U} + \overline{u}_{U} + \overline{\varepsilon}_{;U}))^{2}\right] \\ &= \mathrm{V}\left(\widehat{\alpha}_{0_{\text{Sa}}}^{*} + \widehat{\alpha}_{1_{\text{Sa}}}^{*}\overline{z}_{U}\right) + \mathrm{V}(\overline{u}_{U} + \overline{\varepsilon}_{;U}) \\ &= \overline{\mathrm{z}}_{U}\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right)\overline{\mathrm{z}}_{U}^{\top} + \mathrm{V}(\overline{u}_{U} + \overline{\varepsilon}_{;U}).\end{aligned} \tag{A.33}$$

We assume that $(\widehat{\alpha}_{0_{\text{Sa}}} + \widehat{\alpha}_{1_{\text{Sa}}}\overline{z}_{U})$ and $(\overline{u}_{U} + \overline{\varepsilon}_{;U})$ are independent.

We now derive the $\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right)$ in terms of $\boldsymbol{\Gamma}$ and $\boldsymbol{\beta}$ using the law of total covariance:

$$\begin{aligned}\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right) &= \mathrm{Cov}(\widehat{\boldsymbol{\Gamma}}_{\text{Sa}}\widehat{\boldsymbol{\beta}}_{S}) \\ &= \mathrm{E}_{F}\ [\mathrm{Cov}_{Q}\ (\widehat{\boldsymbol{\Gamma}}_{\text{Sa}}\widehat{\boldsymbol{\beta}}_{S})] + \mathrm{Cov}_{F}\ (\mathrm{E}_{Q}\ [\widehat{\boldsymbol{\Gamma}}_{\text{Sa}}\widehat{\boldsymbol{\beta}}_{S}]) \\ &= \mathrm{E}_{F}\ \left[\widehat{\boldsymbol{\beta}}_{S}^{\top}\Theta\widehat{\boldsymbol{\beta}}_{S} \otimes (\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1}\right] + \mathrm{Cov}_{F}\ (\boldsymbol{\Gamma}\widehat{\boldsymbol{\beta}}_{S}) \\ &= \mathrm{E}_{F}\ \left[\widehat{\boldsymbol{\beta}}_{S}^{\top}\Theta\widehat{\boldsymbol{\beta}}_{S}\right] \otimes (\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1} + \boldsymbol{\Gamma}\mathrm{Cov}_{F}\ (\widehat{\boldsymbol{\beta}}_{S})\boldsymbol{\Gamma}^{\top} \\ &= \mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right](\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1} + \omega^{2}\boldsymbol{\Gamma}(\mathrm{X}_{S}^{\top}\mathrm{X}_{S})^{-1}\boldsymbol{\Gamma}^{\top},\end{aligned} \tag{A.34}$$

where $\mathrm{E}_{F}$ and $\mathrm{Cov}_{F}$ denote expectation and covariance due to the model F, and $\mathrm{E}_{Q}$ and $\mathrm{Cov}_{Q}$ expectation and covariance due to the model Q.

Therefore, $\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right)$ can be expressed as

$$\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right) = \mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right](\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1} + \omega^{2}\boldsymbol{\Gamma}(\mathrm{X}_{S}^{\top}\mathrm{X}_{S})^{-1}\boldsymbol{\Gamma}^{\top}. \tag{A.35}$$

The first term in MSE of $\widehat{\mu}_{\text{HMB}}$, expression Eq. (A.33), can be elaborated using Eq. (A.35) into a generic form as following

$$\begin{aligned}\overline{\mathrm{z}}_{U}\mathrm{Cov}\left(\widehat{\boldsymbol{\alpha}}_{\text{Sa}}^{*}\right)\overline{\mathrm{z}}_{U}^{\top} &= \mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right]\overline{\mathrm{z}}_{U}(\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1}\overline{\mathrm{z}}_{U}^{\top} + \omega^{2}\overline{\mathrm{z}}_{U}\boldsymbol{\Gamma}(\mathrm{X}_{S}^{\top}\mathrm{X}_{S})^{-1}\boldsymbol{\Gamma}^{\top}\overline{\mathrm{z}}_{U}^{\top} \\ &= \mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right]\begin{pmatrix}1 & \overline{z}_{U}\end{pmatrix}(\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1}\begin{pmatrix}1 \\ \overline{z}_{U}\end{pmatrix} + \omega^{2}\begin{pmatrix}1 & (\gamma_{0} + \overline{z}_{U}\gamma_{1})\end{pmatrix}(\mathrm{X}_{S}^{\top}\mathrm{X}_{S})^{-1}\begin{pmatrix}1 \\ (\gamma_{0} + \overline{z}_{U}\gamma_{1})\end{pmatrix} \\ &= \mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right]\begin{pmatrix}1 & \overline{z}_{U}\end{pmatrix}(\mathrm{Z}_{\text{Sa}}^{\top}\mathrm{Z}_{\text{Sa}})^{-1}\begin{pmatrix}1 \\ \overline{z}_{U}\end{pmatrix} + \omega^{2}\begin{pmatrix}1 & \mathrm{E}_{Q}\ \left[\widehat{\overline{x}}_{U}\right]\end{pmatrix}(\mathrm{X}_{S}^{\top}\mathrm{X}_{S})^{-1}\begin{pmatrix}1 \\ \mathrm{E}_{Q}\ \left[\widehat{\overline{x}}_{U}\right]\end{pmatrix} \\ &= \frac{\mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right]}{M}\left(1 + \frac{(\overline{z}_{\text{Sa}} - \overline{z}_{U})^{2}}{\sigma_{z_{\text{Sa}}}^{2}}\right) + \frac{\omega^{2}}{n}\left(1 + \frac{(\overline{x}_{S} - \mathrm{E}_{Q}\ [\widehat{\overline{x}}_{U}])^{2}}{\sigma_{x_{S}}^{2}}\right),\end{aligned} \tag{A.36}$$

knowing that $(\gamma_{0} + \gamma_{1}\overline{z}_{U}) = \mathrm{E}_{Q}\ [\widehat{\overline{x}}_{U}]$.

The expectation $\mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right] = \mathrm{E}_{F}\ \left[\widehat{\boldsymbol{\beta}}_{S}^{\top}\Theta\widehat{\boldsymbol{\beta}}_{S}\right]$, in our case with only one explanatory variable, reduces to

$$\begin{aligned}\mathrm{E}_{F}\ \left[\widehat{\lambda}^{2}\right] &= \mathrm{E}_{F}\ \left[\widehat{\beta}_{1_{S}}^{2}\theta^{2}\right] \\ &= \mathrm{E}_{F}\ \left[\widehat{\beta}_{1_{S}}^{2}\right]\theta^{2} \\ &= (\beta_{1}^{2} + \mathrm{V}_{F}\ (\widehat{\beta}_{1_{S}}))\theta^{2}.\end{aligned} \tag{A.37}$$

To perform analyses in module I, we expressed the variance of the mean of the residual errors $\mathrm{V}(\overline{u}_{U} + \overline{\varepsilon}_{;U})$ as

$$\begin{aligned}
\mathrm{V}(\overline{u}_U+\overline{\varepsilon}_{;U}) &= \mathrm{V}(\overline{u}_U)+\mathrm{V}(\overline{\varepsilon}_{;U}) \\
&= \frac{\lambda^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Corr}(u_i,u_j)+\frac{\omega^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Corr}(\varepsilon_{;i},\varepsilon_{;j}) \\
&= \frac{\lambda^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{G^*}^{\mathrm{dist}_{i,j}}+\frac{\omega^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{F}^{\mathrm{dist}_{i,j}} \\
&= \frac{\beta^2\theta^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{G^*}^{\mathrm{dist}_{i,j}}+\frac{\omega^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{F}^{\mathrm{dist}_{i,j}},
\end{aligned} \tag{A.38}$$

where $\rho_{G^*}$ is a spatial autocorrelation between two neighboring units, given in Table 2 for GEDI AGB – TanDEM-X model.

To accommodate analyses in module III, we expressed the residual error variance $\theta^2$ through the correlation between GEDI and TanDEM-X variables (cf. Eq. 7 Helland, 1987):

$$\theta^2=\gamma_1^2\sigma_{z_{\mathrm{Sa}}}^2\frac{1-\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}{\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}. \tag{A.39}$$

And thus, the MSE expression for the HMB predictor is

$$\begin{aligned}
\mathrm{MSE}(\widehat{\mu}_{\mathrm{HMB}}) &= \frac{\left(\beta_1^2+\dfrac{\beta_1^2\sigma_{x_S}^2\dfrac{1-\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}{\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}}{n\sigma_{x_S}^2}\right)\gamma_1^2\sigma_{z_{\mathrm{Sa}}}^2\dfrac{1-\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}{\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}}{M}(1+\mathrm{DF}_{Z_{\mathrm{HMB}}}^2) \\
&+\frac{\beta_1^2\sigma_{x_S}^2\dfrac{1-\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}{\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}}{n}\left(1+\left(\mathrm{DF}_{X_{\mathrm{HMB}}}+\frac{\gamma_1\sigma_{z_{\mathrm{Sa}}}\mathrm{DF}_{Z_{\mathrm{HMB}}}}{\sigma_{x_S}}\right)^2\right) \\
&+\frac{\beta^2\gamma_1^2\sigma_{z_{\mathrm{Sa}}}^2\dfrac{1-\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}{\mathrm{Corr}^2\ (\mathrm{x},\mathrm{z})}}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{G^*}^{\mathrm{dist}_{i,j}} \\
&+\frac{\beta_1^2\sigma_{x_S}^2\dfrac{1-\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}{\mathrm{Corr}^2\ (\mathrm{y},\mathrm{x})}}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{F}^{\mathrm{dist}_{i,j}},
\end{aligned} \tag{A.40}$$

where disparity factors in GEDI data and TanDEM-X data in HMB inference are expressed as $\mathrm{DF}_{X_{\mathrm{HMB}}}=\frac{\overline{x}_{\mathrm{Sa}}-\overline{x}_S}{\sigma_{x_S}}$ and $\mathrm{DF}_{Z_{\mathrm{HMB}}}=\frac{\overline{x}_U-\overline{x}_{\mathrm{Sa}}}{\sigma_{x_S}}$ respectively (see Appendix A.4 for more details).

### A.4. Disparity factor

In previous sections, a new term was introduced – the *Disparity Factor* (DF), which measures the degree of disparity between the mean value of the explanatory variable in the dataset used to estimate the model parameters and the corresponding mean value in the dataset used for applying the model. It also contains the standard deviation of the explanatory variable in the training dataset in the denominator, pointing at the benefit of large variability between observations when estimating model parameters. In Table A.1, the disparity factor expressions are shown for the three methods compared in this study, i.e. hybrid, conventional model-based, and HMB inference. As can be seen in the previous sections, the disparity factor enters as a component of the MSE and variance expressions.

**Table A.1**
The disparity factor for the three different prediction methods.

| Inference | Disparity factor in the space of the **X** explanatory variables ($\mathrm{DF}_X$), i.e. GEDI data in our cases | Disparity factor in the space of the **Z** explanatory variables ($\mathrm{DF}_Z$), i.e. TanDEM-X data in our cases |
|---|---|---|
| Hybrid | $\frac{\overline{x}_U-\overline{x}_S}{\sigma_{x_S}}$ | – |
| Model-Based | – | $\frac{\overline{z}_U-\overline{z}_S}{\sigma_{z_S}}$ |
| Hierarchical Model-Based | $\frac{\overline{x}_{\mathrm{Sa}}-\overline{x}_S}{\sigma_{x_S}}$ | $\frac{\overline{z}_U-\overline{z}_{\mathrm{Sa}}}{\sigma_{z_{\mathrm{Sa}}}}$ |

Values of DFs in empirical data used for the analyses are presented in Table A.2.

**Table A.2**
Disparity Factor values based on empirical data.

| Baseline settings | | California, USA (TEF) | Costa Rica (LSBS) |
|---|---|---|---|
| Disparity Factor: | in GEDI data for hybrid inference | -0.0000 | -0.0013 |
| | in TanDEM-X data for model-based inference | 0.0105 | -0.3854 |
| | in GEDI data for HMB inference | -0.0000 | -0.0013 |
| | in TanDEM-X data for HMB inference | -0.0020 | -0.0416 |

### *A.5. Additional basic empirical data used in the computations*

The main empirical conditions used in the computations for the two study areas are presented in Table 2 in the main text. In Table A.3, we provide some additional data linked to the variance and MSE expressions.

**Table A.3**
Additional basic empirical data used in the computations for the two study areas (the notations of datasets and components are given in Appendix A.1, A.2 and A.3).

| Dataset | Comp. | TEF Field Data | GEDI | TanDEM-X | LSBS Field Data | GEDI | TanDEM-X |
|---|---|---|---|---|---|---|---|
| $S$ | $\bar{y}_S$ | 426.47 | – | – | 231.63 | – | – |
| | $\sigma^2_{y_S}$ | 62201.45 | – | – | 9335.12 | – | – |
| | $\bar{x}_S$ | – | 826.27 | – | – | 408.49 | – |
| | $\sigma^2_{x_S}$ | – | 338765.20 | – | – | 47809.48 | – |
| | $\bar{z}_S$ | – | – | 36.91 | – | – | 37.08 |
| | $\sigma^2_{z_S}$ | – | – | 38.16 | – | – | 10.12 |
| $Sa$ | $\bar{x}_{Sa}$ | – | 822.43 | – | – | 344.19 | – |
| | $\sigma^2_{x_{Sa}}$ | – | 288367.00 | – | – | 37509.82 | – |
| | $\bar{z}_{Sa}$ | – | – | 37.46 | – | – | 34.25 |
| | $\sigma^2_{z_{Sa}}$ | – | – | 73.86 | – | – | 25.73 |
| $U$ | $\bar{z}_U$ | – | – | 37.31 | – | – | 33.18 |

## References


Andersen, H.-E., Barrett, T., Winterberger, K., Strunk, J., Temesgen, H., 2009. Estimating forest biomass on the western lowlands of the Kenai Peninsula of Alaska using airborne lidar and field plot data in a model-assisted sampling design. In: In Proceedings of the IUFRO Division 4 Conference: Extending Forest Inventory and Monitoring over Space and Time., pp. 19–22.
Asner, G.P., 2001. Cloud cover in landsat observations of the Brazilian Amazon. Int. J. Remote Sens. 22, 3855–3862.
Boudreau, J., Nelson, R.F., Margolis, H.A., Beaudoin, A., Guindon, L., Kimes, D.S., 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. Remote Sens. Environ. 112, 3876–3890.
Cassel, C.-M., Särndal, C.E., Wretman, J.H., 1977. Foundations of Inference in Survey Sampling. Wiley.
Corona, P., Fattorini, L., Franceschi, S., Scrinzi, G., Torresan, C., 2014. Estimation of standing wood volume in forest compartments by exploiting airborne laser scanning information: model-based, design-based, and hybrid perspectives. Can. J. Forest Res. 44, 1303–1311.
Dubayah, R., Armston, J., Healey, S., Bruening, J., Patterson, P., Kellner, J., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., 2022. GEDI launches a new era of biomass inference from space. EarthArXiv (p. Preprint).
Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. Science of Remote Sensing 1, 100002.
Dubayah, R., Goetz, S., Blair, J.B., Fatoyinbo, T., Hansen, M., Healey, S.P., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., 2014. The global ecosystem dynamics investigation. In: In AGU Fall Meeting Abstracts.
Duncanson, L., Kellner, J.R., Armston, J., Dubayah, R., Minor, D.M., Hancock, S., Healey, S.P., Patterson, P.L., Saarela, S., Marselis, S., 2022. Aboveground biomass density models for NASA's global ecosystem dynamics investigation (GEDI) lidar mission. Remote Sens. Environ. 270, 112845.
Gobakken, T., Næsset, E., Nelson, R., Bollandsås, O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O., Astrup, R., 2012. Estimating biomass in hedmark county Norway using national forest inventory field plots and airborne laser scanning. Remote Sens. Environ. 123, 443–456.
Gregoire, T.G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Can. J. Forest Res. 28, 1429–1447.
Gregoire, T.G., Næsset, E., McRoberts, R.E., Ståhl, G., Andersen, H.-E., Gobakken, T., Ene, L., Nelson, R., 2016. Statistical rigor in LiDAR-assisted estimation of aboveground forest biomass. Remote Sens. Environ. 173, 98–108.
Gregoire, T.G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R., Holm, S., 2011. Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County Norway. Can. J, Forest Res. 41, 83–95.
Gregoire, T.G., Valentine, H.T., 2008. Sampling Strategies for Natural Resources and the Environment Volume 1. CRC Press.
Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L.I., Kellner, J.R., Dubayah, R., 2019. The GEDI simulator: a large-footprint waveform lidar simulator for calibration and validation of spaceborne missions. Earth Space Sci. 6, 294–310.
Healey, S., Armston, J.D., Yang, Z., Dubayah, R., Bruening, J., Patterson, P.L., Saarela, S., Ståhl, G., Duncanson, L., Kellner, J.R., 2021. The GEDI gridded biomass product: patterns of coverage and precision after two years of operation. In: In AGU Fall Meeting AGU.
Healey, S.P., Patterson, P.L., Saatchi, S., Lefsky, M.A., Lister, A.J., Freeman, E.A., 2012. A sample design for globally consistent biomass estimation using lidar data from the geoscience laser altimeter system (GLAS). Carbon Balance and Manag. 7, 10.
Healey, S.P., Yang, Z., Gorelick, N., Ilyushchenko, S., 2020. Highly local model calibration with a new GEDI LiDAR asset on google earth engine reduces landsat forest height signal saturation. Remote Sens. 12, 2840.
Helland, I.S., 1987. On the interpretation and use of R2 in regression analysis. Biometrics 61–69.
Holm, S., Nelson, R., Ståhl, G., 2017. Hybrid three-phase estimators for large-area forest inventory using ground plots, airborne lidar, and space lidar. Remote Sens. Environ. 197, 85–97.
Hyyppä, J., Hallikainen, M., 1996. Applicability of airborne profiling radar to forest inventory. Remote Sens. Environ. 57, 39–57.
Hyyppä, J., Inkinen, M., 1999. Detecting and estimating attributes for single trees using laser scanner. Photogramm. J. Finland 16, 27–42.
Magnussen, S., Mandallaz, D., Lanz, A., Ginzler, C., Næsset, E., Gobakken, T., 2016. Scale effects in survey estimates of proportions and quantiles of per unit area attributes. For. Ecol. Manag. 364, 122–129.
Magnussen, S., McRoberts, R.E., Breidenbach, J., Nord-Larsen, T., Ståhl, G., Fehrmann, L., Schnell, S., 2020. Comparison of estimators of variance for forest inventories with systematic sampling-results from artificial populations. For. Ecosyst. 7, 1–19.
McRoberts, R.E., 2010a. Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. Remote Sens. Environ. 114, 1017–1025.
McRoberts, R.E., 2010b. The effects of rectification and global positioning system errors on satellite image-based estimates of forest area. Remote Sens. Environ. 114, 1710–1717.
McRoberts, R.E., Næsset, E., Gobakken, T., Chirici, G., Condés, S., Hou, Z., Saarela, S., Chen, Q., Ståhl, G., Walters, B.F., 2018. Assessing components of the model-based

mean square error estimator for remote sensing assisted forest applications. Can. J, For. Res. 48.
Mehtatalo, L., Lappi, J., 2020. Biometry for Forestry and Environmental Data: With Examples in R. CRC press.
Milenković, M., Schnell, S., Holmgren, J., Ressl, C., Lindberg, E., Hollaus, M., Pfeifer, N., Olsson, H., 2017. Influence of footprint size and geolocation error on the precision of forest biomass estimates from space-borne waveform LiDAR. Remote Sens. Environ. 200, 74–88.
Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80, 88–99.
Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scand. J. For. Res. 19, 164–179.
Neigh, C.S., Nelson, R.F., Ranson, K.J., Margolis, H.A., Montesano, P.M., Sun, G., Kharuk, V., Næsset, E., Wulder, M.A., Andersen, H.-E., 2013. Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. Remote Sens. Environ. 137, 274–287.
Nelson, R., Boudreau, J., Gregoire, T.G., Margolis, H., Næsset, E., Gobakken, T., Ståhl, G., 2009. Estimating Quebec provincial forest resources using ICESat/GLAS. Can. J. For. Res. 39, 862–881.
Nelson, R., Gobakken, T., Næsset, E., Gregoire, T., Ståhl, G., Holm, S., Flewelling, J., 2012. Lidar sampling - using an airborne profiler to estimate forest biomass in Hedmark County. Norway. Remote Sens. Environ. 123, 563–578.
Nelson, R., Krabill, W., Tonelli, J., 1988. Estimating forest biomass and volume using airborne laser data. Remote Sens. Environ. 24, 247–267.
Nelson, R., Short, A., Valenti, M., 2004. Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. Scand. J. For. Res. 19, 500–511.
Patterson, P.L., Healey, S.P., Ståhl, G., Saarela, S., Holm, S., Andersen, H.-E., Dubayah, R. O., Duncanson, L., Hancock, S., Armston, J., 2019. Statistical properties of hybrid estimators proposed for GEDI - NASA's global ecosystem dynamics investigation. Environ. Res. Lett. 14, 065007.
Persson, H.J., Ståhl, G., 2020. Characterizing uncertainty in forest remote sensing studies. Remote Sens. 12, 505.
Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sens. Environ. 253, 112165.
Qi, W., Saarela, S., Armston, J., Ståhl, G., Dubayah, R., 2019. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. Remote Sens. Environ. 232, 111283.
Rishmawi, K., Huang, C., Zhan, X., 2021. Monitoring key forest structure attributes across the conterminous United States by integrating GEDI LiDAR measurements and VIIRS data. Remote Sens. 13, 442.
Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyyppä, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and landsat data as auxiliary information. Remote Sens. Environ. 158, 431–440.
Saarela, S., Holm, S., Grafström, A., Schnell, S., Næsset, E., Gregoire, T.G., Nelson, R.F., Ståhl, G., 2016. Hierarchical model-based inference for forest inventory utilizing three sources of information. Ann. For. Sci. 73, 895–910.
Saarela, S., Holm, S., Healey, S.P., Andersen, H.-E., Petersson, H., Prentius, W., Patterson, P.L., Næsset, E., Gregoire, T.G., Ståhl, G., 2018. Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and landsat data. Remote Sens. 10, 1832.
Saarela, S., Wästlund, A., Holmström, E., Mensah, A.A., Holm, S., Nilsson, M., Fridman, J., Ståhl, G., 2020. Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data, accounting for tree level allometric and LiDAR model errors. For. Ecosyst. 7, 1–17.
Saarela, S., Yang, Z., Ståhl, G., 2022. HMB: Hierarchical Model-Based estimation approach. R package version 2.0.
Solodukhin, V.I., Kulyasov, A.G., Utenkov, B.I., Zhukov, A.Y., Mazhugin, I.N., Emel'yanov, V.P., Korolev, I.A., 1977. Metody izuchenija vertikal'nyh sechenij drevostoev (method of study of vertical sections of forest stands). Lesnoe Khozyaistvo (in Russian) 2, 71–73.
Ståhl, G., Heikkinen, J., Petersson, H., Repola, J., Holm, S., 2014. Sample-based estimation of greenhouse gas emissions from forests - a new approach to account for both sampling and model errors. For. Sci. 60, 3–13.
Ståhl, G., Holm, S., Gregoire, T.G., Gobakken, T., Næsset, E., Nelson, R., 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County Norway. Can. J. For. Res. 41, 96–107.
Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S.P., Patterson, P.L., Magnussen, S., Næsset, E., McRoberts, R.E., Gregoire, T.G., 2016. Use of models for improved estimation in sample-based large-area forest surveys: a review. For. Ecosyst. 3 (5), 1–11.
Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. J. Geodyn. 34, 405–445.