Remote Sensing of Environment 299 (2023) 113854

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Conjugating remotely sensed data assimilation and model-assisted estimation for efficient multivariate forest inventory

Zhengyang Hou [a,b,*], Keyan Yuan [a,b], Göran Ståhl [c], Ronald E. McRoberts [d], Annika Kangas [e], Hao Tang [f], Jingyi Jiang [a,b], Jinghui Meng [a], Qing Xu [g,**], Zengyuan Li [h,i]

[a] The Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China
[b] Ecological Observation and Research Station of Heilongjiang Sanjiang Plain Wetlands, National Forestry and Grassland Administration, Shuangyashan 518000, China
[c] Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden
[d] Department of Forest Resources, University of Minnesota, St. Paul, MN, USA
[e] Natural Resources Institute Finland (Luke), Bioeconomy and Environment Unit, Joensuu, Finland
[f] Department of Geography, National University of Singapore, Singapore
[g] Key Laboratory of National Forestry and Grassland Administration/Beijing for Bamboo & Rattan Science and Technology, International Center for Bamboo and Rattan, Beijing 100102, China
[h] Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland Administration, Beijing 100091, China
[i] Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China

ARTICLE INFO

Edited by Dr. Marie Weiss

Keywords:
Model-assisted estimation
Data assimilation
Survey sampling
Seemingly unrelated regression
Best linear unbiased predictor

ABSTRACT

Remote sensing aims to provide precise information on forest ecosystems under climate and land use changes, much of which is in the form of parameters estimated for biotic and abiotic variables for various official reporting instruments. Model-assisted estimation (MA) that harnesses remote sensing has demonstrated a surpassing ability to balance the tradeoff between robustness and efficiency. However, (1) MA has to unfold in a way complying with rather than overriding a sampling design because modification to sample size and field protocol is usually not allowed for an established setup, thus impeding further increases to inventory precision; and (2) it is inefficient to predict multiple forest attributes with many individual models, producing inconsistencies in the estimates due to lack of preserving the correlations, and offsetting the gains in inventory precision with the cost spent on modeling. Consequently, within the statistical framework of MA, this study proposes a remotely sensed data assimilation procedure, DAMA, to support high-precision multivariate forest inventory. Based on populations in China and Burkina Faso, promising results indicate that (1) the DAMA estimator proposed is approximately design-unbiased with its variance affected by the sampling design, the prediction accuracy, and the type of remotely sensed auxiliaries involved in DA, in descending order; (2) with simple random sampling, DAMA estimator increases the inferential precision on average 14% and 7% for Horvitz-Thompson and MA counterparts; and (3) with two-stage sampling, remarkably, 180% and 57%. Overall, DAMA demonstrates considerable efficiency that would better serve natural resource observation and management.

1. Introduction

Remote sensing aims to provide accurate information about forest ecosystems under climate and land use changes, much of which is in the form of parameters estimated for biotic and abiotic variables of biodiversity, energy and material cycling for national, regional, and international reporting instruments including National Forest Inventory (NFI) and United Nations Framework Convention on Climate Change (UNFCCC) (Tomppo et al., 2010; Vidal et al., 2016). These official instruments require continuous inventory estimates, typically on an annual basis and at population or domain levels where a domain refers to the subset of a population, e.g., a county of a province or a productivity class of a eucalyptus farm (Hou et al., 2022). This indicates that robustness and efficiency are the key to any inferential procedure confronting these demands (Eggleston et al., 2006; Williams and Brown, 2019).

* Corresponding author at: The Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China.
** Correspondence author.
E-mail addresses: houzhengyang@bjfu.edu.cn (Z. Hou), qing.xu@icbr.ac.cn (Q. Xu).

https://doi.org/10.1016/j.rse.2023.113854
Received 7 May 2023; Received in revised form 6 August 2023; Accepted 10 October 2023
Available online 19 October 2023

Remote sensing-assisted forest inventory using model-assisted estimation (MA) has demonstrated an extraordinary ability to balance the tradeoff between robustness and efficiency. MA is a category of design-based inference that combines features of sampling designs and model prediction in population parameter estimation (Särndal et al., 1992) where remote sensing is the key to statistical modeling and prediction (Baffetta et al., 2009; Gregoire et al., 2016; McRoberts et al., 2022; Næsset et al., 2011). For MA estimators, the approximate design-unbiasedness holds valid by probability sampling and a bias-correction term where design-consistency is affected by factors including design, sample size, remotely sensed auxiliary data, and modeling (Hou et al., 2018; Kangas et al., 2016; Mandallaz et al., 2013; Myllymäki et al., 2017; Ståhl et al., 2016; Stehman, 2009). A standard design-based estimator, e.g., Horvitz-Thompson (HT) estimator, is design-unbiased, and the precision of which increases as sample size increases. The MA counterpart of the standard estimator has the advantage of being able to produce a higher precision using the same sample, or equivalently, to achieve a similar precision with much smaller sample sizes (Gregoire et al., 2011; McRoberts et al., 2013).

However, challenges persist for MA:

- MA has to unfold in a way complying with rather than overriding a sampling design.
- It is inefficient to predict multiple forest attributes with many individual models.
- Remotely sensed auxiliary data are not necessarily free.

Modification to sample size and field protocol is usually not allowed for an established sampling setup, thus impeding further increases to inventory precision particularly for NFIs (Hou et al., 2021). Predicting multiple forest attributes with many individual models would produce inconsistencies in the estimates due to lack of preserving the correlations, which offsets the gains in inventory precision with the cost spent on modeling (McRoberts et al., 2017). Remotely sensed quality data are costly. Although LiDAR is increasingly considered ideal for predicting forest variables of interests (VOIs) (Xu et al., 2018, 2019), frequent acquisitions for large areas are still impractical. Ideal auxiliary data would be timely and effective, and preferably free of charge. Hence, alternatives that overcome these challenges must be sought.

Data assimilation (DA) is a viable option for improving MA. DA is a broad category of mathematical procedures that improves prediction or estimation by adjusting the parameters of the marginal distribution of a random variable to the parameters of its conditional distribution given the jointly distributed variables observed (Fletcher, 2017; Lahoz et al., 2010). In forest inventory, DA has drawn increasing attention for applications entailing updating parameters, calibration of predictions, and prediction of missing values, with various procedures rooted in estimation theory including Kalman filter, mixed estimator, and Bayesian statistics (Czaplewski, 1990; Hou et al., 2021). Best linear unbiased predictor (BLUP) is a frequentist counterpart formula of Bayesian statistics that estimates the conditional expectation for multivariate normal vectors (Henderson, 1975; Robinson, 1991) and is the equation from which the Kalman filter, the mixed estimator, and Kriging were mathematically derived (Kalman, 1960; Robinson, 1991; Theil, 1963). With BLUP, Hou et al. (2019) and Xu et al. (2023) devised efficient DA procedures that leverage cross-model correlation and spatial autocorrelation for model-based inference.

Consequently, the objectives of this study are threefold: (1) to develop and demonstrate a data assimilation procedure, DAMA, that improves the inferential precision of MA through integrating sampling design, remote sensing, simultaneous prediction, and BLUP for inventorying multiple VOIs; (2) to compare HT, MA and DAMA estimators for simple random and two-stage samplings; and (3) to compare the performance of remotely sensed auxiliary data obtained with multispectral satellites that are free of charge or inexpensive to MA and DAMA. A highlight of DAMA resides in its approximate design-unbiasedness because of its probabilistic nature.

2. Materials

2.1. Populations and samples

Two study areas were selected from southwestern China and southeastern Burkina Faso for evaluating the DAMA estimators proposed in the method section (Fig. 1). Generalizability of the evaluation would benefit from these disparate albeit typical populations in respective ecosystems. A county in China, Pingxiang, forms the first population of 64,766 ha (22°00′N, 106°50′W), and the Kou region in Burkina Faso forms the second population of 10,712 ha (11°45′N, 1°57′W). Vegetation in the first population is subtropical monsoon rainforest with moderate to dense coverage, and coverage and vegetation in the second population is dry savanna.

There are 719,624 elements in the China population, and 119,025 elements in the Burkina Faso population. Each element encompasses 30 by 30 m. Different sampling designs produced the sample of respective populations (Fig. 1). For China, according to the Chinese NFI protocol (NFGA, 2010), a simple random sample of 145 elements was selected and field measured in the second half of 2015. For Burkina Faso, a two-stage sample of 160 elements was chosen. Selection and field measurements were carried out between November 2013 and February 2014 according to the protocol of the Land Degradation Surveillance Framework (Vågen et al., 2015). The VOIs for both populations were densities of stem volume (m^3/ha) and basal area (m^2/ha), two standard forest attributes that are often correlated.

2.2. Remotely sensed auxiliary variables

RapidEye satellite sensor (RE) and Landsat 8 Operational Land Imager (L8) data were collected for both populations and georeferenced to WGS84/UTM Zone 48 N and 30 N. Single scenes of respective sensors covered respective populations and were acquired for the field season. RE data with a spatial resolution of 5 m were purchased at 1.3 USD/km^2 and processed to Level-3 A with a resampled spatial resolution of 30 m. L8 data with a spatial resolution of 30 m were Provisional Surface Reflectance product freely available from the U.S. Geological Survey.

Auxiliary variables that are candidates for modeling and DA calibration were calculated from the spectral bands, spectral indices (Table 1), the first principal component of the spectral bands (PCA), the textures of PCA, and textures for the respective spectral indices. Textures included the mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment and correlation (Haralick et al., 1973).

3. Methods

3.1. Overview

Sections within Methods are nested and structured in a top-down fashion. First, Section 3.2 introduces design-based inference, a gold standard for NFIs, with classical point and variance estimators that do not require models or remotely sensed auxiliary data, but only field observations regarding VOI. Point estimator uses the observed values and inclusion probabilities of a sample to infer the expected value of a VOI in a spatial population. Variance estimator quantifies the uncertainty in the corresponding point estimator due to the random inclusion of elements or sample plots into the sample. The ideal inference or inventory is consistent with the desired properties of an estimator, such as unbiasedness and small variance, at little cost.

Second, to this end, Section 3.3 introduces MA where corresponding estimators from Section 3.2 are developed. These estimators are also design-based, but more efficient than the estimators in Section 3.2, as they further exploit remotely sensed auxiliary data and models to reduce variance. Third, Section 3.4 further introduces simultaneous modeling

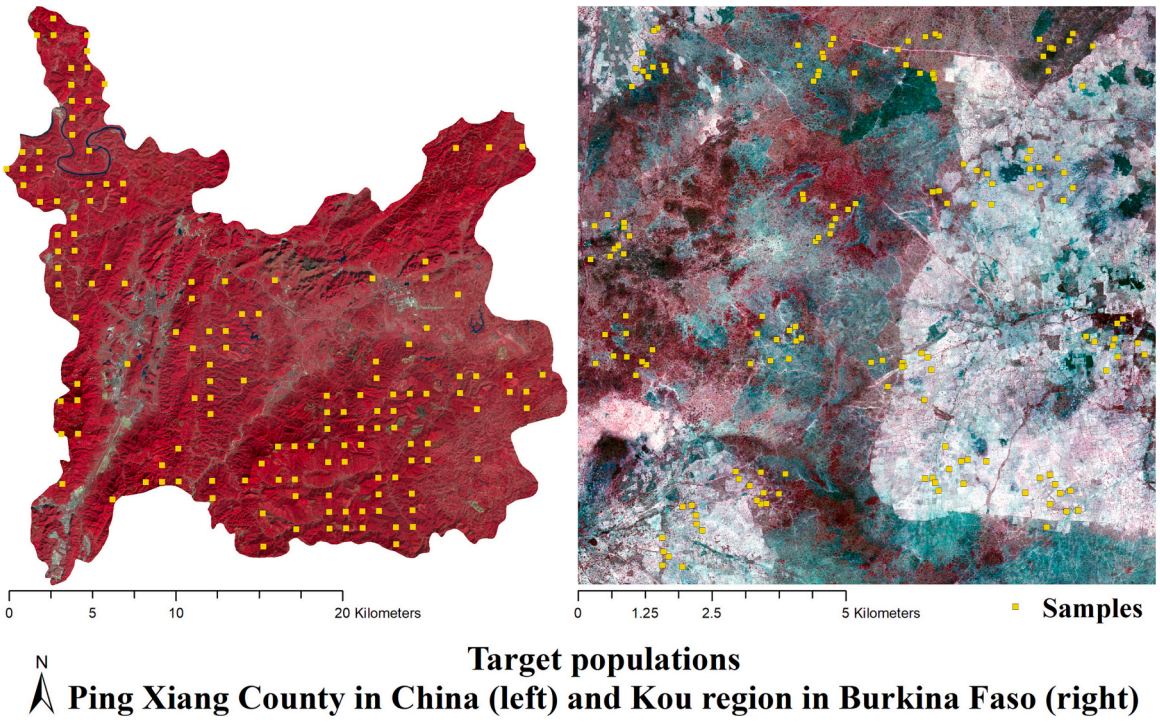

Fig. 1. Study areas in China and Burkina Faso.

Table 1
Spectral indices calculated from multispectral data.

Spectral indices	Formula
Enhanced vegetation index (EVI)	$2.5(NIR-R)/(NIR+6R-7.5B+1)$
Generalized Difference Vegetation Index (GDVI)	$(NIR^2-R^2)/(NIR^2+R^2)$
Normalized Difference Vegetation Index (NDVI)	$(NIR-R)/(NIR+R)$
Normalized Difference Water Index (NDWI)	$(NIR-SWIR2)/(NIR+SWIR)$
* Specific Leaf Area Vegetation Index (SLAVI)	$NIR/(R+SWIR2)$
* Simple Ratio (SR)	NIR/R

* Not available for RE in the absence of SWIR and SWIR2 bands. NIR, near infrared band; R, red band; B, blue band; SWIR, short-wave infrared band.

of multiple VOIs with seemingly unrelated regressions (SUR). This is not only because forest inventory is multivariate in nature, and the SUR model system serves MA in Section 3.3, but also the constructed SUR model system readily provides vectors and matrices required by BLUP in Section 3.5. Fourth, within the framework of MA, Section 3.5 develops the statistical structure of BLUP, which leads to the DAMA estimators. Clearly, Sections 3.3 to 3.5 together build DAMA. Finally, Section 3.6 introduces a measurement statistic for comparing the inference precision of the different estimators in the previous sections. In MA and DAMA estimators, remote sensing plays a central role in reducing variance.

3.2. Design-based inference

Design-based inference is a category of statistical procedures with an objective of estimating the unknown population parameter, such as the mean, μ, of a VOI, using a probability sample selected from the target population. In forest inventory, this population is finite and can be any spatial area of interest that is tessellated with smaller grids of a given size serving as population elements. A VOI can be any forest attribute such as stem volume, basal area, biomass, or carbon. The probability sample is selected from the population by a sampling design (or design), and for each element in the sample, i.e., sample plot, the VOI value is to be observed.

Essentially, design is the function defining a probability distribution, $p(\bullet)$, on the set of samples that are possible to obtain. Different designs may lead to different estimators, but the HT estimator, alias π-estimator, works for any design (Särndal et al., 1992, p.43). Since a design assigns a positive inclusion probability, π, to each element, the π-estimator can estimate μ by expanding the observed VOI values by their respective πs. The π-estimator has a general form of

$$\hat{\mu}_\pi = \frac{1}{N} \bullet \sum_{k\in u} \frac{y_k}{\pi_k} \tag{1}$$

with an unbiased estimator for its variance, $Var(\hat{\mu}_\pi)$, as

$$\widehat{Var}(\hat{\mu}_\pi) = \frac{1}{N^2} \bullet \sum_{k\in u}\sum_{l\in u} \left(\frac{\Delta_{kl}}{\pi_{kl}} \bullet \frac{y_k}{\pi_k} \bullet \frac{y_l}{\pi_l} \right) \tag{2}$$

where N is the population size; for the k- or l-th element in the sample u, y_k is the observed VOI value; $\pi_k = \sum_{k\in u} p(u)$ is the inclusion probability for element k; $\pi_{kl} = \sum_{k\&l\in u} p(u)$ is the joint probability of inclusion for elements k and l; $\Delta_{kl} = \pi_{kl} - \pi_k\pi_l$ for $k \neq l$, and $\Delta_{kk} = \pi_k(1-\pi_k)$ for $k = l$. The π-estimator is design-unbiased, i.e., $E(\hat{\mu}_\pi) = \mu$, but $\widehat{Var}(\hat{\mu}_\pi)$ depends on the specific design, sample size, and which elements happen to be included in the sample.

In this study, we consider sampling designs that are common in NFI programs and operational forest management planning with estimators reduced from the π-estimator, including simple random sampling (SRS) and two-stage sampling (TS). Thorough descriptions about respective designs and mathematical derivations are available in classic sampling textbooks (Cochran, 1977; Gregoire and Valentine, 2007; Särndal et al., 1992; Thompson, 2012).

With SRS without replacement, μ can be estimated using an expansion estimator (Cochran, 1977). This estimator is simple to use, even

compatible with alternative designs (Magnussen et al., 2020). The expansion estimator takes the form

$$\hat{\mu}_1 = \frac{1}{n} \bullet \sum_k^n y_k \tag{3}$$

with its variance estimator as

$$\widehat{Var}(\hat{\mu}_1) = \frac{N-n}{N} \bullet \frac{\sigma^2}{n} \tag{4}$$

where n is the sample size; and $\sigma^2 = \sum_{k\in u}(y_k - \hat{\mu}_1)^2/(n-1)$ is the sample variance. The finite-population correction factor, $(N-n)/N$, has effect on reducing $\widehat{Var}(\hat{\mu}_1)$, but can be omitted for populations that are large relative to the sample size. For the China population in Section 2.1, $N = 719,624$ and $n = 145$. NFI programs in Belgium and China use this estimator for official reporting although the sample is (quasi) systematic (Tomppo et al., 2010).

With TS, the population of N elements is partitioned into primary and secondary sampling units (PSUs and SSUs) with sampling required at both stages. Let C denote the number of PSUs in the population and S_i the number of SSUs in the i-th PSU, so $N = \sum_i^C S_i$. A first stage sample of size c is drawn from PSUs, and then a second stage sample of size s_i is drawn from the i-th PSU in c, so $n = \sum_i^c s_i$. When SRS without replacement is applied at each stage, the TS estimator takes the form (Särndal et al., 1992, p.142)

$$\hat{\mu}_2 = \frac{1}{N} \bullet \frac{C}{c} \bullet \sum_i^c \hat{\tau}_i \tag{5}$$

with its variance estimator as

$$\widehat{Var}(\hat{\mu}_2) = \frac{1}{N^2}\left[C(C-c)\frac{\sigma_I^2}{c}\right] + \frac{1}{N^2}\left[\frac{C}{c} \bullet \sum_i^c S_i(S_i - s_i)\frac{\sigma_i^2}{s_i}\right] \tag{6}$$

where $\hat{\tau}_i = \frac{S_i}{s_i}\sum_k^{s_i} y_{ik}$ estimates the total y-value of the i-th PSU; $\sigma_I^2 = \sum_i^c(\hat{\tau}_i - \bar{\hat{\tau}}_i)^2/(c-1)$ with $\bar{\hat{\tau}}_i = \sum_i^c \hat{\tau}_i/c$; and $\sigma_i^2 = \sum_k^{s_i}(y_{ik} - \bar{y}_i)^2/(s_i - 1)$ with $\bar{y}_i = \sum_k^{s_i} y_{ik}/s_i$. σ_I^2 is the sample variance among PSU totals, and σ_i^2 is the sample variance within the i-th PSU. $\widehat{Var}(\hat{\mu}_2)$ decomposes into two terms, indicating two sources of variation. The first term in Eq. (6) arises from the selection of PSUs and is the variance that would be obtained if all SSUs in a selected PSU were observed, while the second term arises from the subsampling of SSUs within selected PSUs and is the variance due to estimating τ_i. For the Burkina Faso population in Section 2.1, $N = C \bullet S = 115 \bullet 1035$ and $n = c \bullet s = 16 \bullet 10$. While TS is common in operational forest management planning, it is relatively rare for NFI programs with Croatia one of the few exceptions (Tomppo et al., 2010).

3.3. Model-assisted estimation

MA is a category of design-based inference that combines design and model considerations in estimation. Benefiting from remotely sensed auxiliary variables, the (generalized) regression estimator is an important class of MA estimators, seeking approximate unbiasedness under the design and using a variance estimator that is robust against departures from the assisting model. Resembling the π-estimator principle, the regression estimator has a general form of

$$\hat{\mu}_r = \frac{1}{N} \bullet \left(\sum_{k\in U} \hat{y}_k + \sum_{k\in u} \frac{e_k}{\pi_k}\right) \tag{7}$$

where U denotes the population of elements; and $\hat{y}_k$ is the prediction for element k using a model fit to sample data; and for $k \in u$, $e_k = y_k - \hat{y}_k$ is the residual.

Through Taylor linearization, the approximate variance of $\hat{\mu}_r$ has a general form of

$$Var(\hat{\mu}_r) = \frac{1}{N^2} \bullet \sum_{k\in U}\sum_{l\in U}\left(\Delta_{kl} \bullet \frac{E_k}{\pi_k} \bullet \frac{E_l}{\pi_l}\right) \tag{8}$$

where the error E_k results from a population fit model; Δ_{kl} and π_k are defined in Eq. (2). Subject to different perspectives on the unobservable E_k, two variance estimators are available based on large sample approximations. The first one uses a predicted g-weight, g_k, to calibrate e_k due to utilizing a sample fit rather than a population fit model (Särndal et al., 1992, Eq. 6.6.4, p.235),

$$\widehat{Var}(\hat{\mu}_r)_1 = \frac{1}{N^2} \bullet \sum_{k\in u}\sum_{l\in u}\left(\frac{\Delta_{kl}}{\pi_{kl}} \bullet \frac{g_k e_k}{\pi_k} \bullet \frac{g_l e_l}{\pi_l}\right) \tag{9}$$

The second one simply replaces the unobservable E_k by the observable e_k, leading to (Särndal et al., 1992, Eq. 6.6.11, p.237)

$$\widehat{Var}(\hat{\mu}_r)_2 = \frac{1}{N^2} \bullet \sum_{k\in u}\sum_{l\in u}\left(\frac{\Delta_{kl}}{\pi_{kl}} \bullet \frac{e_k}{\pi_k} \bullet \frac{e_l}{\pi_l}\right) \tag{10}$$

Relative to $\widehat{Var}(\hat{\mu}_r)_1$, $\widehat{Var}(\hat{\mu}_r)_2$ is slightly less efficient without the g-weight calibration. However, the derivation for a g-weight predictor is nontrivial due to its design- and linear model-specific features. Regardless, the g-weight is asymptotically $g_k = 1$ and often omitted to ease computation, in practice making $\widehat{Var}(\hat{\mu}_r)_1$ collapse to $\widehat{Var}(\hat{\mu}_r)_2$ (Gregoire et al., 2011; Mandallaz, 2015; Särndal et al., 1992, Chapters 6–8).

We directly employ $\widehat{Var}(\hat{\mu}_r)_2$, because it works for any assisting parametric models such as linear, nonlinear, generalized linear mixed model with some fixed, some random effects, and even works for semi- and non-parametric models (Kangas et al., 2016). This capability to handle such a wide array of model types greatly increases the utility of MA in forest inventory. Breidt and Opsomer (2017) reviewed many of these prediction techniques for MA.

Consequently, the regression estimator counterpart to $\hat{\mu}_1$ takes the form (Särndal et al., 1992, Eq. 6.5.3, p.231)

$$\hat{\mu}_{r1} = \frac{1}{N} \bullet \sum_k^N \hat{y}_k + \frac{1}{n} \bullet \sum_k^n e_k \tag{11}$$

with its variance estimator as

$$\widehat{Var}(\hat{\mu}_{r1}) = \frac{N-n}{N} \bullet \frac{\sigma_e^2}{n} \tag{12}$$

where $\sigma_e^2 = \sum_{k\in u}(e_k - \sum_k^n e_k/n)^2/(n-1)$ is the residual variance.

The regression estimator counterpart to $\hat{\mu}_2$ takes the form (Särndal et al., 1992, Eq. 8.9.5, p.323)

$$\hat{\mu}_{r2} = \frac{1}{N} \bullet \sum_k^N \hat{y}_k + \frac{1}{N} \bullet \frac{C}{c} \bullet \sum_i^c \hat{\tau}_{ei} \tag{13}$$

with its variance estimator as

$$\widehat{Var}(\hat{\mu}_{r2}) = \frac{1}{N^2}\left[C(C-c)\frac{\sigma_{eI}^2}{c}\right] + \frac{1}{N^2}\left[\frac{C}{c} \bullet \sum_i^c S_i(S_i - s_i)\frac{\sigma_{ei}^2}{s_i}\right] \tag{13}$$

where $\hat{\tau}_{ei} = \frac{S_i}{s_i}\sum_k^{s_i} e_{ik}$ estimates the total residual-value of the i-th PSU; $\sigma_{eI}^2 = \sum_i^c(\hat{\tau}_{ei} - \bar{\hat{\tau}}_{ei})^2/(c-1)$ with $\bar{\hat{\tau}}_{ei} = \sum_i^c \hat{\tau}_{ei}/c$; and $\sigma_{ei}^2 = \sum_k^{s_i}(e_{ik} - \bar{e}_i)^2/(s_i - 1)$ with $\bar{e}_i = \sum_k^{s_i} e_{ik}/s_i$.

3.4. Modeling at the element level with seemingly unrelated regressions

With the regression estimators, the remote sensing-based model is not required to be "true" in terms of correctly depicting some process by which the population is generated. If the population is well-described by an assumed model, the regression estimator normally brings a large variance reduction; otherwise, the reduction may be modest, but the

regression estimator still is approximately unbiased (Särndal et al., 1992, p.238).

With multivariate forest inventory, we are interested in surveying multiple VOIs instead of just one. When these VOIs are correlated, the seemingly unrelated regressions (SUR), readily available in R-package "systemfit" (Henningsen and Hamann, 2022), can be used to construct a model system for making simultaneous predictions (Zellner, 1962). The SUR model system takes the form

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_K \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{X}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{X}_K \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_K \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_K \end{bmatrix} \tag{15}$$

i.e., $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$, where K is the number of submodels; $\boldsymbol{y}_k$ is the vector of the kth VOI, $k = \{1, 2, \ldots, K\}$, $\boldsymbol{y}_k' = [y_{k1} \cdots y_{kn}]$ with n indicating the number of field observations; $\boldsymbol{X}_k$ is the design matrix of remotely sensed variables in the form of a block diagonal matrix, $\boldsymbol{\beta}_k$ the vector of model parameters, and $\boldsymbol{e}_k$ the vector of errors.

The SUR model system captures VOI dependencies through a joint variance-covariance matrix of the system's errors, $\boldsymbol{e}$, which is a multivariate normal vector with $\mathrm{E}(\boldsymbol{e}) = \boldsymbol{0}$ and variance-covariance matrix,

$$Var(\boldsymbol{e}) = \begin{bmatrix} \sigma_{11}\boldsymbol{I}_n & \sigma_{12}\boldsymbol{I}_n & \cdots & \sigma_{1K}\boldsymbol{I}_n \\ \sigma_{21}\boldsymbol{I}_n & \sigma_{22}\boldsymbol{I}_n & \cdots & \sigma_{2K}\boldsymbol{I}_n \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1}\boldsymbol{I}_n & \sigma_{K2}\boldsymbol{I}_n & \cdots & \sigma_{KK}\boldsymbol{I}_n \end{bmatrix}$$

$$= \boldsymbol{\Sigma}_{K\times K} \bigotimes \boldsymbol{I}_n \tag{16}$$

where $\boldsymbol{\Sigma}_{K\times K}$ is a positive definite symmetric matrix comprising submodel-specific error variances and cross-submodel error covariances. The $\boldsymbol{\beta}$ vector can be estimated using feasible generalized least squares by estimating each element in $\boldsymbol{\Sigma}_{K\times K}$ with a consistent estimator, $\widehat{\sigma}_{ij} = \frac{1}{n}\widehat{e}_i'\widehat{e}_j$, where $\widehat{e}_i$ results from an initial OLS fit, and then by inserting $\widehat{\boldsymbol{\Sigma}}_{K\times K}$ into $\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}_{K\times K}^{-1} \bigotimes \boldsymbol{I}_n\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}_{K\times K}^{-1} \bigotimes \boldsymbol{I}_n\boldsymbol{y}$ (Hansen, 2007).

Assisting models were constructed separately and then combined as sub-models for estimating $\boldsymbol{\Sigma}_{K\times K}$ and $\boldsymbol{\beta}$ of the SUR model system. This system comprises six sub-models, with the first two predicting respectively stem volume and basal area using RapidEye independent variables and the last four predicting residuals (Appendix Table A1). We use RE for modeling VOIs because preliminary analyses indicate a slightly greater prediction accuracy for RE than L8 (Hou et al., 2018). Independent variables were selected parsimoniously using the "bootstrap stepAIC" procedure (Rizopoulos, 2022).

3.5. Data assimilation at the element level with BLUP

For the SUR model system, the residuals of a submodel can be decreased with correlated residuals of other submodels. Although the regression estimators do not rely on a "true" model, the smaller the residuals, the smaller the variance of the regression estimators. Therefore, we resort to a calibration procedure of DA developed from BLUP to achieve this purpose. This procedure was introduced in Hou et al. (2019) for which an example with code is readily available (Hou, 2019).

BLUP calibrates the parameters of the marginal distribution of a random variable to the parameters of its conditional distribution given the jointly distributed variables observed. The SUR model system and its $\widehat{\boldsymbol{\Sigma}}_{K\times K}$ provide all required components for implementing BLUP. Assume that a random vector, $\boldsymbol{y}$, is multivariate normal and partitioned into two parts, $\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_a \\ \boldsymbol{y}_b \end{bmatrix}$, where $\boldsymbol{y}_a$ and $\boldsymbol{y}_b$ are two sub-vectors, each of which comprises at least one SUR sub-model y-vector of VOIs. Given that the first and second order properties of these sub-vectors are known, i.e. $E(\boldsymbol{y}_a) = \boldsymbol{\mu}_a, E(\boldsymbol{y}_b) = \boldsymbol{\mu}_b$, $Var(\boldsymbol{y}_a) = \boldsymbol{\Sigma}_a$, $Var(\boldsymbol{y}_b) = \boldsymbol{\Sigma}_b$, and $Cov(\boldsymbol{y}_a, \boldsymbol{y}_b') = \boldsymbol{\Sigma}_{ab}$, these relationships denote $\begin{bmatrix} \boldsymbol{y}_a \\ \boldsymbol{y}_b \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_a & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ab}' & \boldsymbol{\Sigma}_b \end{bmatrix} \right)$ (McCulloch and Searle, 2001).

If $\boldsymbol{y}_b$ is observed, the best linear predictor (BLP) of $\boldsymbol{y}_a$ is the conditional expectation $BLP(\boldsymbol{y}_a) = E(\boldsymbol{y}_a|\boldsymbol{y}_b)$ and $(\boldsymbol{y}_a|\boldsymbol{y}_b) \sim N(\boldsymbol{\mu}_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_b^{-1}(\boldsymbol{y}_b - \boldsymbol{\mu}_b), \boldsymbol{\Sigma}_a - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_b^{-1}\boldsymbol{\Sigma}_{ab}')$. If $\boldsymbol{\mu}_a$, $\boldsymbol{\mu}_b$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_a & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ab}' & \boldsymbol{\Sigma}_b \end{bmatrix}$ are replaced by their estimates extracted from the SUR model system, i.e., the sub-models in $\boldsymbol{X}\widehat{\boldsymbol{\beta}}$ and the corresponding submatrices in $\widehat{\boldsymbol{\Sigma}}_{K\times K}$ as demonstrated in Hou et al. (2019, Appendix), BLUP is the resulting empirical predictor,

$$\widehat{\boldsymbol{y}}_{a.BLUP} = \widehat{\boldsymbol{\mu}}_a + \widehat{\boldsymbol{\Sigma}}_{ab}\widehat{\boldsymbol{\Sigma}}_b^{-1}(\boldsymbol{y}_b - \widehat{\boldsymbol{\mu}}_b) \tag{17}$$

with

$$\widehat{Var}(\widehat{\boldsymbol{y}}_{a.BLUP}) = \widehat{\boldsymbol{\Sigma}}_a - \widehat{\boldsymbol{\Sigma}}_{ab}\widehat{\boldsymbol{\Sigma}}_b^{-1}\widehat{\boldsymbol{\Sigma}}_{ab}' \tag{18}$$

Note that the second terms in $\widehat{\boldsymbol{y}}_{a.BLUP}$ and $\widehat{Var}(\widehat{\boldsymbol{y}}_{a.BLUP})$ are respectively the calibration term for $\widehat{\boldsymbol{\mu}}_a$ and $\widehat{\boldsymbol{\Sigma}}_a$, forming the basis of DA calibration for the SUR model system at the element level. Because the conditional expectation of $\boldsymbol{y}_a$ given $\boldsymbol{y}_b$, $E(\boldsymbol{y}_a|\boldsymbol{y}_b)$, is always the best predictor that has the smallest variance (Mehtätalo and Lappi, 2020, p.60; Robinson, 1991), the calibrated residuals, $\boldsymbol{e}_{a.BLUP} = \boldsymbol{y}_a - \widehat{\boldsymbol{y}}_{a.BLUP}$, are theoretically smaller than the original residuals resulting from the SUR model system, $\boldsymbol{e}_a = \boldsymbol{y}_a - \widehat{\boldsymbol{y}}_a$, thus leading directly to DAMA estimators that are counterparts to Eq. (11) to Eq. (14). Because the SUR model system supports simultaneous predictions for multiple VOIs, DA supports simultaneous estimation with extended vectors and matrices.

For implementing Eq. (17), wall-to-wall observations are required in the form of $\boldsymbol{y}_b$. This is conveniently achieved with remote sensing in three steps. First, regard residuals resulting from predicting stem volume or basal area as a new variable and model them with either RE or L8 independent variables that are available wall-to-wall. Second, predict wall-to-wall residuals and regard these predicted residuals as the dependent variable of an intercept model for maintaining correlation with the residuals resulting from predicting the VOI. Third, combine all sub-models and estimate parameters for the SUR model system. Thereby, the last four sub-models predict respectively RE predicted residuals for stem volume, RE predicted residuals for basal area, L8 predicted residuals for stem volume, and L8 predicted residuals for basal area. With these four sub-models whose observations and predictions are available wall-to-wall, we can calibrate the predicted values, or equivalently, the residuals for any one of the first two sub-models, using RE, L8 or both.

3.6. Comparison of inferential precision

With design-based inference, the smaller the coefficient of variation, $CV_{\%}$, the greater the inferential precision, or equivalently, the less the inferential uncertainty. $CV_{\%}$ enables comparisons among estimators by quantifying the uncertainty on a percentage basis, taking the form

$$CV_{\%} = 100 \times \frac{\sqrt{\widehat{Var}(\widehat{\mu})}}{\widehat{\mu}} \tag{19}$$

The $CV_{\%}$, officially used by NFI programs such as in China and the USA, can further be used to approximate 67% and 95% confidence intervals under a normality assumption for the sampling distribution of estimates (Bechtold and Patterson, 2005).

4. Results and discussion

4.1. Comparing prediction and data assimilation at the element level

Flexibility is a feature of the constructed SUR model system because it supports both individual and simultaneous prediction for VOIs. Both the original prediction (i.e., SUR model prediction) and the calibrated prediction (i.e., SUR model prediction calibrated using BLUP) are summarized in Table 2. Other seemingly trivial but useful details are reported in the Appendix, including the SUR models constructed for the populations in China and Burkina Faso (Table A1), empirical densities and correlations for the residual errors (Fig. A1 and A2), and the wall-to-wall observations produced for sub-models 3–6 using RE or L8 (Table A2).

Table 2 compares the original and calibrated prediction accuracy evaluated with root mean square error (RMSE). Three findings are noteworthy. First, DA increased the prediction accuracy on average 9.5% from the original, regardless of population or VOI, complying with the theory that the conditional expectation is always the best predictor that has the smallest variance, also indicating the validity of the proposed DA as supported by Hao et al. (2022) using other populations and VOIs. Second, while the joint use of RE and L8 brings the largest increase to prediction accuracy, L8 alone ensured an increase in prediction accuracy, free of charge, thus markedly increasing the utility of incorporating this DA procedure for thematic mapping. This favorable effect stems from the spectral band complementarity of RE and L8, which provides additional information that can be used to better predict residuals. Third, the key to DA is $\widehat{\boldsymbol{\Sigma}}_{K\times K}$, i.e., the correlation between residuals for VOIs. While variables explaining the remaining residual variability are probably common albeit unfound, DA circumvents this difficulty by using the unexplained variability in the form of residual covariances to increase the prediction accuracy, using the whole set of RE and L8 (sub-models 3–6) or the subset of RE (sub-models 3–4) or L8 (sub-models 5–6).

4.2. Comparing estimators at the population level

Table 3 summarizes estimates for the HT, MA and DAMA estimators. There are four major findings. First, DAMA estimators were the most efficient in the respective designs, suggesting similar performances expected for other designs beyond SRS and TS. With SRS, DAMA estimator increased the inferential precision on average 14% and 7% compared to the HT and MA counterparts; and with TS, remarkably, 180% and 57%.

Table 2
Comparison for the original and DA calibrated prediction accuracy.

Pop.	Model	Dependent variable	Prediction	RMSE	RMSE %
China	Submodel-1	Stem volume (m^3/ha)	Original	63.60	60.99
			DA with RE & L8	56.95	54.62
			DA with RE	61.58	59.05
			DA with L8	58.11	55.73
	Submodel-2	Basal area (m^2/ha)	Original	7.77	48.46
			DA with RE & L8	6.96	43.42
			DA with RE	7.47	46.59
			DA with L8	7.14	44.54
Burkina Faso	Submodel-1	Stem volume (m^3/ha)	Original	9.04	53.29
			DA with RE & L8	7.96	46.94
			DA with RE	8.32	49.02
			DA with L8	8.26	48.67
	Submodel-2	Basal area (m^2/ha)	Original	2.6	48.69
			DA with RE & L8	2.16	40.31
			DA with RE	2.31	43.23
			DA with L8	2.26	42.34

Table 3
Comparing the HT, MA and DAMA estimators for SRS and TS.

VOI	Group	SRS (China)			TS (Burkina Faso)		
		$\widehat{\mu}$	$\widehat{Var}(\widehat{\mu})$	$CV_{\%}$	$\widehat{\mu}$	$\widehat{Var}(\widehat{\mu})$	$CV_{\%}$
Stem volume (m^3/ha)	HT	104.28	34.21	5.61	16.96	4.02	11.82
	MA	101.39	28.09	5.23	17.05	0.95	5.72
	DAMA	98.09	22.53	4.84	16.99	0.27	3.06
Basal area (m^2/ha)	HT	16.03	0.51	4.45	5.35	0.43	12.26
	MA	15.53	0.42	4.17	5.41	0.11	6.13
	DAMA	15.07	0.34	3.85	5.41	0.03	3.20

Second, differences among point estimates were negligible, suggesting approximate unbiasedness holds valid for these design-based estimators. Third, although model predictions and calibrated predictions were only intermediate products enroute to inferences for population parameters, the inferential precision of MA and DAMA increased as the prediction accuracy increases. Fourth, it is the design that decides the route of residual propagation in MA. The routes for MA and DAMA to reducing $\widehat{Var}(\widehat{\mu})$ are through decreasing residual variances in Eq. (12) and Eq. (14), with the reduction rate faster for TS than SRS.

4.3. Effects of two-stage design

Model-assisted estimation relies on design, and it is well noted that for SRS the inferential precision increases as the sample size increases, so our primary focus is TS. The effects of stage partitioning were evaluated for respective TS estimators, with resulting estimates summarized in Table 4. In general, estimates were only negligibly different for point estimators, but substantially greater for variance estimators, indicating that these effects manifested themselves by affecting the precision rather than the approximate unbiasedness of design-based inference.

There are three relevant findings. First, the partition into PSU and SSU affected the inferential precision, particularly for HT estimator, but not as much for MA and DAMA, indicating an improved consistency of model-assisted estimation. Second, DAMA enhanced TS cost-efficiency because of deflation to the two variance components stemming from the tradeoff between PSU and SSU sizes. The size here refers to the number of selectable PSUs or SSUs, not the areal extent. The larger the PSU size, the smaller the SSU size, and thus the greater the between-PSUs heterogeneity that inflates the first variance term, and meanwhile the greater the within-PSU homogeneity that deflates the second variance term (Table 4). For reducing inventory cost, PSU partitions are large in practice, typically on the order of dozens to hundreds for a population. Although such partition is not necessarily VOI-specific or optimal, DAMA appears efficient in a way that the uncertainty ascribed to the tradeoff becomes marginal, paving a cost-effective path for TS towards inventorying multiple VOIs. Third, design flexibility at respective stages further improved the inferential precision for TS. Although SRS was examined at each stage, TS supports more efficient designs at each stage to deflate the first and second variance terms for DAMA. Generalized from TS, multistage sampling (Mandallaz, 2015) cohering with the hierarchy of a large population such as globe, region, country, province, state, and county could be a viable framework harmonizing international reporting instruments.

TS is a converted stratified sampling that circumvents classification. As the PSU size decreases, the first variance term in Eq. (6) diminishes for TS, and the second variance term equals the variance of stratified sampling (Särndal et al., 1992, p.137, Remark 4.3.2). A main difference between stratified sampling and TS resides in the classification or partitioning being contextual or not. Contextual classification uses classifiers developed with pattern recognition algorithms for strata generation (Swain et al., 1981; Phiri and Morgenroth, 2017). Misclassification or a wrong contextual variable used in classification deflates the inferential precision. Even with an ideal classification and contextual variable, strata are still VOI-specific with resulting challenges interpreted as

Table 4
Effects of stage partitioning on TS estimators with variance decomposition.

VOI	Two-stage Partition		Estimation				Variance Decomposition	
	PSU	SSU	Group	$\widehat{\mu}$	$\widehat{Var}(\widehat{\mu})$	$CV_{\%}$	Term 1	Term 2
Stem volume (m^3/ha)	115	1035	HT	16.96	4.02	11.82	3.9263	0.0904
			MA	17.05	0.95	5.72	0.8841	0.0635
			DAMA	16.99	0.27	3.06	0.2166	0.0571
	25	4761	HT	16.96	2.06	8.46	1.6419	0.4189
			MA	17.05	0.66	4.76	0.3697	0.2941
			DAMA	16.99	0.36	3.53	0.0906	0.2645
Basal area (m^2/ha)	115	1035	HT	5.35	0.43	12.26	0.4234	0.0064
			MA	5.41	0.11	6.13	0.1013	0.0048
			DAMA	5.41	0.03	3.20	0.0281	0.0040
	25	4761	HT	5.35	0.21	8.57	0.1771	0.0297
			MA	5.41	0.06	4.53	0.0424	0.0223
			DAMA	5.41	0.03	3.20	0.0118	0.0184

reduced cost-efficiency. However, TS does not necessarily partition a population as per a contextual variable but in a flexible way that circumvents the detrimental effects associated with classification. The DAMA estimator would in such scenario help to improve the inferential precision.

4.4. Effects of remotely sensed data assimilation

Unbiasedness is a desired property of an estimator or inventory procedure, crucial to the credibility of any reporting instruments (Eggleston et al., 2006). While Hou et al. (2019) devised a model-unbiased DA procedure for model-based inference, DAMA is devised for model-assisted estimation. The approximate design-unbiasedness of DAMA is rooted in the unbiasedness of BLUP (Robinson, 1991), and the approximate design-unbiasedness of MA (Särndal et al., 1992). As per Table 5, remotely sensed auxiliary variables associated with DA do not affect the approximate unbiasedness of DAMA, regardless of populations, designs or VOIs.

As for the inferential precision, there are four findings. First, the greater the DA prediction accuracy, the greater the DAMA inferential precision (Tables 2 and 5). This is because a variance estimator is constructed on residuals for the model-assisted estimation. Second, the inferential precision was decided by the number of remotely sensed auxiliary variables in the form of sub-models 3–6 involved with DA. While the more the better, L8 ensured a minimum increase in precision, free of charge. Third, design affected the efficiency of DAMA more than the number of remotely sensed auxiliaries involved with DA. Fourth, DAMA is compatible with MA in a way that a sub-model can be used independently in MA or jointly in DAMA, extending the flexibility for ad- and post-hoc DAMA utilities.

DAMA is compatible with established and alternative designs. The convenience of DAMA resides in its compatibility with NFI programs by not imposing any modifications to sampling and field protocols, paving a path towards building a common reporting instrument by preserving comparability for estimates over years and among populations (Tomppo et al., 2010). DAMA also supports alternative designs including network sampling that is efficient for elusive populations or domains (Hou et al., 2022; Thompson, 2012; Xu et al., 2021).

Table 5
Effects of remotely sensed auxiliaries on DAMA.

VOI	Group	SRS (China)			TS (Burkina Faso)		
		$\widehat{\mu}$	$\widehat{Var}(\widehat{\mu})$	$CV_{\%}$	$\widehat{\mu}$	$\widehat{Var}(\widehat{\mu})$	$CV_{\%}$
Stem volume (m^3/ha)	RE&L8	98.09	22.53	4.84	16.99	0.27	3.06
	RE	99.99	26.33	5.13	16.65	0.39	3.75
	L8	98.78	23.45	4.90	17.37	0.47	3.95
Basal area (m^2/ha)	RE&L8	15.07	0.34	3.85	5.41	0.03	3.20
	RE	15.32	0.39	4.06	5.28	0.04	3.79
	L8	15.18	0.35	3.92	5.52	0.05	4.05

In terms of versatility, DAMA is expected to be more efficient when used with more effective remotely sensed alternatives such as airborne laser scanning (ALS). However, there are pros and cons to the coverage of auxiliary data and existing MA theory. First, with wall-to-wall coverage, ALS submodels would outperform RE or L8 counterparts, improving prediction accuracy at the element level and thus improving DAMA precision. Although further validation is needed, this conjecture is supported by extensive studies conducted over the past few decades in which there are relatively high correlations between ALS auxiliary variables and various forest attributes of interest (e.g., Maltamo et al., 2014). However, at present, due to the tradeoff between ALS coverage and cost, DAMA combined with ALS is more practical in most countries for small-scale inventories than for large-scale NFIs. Second, another option is to make a theoretical breakthrough in MA to derive new estimators with native support for non-wall-to-wall auxiliary data. This need guides our ongoing research aimed at relaxing the spatial scale constraints.

In forestry inventory, alternative DA procedures have been developed using BLUP to reduce variance. BLUP requires a covariance matrix and associated observations to update or calibrate. According to the origin of covariance, DA can be divided into three types: temporal, spatial, and cross-model DAs. First, temporal DA exploits serial correlation, including Kalman Filter and mixed estimator, and is often used to update annual estimates at the population or domain level (Ehlers et al., 2013; Hou et al., 2021). Second, spatial DA utilizes semivariograms, including Kriging and BLUP-calibrated spatial regression. The change of support from a population to a domain is convenient in spatial statistics, and calibrated parameters can be estimated by the block Kriging principle (Schabenberger and Gotway, 2005; Xu et al., 2023). Third, cross-model DA exploits residual covariance between models, as in DAMA for model-assisted estimation and Hou et al. (2019) for model-based inference. The higher the cross-model correlation through the residuals, the more effective the DA. This type of DA is attractive for forest inventories not only because of its ability to calibrate estimates using correlated residuals, but also because of its ability to simultaneously predict or estimate multiple VOIs in a population or domain. Finally, with the exception of DAMA, which is design-based and approximately design-unbiased, these DA procedures are model-based and model-unbiased. Mohamedou et al. (2022) reviewed recent DA development in forest inventory.

5. Conclusions

This study proposed and demonstrated a new procedure, DAMA, that conjugates remotely sensed data assimilation and model-assisted estimation to support high-precision multivariate forest inventory. Three conclusions are relevant. First, the DAMA estimator proposed is approximately design-unbiased with its variance affected by the sampling design, the prediction accuracy, and the type of remotely sensed

auxiliary data involved in DA, in descending order. Second, the DAMA estimator is the most efficient in respective designs compared to HT and MA estimators. With SRS, the DAMA estimator increased the inferential precision on average 14% and 7% compared to HT and MA counterparts; and with TS, remarkably, 180% and 57%. Third, the partition into PSUs and SSUs affects the inferential precision for TS, particularly for HT estimator, but not as much for MA and DAMA, indicating an improved consistency of model-assisted estimation. Overall, DAMA demonstrates considerable efficiency that would better serve natural resource monitoring and management, particularly when used with remotely sensed auxiliaries that are free of charge such as Landsat 8 and Sentinel.

CRediT authorship contribution statement

Zhengyang Hou: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Keyan Yuan:** Software, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Göran Ståhl:** Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Ronald E. McRoberts:** Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Annika Kangas:** Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Hao Tang:** Validation, Formal analysis, Investigation, Writing – original draft. **Jingyi Jiang:** Data curation, Validation, Formal analysis, Investigation. **Jinghui Meng:** Data curation, Formal analysis, Investigation. **Qing Xu:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Funding acquisition. **Zengyuan Li:** Validation, Formal analysis, Investigation.

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.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Social Science Fund of China (Grant No. 22BTJ005); and the Special Fund on Science and Technology Innovation of Xiong'an New Area, Ministry of Science and Technology of China (Grant No. 2022XACX1000). Dr. Qing Xu was also supported by the National Natural Science Foundation of China (Grant No. 32001252); and the International Center for Bamboo and Rattan (Grant No. 1632020029 and Grant No. 1632021024). Review comments provided by the editor and anonymous reviewers are sincerely acknowledged.

Appendix A

Table A1
Constructed SUR model systems.

Pop.	SUR model	Dep. variable	Ind. variable	Estimate	Std. Error	RMSE%
China	Submodel-1	vol/ha	(Intercept)	1489.48	446.13	60.99
			RE.Band3	−7.05	1.76	
			RE.NDVI	9993.95	3593.50	
			RE.GDVI	−5478.50	2074.22	
			RE.SR	−587.32	191.09	
	Submodel-2	ba/ha	(Intercept)	163.41	37.04	48.46
			RE.Band3	−0.94	0.23	
			RE.NDVI	1037.16	381.98	
			RE.GDVI	−559.07	200.59	
			RE.SR	−62.38	22.24	
	Submodel-3	RE predicted residuals for vol/ha	(Intercept)	0.12	1.24	
	Submodel-4	RE predicted residuals for ba/ha	(Intercept)	0.04	0.15	
	Submodel-5	L8 predicted residuals for vol/ha	(Intercept)	0.28	2.13	
	Submodel-6	L8 predicted residuals for ba/ha	(Intercept)	−0.08	0.23	
Burkina Faso	Submodel-1	vol/ha	(Intercept)	23.15	8.65	53.29
			RE.Band1	−0.25	0.02	
			RE.NDVI.variance	0.26	0.06	
			RE.SR.variance	−0.57	0.14	
	Submodel-2	ba/ha	(Intercept)	5.4	0.21	48.69
			RE.PCA	1.24	0.1	
	Submodel-3	RE predicted residuals for vol/ha	(Intercept)	-2.16E-14	0.26	
	Submodel-4	RE predicted residuals for ba/ha	(Intercept)	-4.09E-15	0.08	
	Submodel-5	L8 predicted residuals for vol/ha	(Intercept)	−3.97E-14	0.29	
	Submodel-6	L8 predicted residuals for ba/ha	(Intercept)	−4.09E-15	0.08	

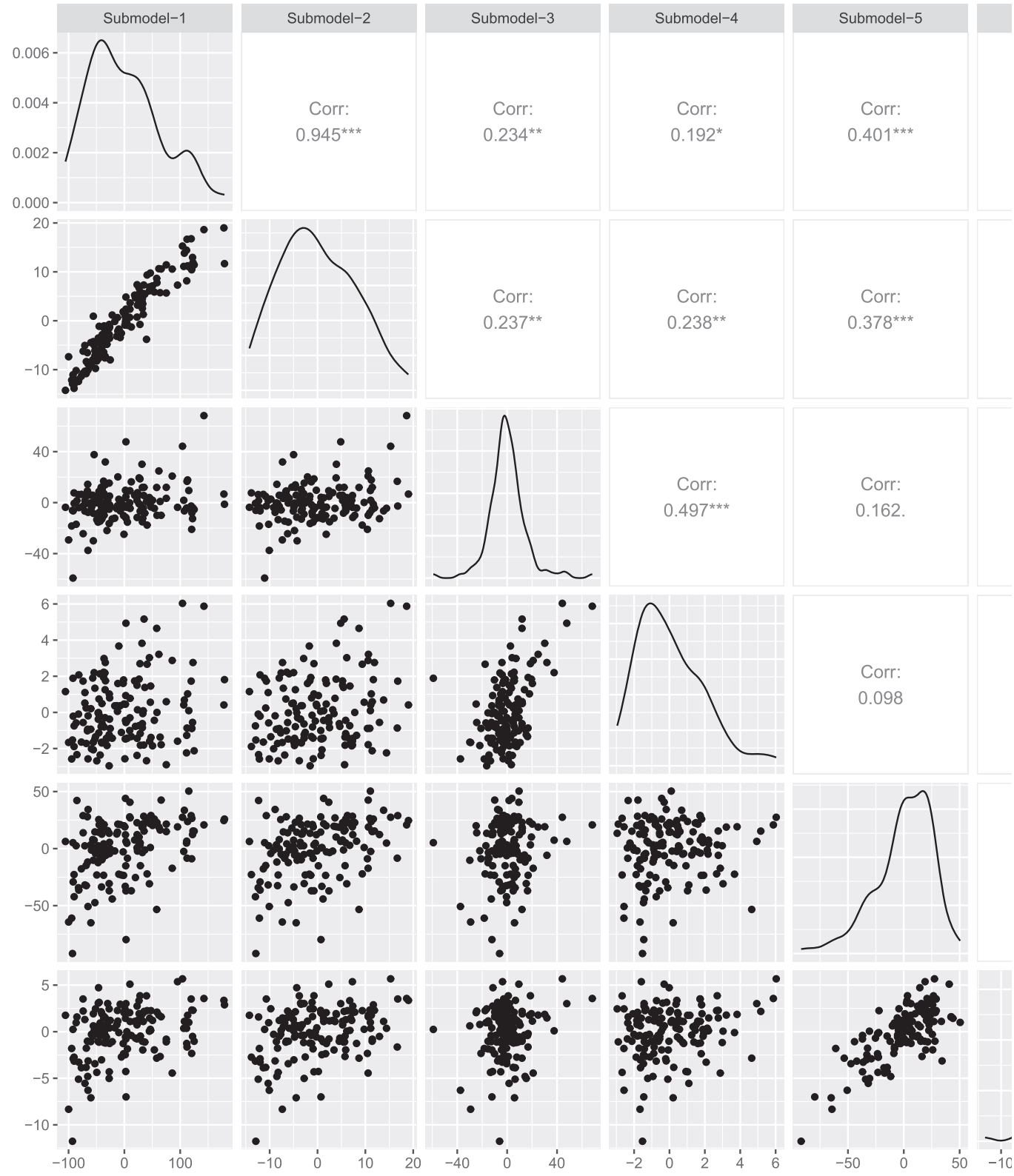

Fig. A1. Empirical density (diagonal) and correlation (off-diagonal) for the residuals of the constructed SUR model system for the population in China.

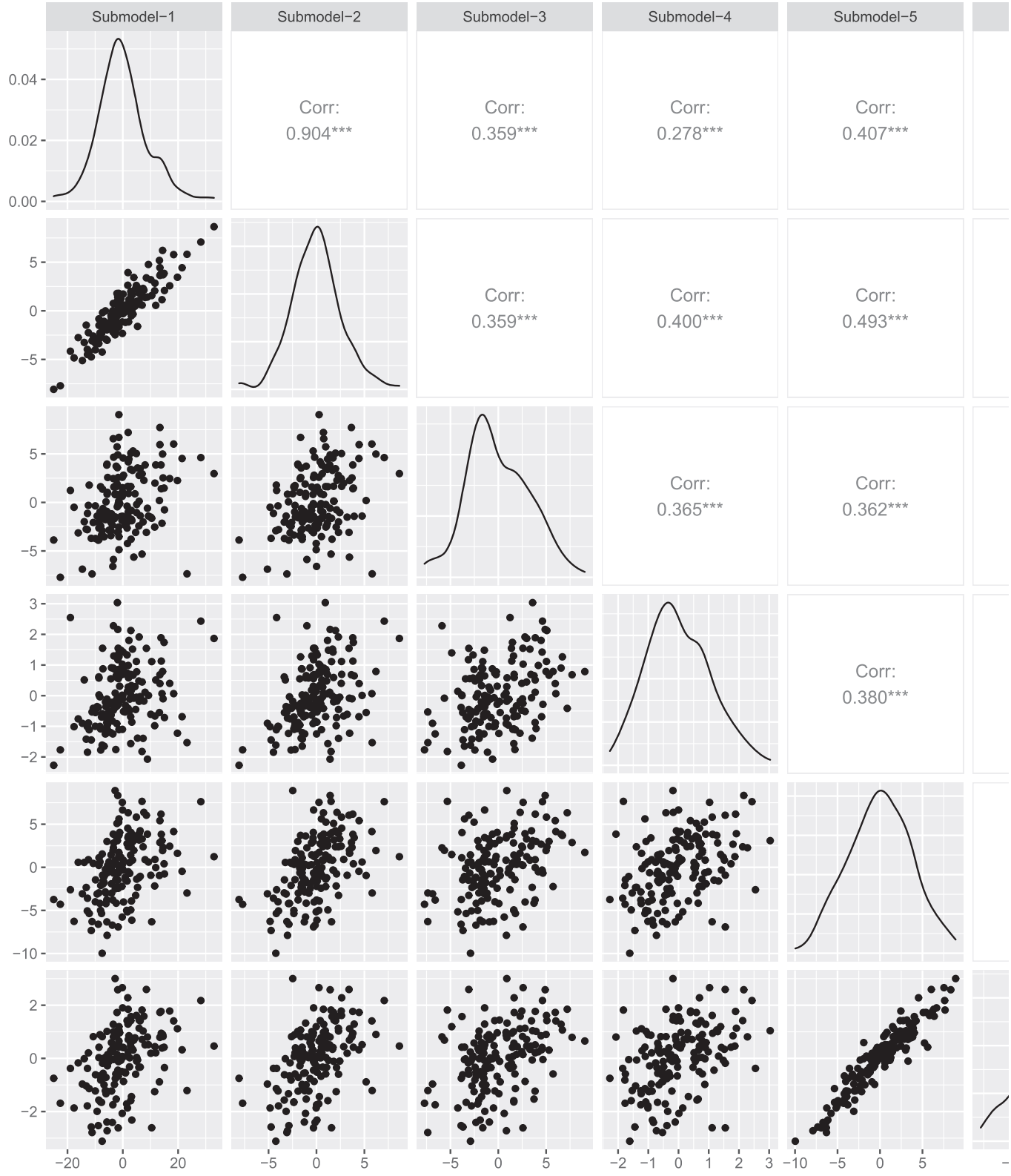

Fig. A2. Empirical density (diagonal) and correlation (off-diagonal) for the residuals of the constructed SUR model system for the population in Burkina Faso.

Table A2
Wall-to-wall observations produced for sub-models 3–6 using RE or L8.

Population	Sensor	Dependent variable	Independent variable	Estimate	Std. Error	R^2
China	RE	Residuals for vol/ha (w.r.t. submodel-3)	SR.entropy	−156.41	59.06	0.06
			SR.2ndmoment	−486.53	175.35	
			EVI.entropy	157.97	60.07	
			EVI.2ndmoment	481.97	168.69	
		Residuals for ba/ha (w.r.t. submodel-4)	SR.entropy	−156.41	59.06	0.06
			SR.2ndmoment	−486.53	175.35	
			EVI.entropy	157.97	60.07	
			EVI.2ndmoment	481.97	168.69	
	L8	Residuals for vol/ha (w.r.t. submodel-5)	PCA.variance	−2.10	0.48	0.16
			NDVI.dissimilarity	−115.65	38.13	
			GDVI.mean	289.99	60.63	
			GDVI.dissimilarity	87.21	34.50	
			NDWI.contrast	17.22	7.55	
		Residuals for ba/ha (w.r.t. submodel-6)	PCA.variance	−0.20	0.05	0.13
			NDWI	59.85	13.45	
			NDWI.contrast	3.60	1.33	
			SLAVI.contrast	−1.15	0.44	
Burkina Faso	RE	Residuals for vol/ha (w.r.t. submodel-3)	(Intercept)	11.68	5.63	0.13
			PCA.mean	20.45	7.33	
			PCA.correlation	−23.04	6.81	
			GDVI.correlation	23.25	7.84	
			NDVI.entropy	−6.38	1.97	
		Residuals for ba/ha (w.r.t. submodel-4)	(Intercept)	−8.39	2.21	0.16
			Band5	0.04	0.01	
			PCA.mean	15.83	2.98	
			SR.entropy	−1.86	0.43	
	L8	Residuals for vol/ha (w.r.t. submodel-5)	(Intercept)	−46.49	9.74	0.17
			Band7	0.011	0.003	
			SLAVI.mean	148.54	27.86	
			SLAVI.variance	−0.15	0.03	
		Residuals for ba/ha (w.r.t. submodel-6)	(Intercept)	−16.69	2.85	0.23
			Band7	0.004	0.001	
			NDWI.variance	0.02	0.01	
			SLAVI.variance	−0.05	0.01	
			SR.mean	30.95	7.16	

References

Baffetta, F., Fattorini, L., Franceschi, S., Corona, P., 2009. Design-based approach to k-nearest neighbors technique for coupling field and remotely sensed data in forest surveys. Remote Sens. Environ. 113, 463–475.

Bechtold, W.A., Patterson, P.L., 2005. The enhanced forest inventory and analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC.

Breidt, F.J., Opsomer, J.D., 2017. Model-assisted survey estimation with modern prediction techniques. Stat. Sci. 32, 190–205.

Cochran, W.G., 1977. Sampling Techniques. Wiley, New York.

Czaplewski, R.L., 1990. Kalman filter to update forest cover estimates. In State-of-the-art methodology of forest inventory. In: V.J. Labau, T. Cunia (Eds.). USDA For. Serv., Pacific Northwest Research Station. GTR PNW-263, pp. 457-465.

Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., 2006. IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4: Agriculture, Forestry and Other Land Use. Institute for Global Environmental Strategies, Hayama, Japan.

Ehlers, S., Grafström, A., Nyström, K., Olsson, H., Ståhl, G., 2013. Data assimilation in stand-level forest inventories. Can. J. For. Res. 43, 1104–1113.

Fletcher, S., 2017. Data Assimilation for the Geosciences: from Theory to Applications. Elsevier 1–4.

Gregoire, T.G., Næsset, E., McRoberts, R.E., Stahl, 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. For. Res. 41, 83–95.

Gregoire, T.G., Valentine, H.T., 2007. Sampling Strategies for Natural Resources and the Environment. CRC Press.

Hansen, C.B., 2007. Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. J. Econom. 140, 670–694.

Hao, Y., Widagdo, F.R.A., Liu, X., Quan, Y., Liu, Z., Dong, L., Li, F., 2022. Estimation and calibration of stem diameter distribution using UAV laser scanning data: a case study for larch (Larix olgensis) forests in Northeast China. Remote Sens. Environ. 268, 112769.

Haralick, R.M., Shanmugam, K., Dinsteein, J., 1973. Textural features for image classification. IEEE T. Syst. Man. Cyb. 3, 610–621.

Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423–447.

Henningsen, A., Hamann, J.D., 2022. Systemfit: Estimating Systems of Simultaneous Equations accessed 2022.06.13. http://www.systemfit.org.

Hou, Z., 2019. houzhengyang/BLUP_sample_code: BLUP_sample_code,Zenodo. https://doi.org/10.5281/zenodo.3364821.

Hou, Z., Domke, G., Russell, M., Coulston, J., Nelson, M., Xu, Q., McRoberts, R.E., 2021. Updating annual state- and county-level forest inventory estimates with data assimilation and FIA data. Forest Ecol. Manag. 483, 118777.

Hou, Z., McRoberts, R.E., Ståhl, G., Packalen, P., Greenberg, J., Xu, Q., 2018. How much can natural resource inventory benefit from a finer resolution auxiliary data? Remote Sens. Environ. 209, 31–40.

Hou, Z., McRoberts, R.E., Zhang, C., Ståhl, G., Zhao, Wang, X., Li, B., Xu, Q., 2022. Cross-classes domain inference with network sampling for natural resource inventory. Forest Ecosyst. 9, 100029.

Hou, Z., Mehtatalo, L., McRoberts, R.E., Ståhl, G., Tokola, T., Rana, P., Siipilehto, J., Xu, Q., 2019. Remote sensing-assisted data assimilation and simultaneous inference for forest inventory. Remote Sens. Environ. 234, 111431.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Basic Eng. T ASME 82 (Series D), 35–45.

Kangas, A., Myllymäki, M., Gobakken, T., Næsset, E., 2016. Model-assisted forest inventory with parametric, semiparametric, and nonparametric models. Can. J. For. Res. 46, 855–868.

Lahoz, W., Khattatov, B., Menard, R., 2010. Data Assimilation: Making Sense of Observations. Springer.

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. Forest Ecosyst. 7, 17.

Maltamo, M., Næsset, E., Vauhkonen, J., 2014. Forestry applications of airborne laser scanning: concepts and case studies. In: Managing Forest ecosystems 27. Springer, Dordrecht, the Netherlands.

Mandallaz, D., 2015. Mathematical details of two-phase/two-stage and three-phase/two-stage regression estimators in forest inventories. Design-based Monte Carlo approach. Available online at: https://www.research-collection.ethz.ch/handle/20.500.11850/101175.

Mandallaz, D., Breschan, J., Hill, A., 2013. New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-

based Monte Carlo approach with applications to small-area estimation. Can. J. For. Res. 43, 1023–1031.
McCulloch, C.E., Searle, S.R., 2001. Generalized, Linear, and Mixed Models. In: Wiley, New York, p. 247.
McRoberts, R.E., Chen, Q., Walters, B.F., 2017. Multivariate inference for forest inventory using auxiliary airborne laser scanning data. Forest Ecol. Manag. 401, 295–303.
McRoberts, R.E., Næsset, E., Heikkinen, J., Chen, Q., Strimbu, V., Esteban, J., Hou, Z., Giannetti, F., Mohammadi, J., Chirici, G., 2022. On the model-assisted regression estimators using remotely sensed auxiliary data. Remote Sens. Environ. 281, 113168.
McRoberts, R.E., Næsset, E., Gobakken, T., 2013. Inference for lidar-assisted estimation of forest growing stock volume. Remote Sens. Environ. 128, 268–275.
Mehtätalo, L., Lappi, J., 2020. Biometry for Forestry and Environmental Data with Examples in R. Chapman & Hall/CRC.
Mohamedou, C., Kangas, A., Hamedianfar, A., Vauhkonen, J., 2022. Potential of bayesian formalism for the fusion and assimilation of sequential forestry data in time and space. Can. J. For. Res. 52, 439–449.
Myllymäki, M., Gobakken, T., Næsset, E., Kangas, A., 2017. The efficiency of poststratification compared with model-assisted estimation. Can. J. For. Res. 47, 515–526.
Næsset, E., Gobakken, T., Solberg, S., Gregoire, T.G., Nelson, R., Stahl, G., Weydahl, D., 2011. Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: a case study from a boreal forest. Remote Sens. Environ. 115, 3599–3614.
NFGA, 2010. Technical regulations for inventory for forest management planning and design. China National Forestry and Grassland Administration. http://c.gb688.cn/bzgk/gb/showGb?type=online&hcno=18D1658E2885CD74F5976BEA0A3B2991.
Phiri, D., Morgenroth, J., 2017. Developments in landsat land cover classification methods: a review. Remote Sens. 9, 967.
Rizopoulos, D., 2022. Bootstrap stepAIC. https://cran.r-project.org/web/packages/bootStepAIC/index.html accessed 2022.06.13.
Robinson, G.K., 1991. That BLUP is a good thing: the estimation of random effects. Stat. Sci. 6, 15–32.
Särndal, C.E., Swensson, B., Wretman, J.H., 1992. Model Assisted Survey Sampling. Springer, New York.
Schabenberger, O., Gotway, C.A., 2005. Statistical Methods for Spatial Data Analysis, (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315275086.
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 in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. Forest Ecosyst. 3, 5.
Stehman, S.V., 2009. Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing. Remote Sens. Environ. 113, 2455–2462.
Swain, P.H., Vardman, S.B., Tilton, J.C., 1981. Contextual classification of multispectral image data. Pattern Recogn. 13, 429–441.
Theil, H., 1963. On the use of incomplete prior information in regression analysis. J. Am. Stat. Assoc. 58, 401–414.
Thompson, S.K., 2012. Sampling, third ed. Wiley, New Jersey, United States.
Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R.E., 2010. National Forest Inventories: Pathways for common reporting. Springer, Berlin.
Vågen, T.G., Winowiecki, L.A., Tamene Desta, L., Tondoh, J.E., 2015. The Land Degradation Surveillance Framework (LDSF) - Field Guide v4.1. World Agroforestry Centre, Nairobi, Kenya.
Vidal, C., Alberdi, I., Hernández, L., Redmond, J., 2016. National Forest Inventories: Assessment of Wood Availability and Use. Springer, Cham, Switzerland.
Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R.E., 2010. National Forest Inventories: Pathways for Common Reporting. Springer, Dordrecht, The Netherlands.
Williams, B.K., Brown, E.D., 2019. Sampling and analysis frameworks for inference in ecology. Method. Ecol. Evol. 10, 1832–1842.
Xu, Q., Li, B., Maltamo, M., Tokola, T., Hou, Z., 2019. Predicting tree diameter using allometry described by non-parametric locally-estimated copulas from tree dimensions derived from airborne laser scanning. For. Ecol. Manag. 434, 205–212.
Xu, Q., Li, B., McRoberts, R.E., Li, Z., Hou, Z., 2023. Harnessing data assimilation and spatial autocorrelation for forest inventory. Remote Sens. Environ. 288, 113488.
Xu, Q., Man, A., Fredrickson, M., Hou, Z., Pitkanen, J., Wing, B., Ramirez, C., Li, B., Greenberg, J.A., 2018. Quantification of uncertainty in aboveground biomass estimates derived from small-footprint airborne LiDAR. Remote Sens. Environ. 216, 514–528.
Xu, Q., Stahl, G., McRoberts, R.E., Li, B., Tokola, T., Hou, Z., 2021. Generalizing systematic adaptive cluster sampling for forest ecosystem inventory. For. Ecol. Manag. 489, 119051.
Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J. Am. Stat. Assoc. 57, 348–368.