Surveys in Geophysics (2019) 40:1001–1015
https://doi.org/10.1007/s10712-019-09507-1

Using a Finer Resolution Biomass Map to Assess the Accuracy of a Regional, Map-Based Estimate of Forest Biomass

Ronald E. McRoberts[1] · Erik Næsset[2] · Greg C. Liknes[1] · Qi Chen[3] · Brian F. Walters[1] · Sassan Saatchi[4] · Martin Herold[5]

Received: 21 July 2018 / Accepted: 8 January 2019 / Published online: 8 February 2019
© Springer Nature B.V. 2019

Abstract

National greenhouse gas inventories often use variations of the *gain–loss* approach whereby emissions are estimated as the products of estimates of areas of land-use change characterized as *activity data* and estimates of emissions per unit area characterized as *emission factors*. Although the term *emissions* is often intuitively understood to mean release of greenhouse gases from terrestrial sources to the atmosphere, in fact, emission factors can also be negative, meaning removal of the gases from the atmosphere to terrestrial sinks. For remote and inaccessible forests for which ground sampling is difficult if not impossible, emission factors may be based on map-based estimates of biomass or biomass change obtained from regional maps. For the special case of complete deforestation, the emission factor for the aboveground biomass pool is simply mean aboveground, live-tree, biomass per unit area prior to the deforestation. If biomass maps are used for these purposes, estimates must still comply with the first IPCC good practice guideline regarding accuracy relative to the true value and the second guideline regarding uncertainty. Accuracy assessment for a map-based estimate entails comparison of the estimate to a second estimate obtained using independent reference data. Assuming ground sampling is not feasible, a map of greater quality than the regional map may be considered as a source of reference data where greater quality connotes attributes such as finer resolution and/or greater accuracy. For a local, sub-regional study area in Minnesota in the USA, the accuracy of an estimate of mean aboveground, live-tree biomass per unit area (AGB, Mg/ha) obtained from a coarser resolution, regional, MODIS-based biomass map was assessed using reference data sampled from a finer resolution, local, airborne laser scanning (ALS)-based biomass map. The rationale for a local assessment of a regional map is that, although assessment of a regional map would be difficult for the entire extent of the map, it can likely be assessed for multiple local sub-regions in which case expected local regional accuracy for the entire map can perhaps be inferred. For this study, the local assessment was in the form of a test of the hypothesis that the local sub-regional estimate from the regional map did not deviate from the local true value. A hybrid approach to inference was used whereby design-based inferential techniques were used to estimate uncertainty due to sampling from the finer resolution map, and model-based inferential techniques were used to estimate uncertainty resulting from using the finer resolution map unit values which were subject to prediction error as reference data. The test revealed no statistically significant difference between the MODIS-based and ALS-based map estimates, thereby indicating that for the local sub-region, the regional, MODIS-based estimate complied with the first IPCC good practice guideline for accuracy.

Extended author information available on the last page of the article

Keywords Hybrid inference · Design-based inference · Model-based inference · Greenhouse gas inventory · IPCC good practice guidelines

1 Introduction

National greenhouse gas (GHG) inventories assess the scale of emissions for multiple land uses including the Agriculture, Forestry and other land use sectors (IPCC 2006) and are typically implemented using either the *stock-change* or *gain–loss* approach (IPCC 2006, p. 4.11; GFOI 2016, p. 22). The stock-change approach, for which emissions are estimated as differences in stocks for two dates, is well suited for countries with established forest sampling programs such as national forest inventories (NFI). For countries with remote and inaccessible forests that are difficult to sample, the gain–loss approach is more often used. With the latter approach, emissions are estimated as products of land-use change area estimates, characterized as *activity data*, and estimates of emissions per unit area, characterized as *emission factors* or *removal factors*. Although the term *emissions* is often intuitively understood to mean release of greenhouse gases from terrestrial sources to the atmosphere, in fact, emission factors can also be negative, meaning removal of the gases from the atmosphere to terrestrial sinks.

For purposes of estimating activity data, methods that address ground sampling difficulties for remote regions have been widely reported (e.g., Olofsson et al. 2013, 2014; McRoberts et al. 2018b), but such is not case for estimating emission factors. One approach that circumvents ground sampling is to use global or regional emission factors (e.g., Pearson et al. 2017), but the adverse effects on bias and precision are mostly unknown (Pelletier et al. 2012). Another approach is to convert estimates obtained from regional biomass maps to carbon estimates where the term *regional* could refer to a global map, a global map within latitude limits, a national map or simply a large area map (Saatchi et al. 2011; Baccini et al. 2012). For the special case of complete deforestation as the land-use change class, the emission factor for the aboveground biomass pool is simply mean aboveground, live-tree biomass per unit area (AGB, Mg/ha) prior to the deforestation. Regardless of the approach, the Intergovernmental Panel on Climate Change (IPCC) specifies two good practice guidelines for GHG inventories: (1) "neither over- nor underestimates so far as can be judged," and (2) "uncertainties are reduced as far as is practicable" (GFOI 2016, p. 15). For the first guideline, the standard for assessing under- or over-estimation was assumed to be the true value; for the second guideline, the presupposition was that uncertainty must first be rigorously estimated before it can be reduced.

Multiple methods can be used to assess the degree to which an estimate from a regional map satisfies the first IPCC good practice guideline regarding accuracy relative to the true value. Generally, such methods entail comparing the regional map-based estimate to an estimate based on independent reference data of at least greater quality than the regional map data. For this study, an underlying assumption was that ground sampling was not a feasible source of reference data for reasons such as the remoteness and/or inaccessibility of the forests. For the latter situation, an alternative is a method analogous to that used for acquiring reference data for estimating activity data, i.e., reference data in the form of visual interpretations of aerial or satellite imagery, albeit subject to the constraint that the interpretations are of greater quality than the map data (Stehman 2009; GFOI 2016, p. 125; Olofsson et al. 2013; McRoberts et al. 2018b). Direct extension of this approach to visual assessment of biomass from interpreted imagery would likely fail to satisfy the criterion

that the reference data must be of greater quality than the map data. A potentially more viable approach would be to use a finer resolution biomass map as a source of reference data subject to the criterion that the finer resolution map data are of greater accuracy than the regional map data. However, even if reference data from a finer resolution map are of greater quality than the regional map unit values, they are still subject to uncertainty.

The study objective was to illustrate a statistically rigorous method for testing the hypothesis that an estimate of mean AGB from a regional map for a sub-regional study area complies with the first IPCC good practice guideline, i.e., the estimate does not deviate from the true value. Although ground reference data were, in fact, available for the study area, for illustrative purposes the analyses assumed such data were not available such as could be the case for tropical forest countries lacking sufficiently extensive ground plot networks. Thus, the test took the form of a comparison of the regional map-based estimate for the sub-region to an estimate based on a sample from a finer resolution map.

2 Data

2.1 Study Area

The 7583-km^2 study area consisted of the entirety of Itasca County in north central Minnesota in the USA (Fig. 1). Land cover includes water, wetlands and approximately 80% forest consisting of uplands with deciduous mixtures of pines (*Pinus* spp.), spruce (*Picea* spp.) and balsam fir (*Abies balsamea* (L.) Mill.) and lowlands with spruce (*Picea* spp.), tamarack (*Larix laricina* (Du Roi) K. Koch), white cedar (*Thuja occidentalis* (L.)) and black ash (*Fraxinus nigra* Marsh.). Forest stands in the study area are typically naturally regenerated, uneven-aged, and mixed species.

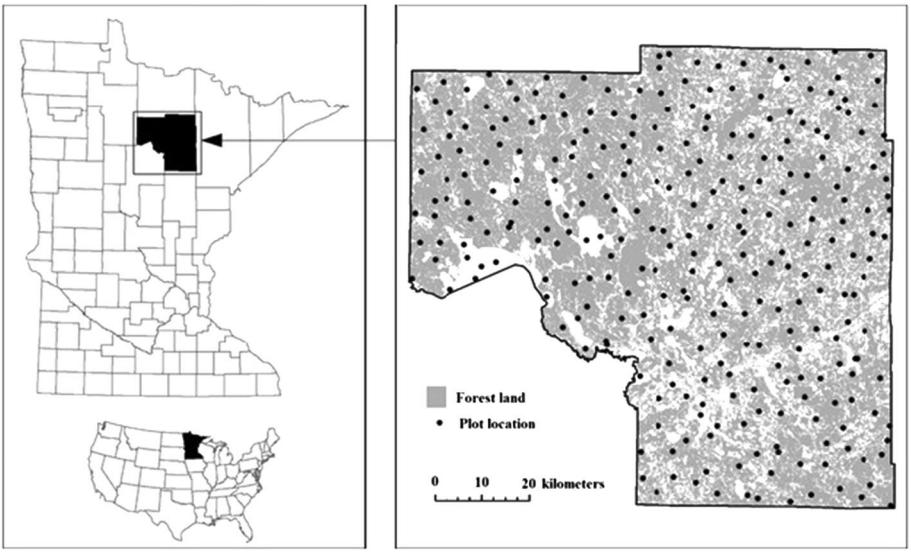

Fig. 1 Study area in Itasca County, Minnesota, USA

2.2 Maps

2.2.1 *Coarser Resolution (CR) Regional Map*

The 250-m × 250-m coarse resolution (CR) regional map used for this study was constructed by the Forest Inventory and Analysis (FIA) program of the US Forest Service which conducts the NFI of the USA (McRoberts et al. 2010). The map was based on regression trees, AGB for FIA plots measured between 1990 and 2003, and predictor variables from multiple sources including 2001 Moderate Resolution Imaging Spectrometer (MODIS) image products (Hansen et al. 2003; Huete et al. 2002; Vermote and Vermueulen 1999), the 1992 National Land Cover Dataset (Vogelmann et al. 2001), and topographic and climatic variables (Blackard et al. 2008). Validation techniques included pixel-level comparisons and comparisons of mean plot-level AGB and mean model AGB predictions for polygons of various sizes. In general, the map tended to over-predict for areas of small AGB and under-predict for areas of large AGB. Reported correlations between aggregations of FIA plot AGB values and map values ranged from 0.31 to 0.73, depending on the region of the country; for the region that included the study area the correlation was 0.46. Blackard et al. (2008) provide more details regarding the regional CR map.

2.2.2 *Finer Resolution (FR) Map*

A 13-m × 13-m, local, finer resolution (FR) AGB map was constructed using 2012 airborne laser scanning (ALS) data and FIA plot data obtained for an equal probability sample. Allometric models were used to predict AGB for the central, circular, 7.32-m (24-ft) radius subplots for 541 FIA plots measured between 2010 and 2014. For this study, uncertainty in the allometric model predictions was considered negligible relative to the effects of sampling variability (McRoberts et al. 2016). The forest inventory data are described in greater detail in McRoberts et al. (2018a). Wall-to-wall airborne laser scanning (ALS) data were acquired in April 2012 with a nominal pulse density of 0.67 pulses/m^2. Distributions of all first return heights were constructed for the 168.3-m^2 circular plots and the 169-m^2 square cells that tessellated the study area and served as FR map units. Standard ALS metrics included the mean, quadratic mean, standard deviation, skewness, kurtosis of the ALS height distributions and deciles of the height and canopy density distributions. The ALS data are described in greater detail in McRoberts et al. (2018a).

A model of the relationship between plot-level AGB as the dependent variable and the ALS metrics as independent variables was formulated as

$$\begin{aligned} \mathrm{AGB}_i &= f\left(\boldsymbol{x}_i;\boldsymbol{\beta}\right) + \varepsilon_i \\ &= \beta_0 \cdot x_{1i}^{\beta_1} \cdot \exp\left(\beta_2 \cdot x_{2i} + \cdots + \beta_{\mathrm{p}} \cdot x_{\mathrm{p}i}\right) + \varepsilon_i \end{aligned} \tag{1}$$

where i indexes plots, x_{ji} is an ALS metric, $\beta_j \in \boldsymbol{\beta}$ is a parameter to be estimated and ε_i is a residual assumed to be distributed $N(0,\boldsymbol{\Sigma})$. The model was fit to the data using weighted nonlinear least squares where $\hat{\boldsymbol{\Sigma}}^{-1}$ was used to weight the observations. A forward selection procedure was used to select independent variables for inclusion in the model if they statistically significantly increased the quality of fit of the model to the data at a nominal $\alpha = 0.05$ level. The resulting model was used to predict AGB for all 13-m × 13-m cells in the study area, thereby producing the FR map. For illustrative purposes for this study, the FR map was henceforth considered to be the only available source of reference data for

assessing the accuracy of the regional, CR map-based estimate of mean AGB. Although the local FR map was constructed using ground plot data acquired from within the study area, this is not a necessary condition (McRoberts et al. 2014).

Heteroscedastic (nonconstant) residual variance as required for the weight matrix, $\widehat{\mathbf{\Sigma}}^{-1}$, was estimated using a 4-step procedure (McRoberts et al. 2016): (1) for $\hat{\varepsilon}_i = y_i - \hat{y}_i$, the sample pairs $\left(\hat{y}_i, \hat{\varepsilon}_i\right)$ were ordered with respect to $\hat{y}_i$, (2) the pairs were aggregated into groups of size 15; (3) within each group, g, the mean of the predictions, $\bar{\hat{y}}_g$, and the standard deviation, $\hat{\sigma}_{\mathrm{g}}$, of the estimated residuals, $\hat{\varepsilon}_i$, were calculated; and (4) the relationship between the group standard deviations and the group prediction means was represented using a model of the form

$$\hat{\sigma}_{\mathrm{g}} = \gamma \cdot \bar{\hat{y}}_{\mathrm{g}}, \tag{2}$$

where γ is a parameter to be estimated. To accommodate heteroscedasticity when fitting the model, the *i*th diagonal element of $\widehat{\mathbf{\Sigma}}$ was $\hat{\sigma}_i^2 = \left(\hat{\gamma} \cdot \hat{y}_i\right)^2$ from Eq. (2) where $\hat{y}_i$ is the model prediction for the *i*th plot. Because distances between plot locations mostly exceeded the range of spatial correlation, the off-diagonal elements of $\widehat{\mathbf{\Sigma}}$ were set to 0.

3 Methods

3.1 The Test

Three assumptions underlay the study: (1) the only information available for the local portion of the regional CR map was the map unit values, (2) ground sampling as a source of reference data for the study area was not feasible, and (3) the FR map included sufficient meta-data to estimate mean AGB and the uncertainty of the estimate.

The objective was to test the hypothesis that the regional CR map-based estimate of mean AGB for the local sub-region complies with the first IPCC good practice guideline. In particular, the test focused on the accuracy of the CR map-based estimate where *accuracy* is related to *bias* and refers to agreement between the true value and the average of repeated independent estimates (IPCC, 2006, p. 3.7). Because the true value is not known, the best alternative is an estimate obtained using reference data that are at least of greater quality than the CR map data (Stehman 2009; GFOI 2016, p. 125). Further, because typically only a single sample of model calibration data is available, the test result can only be expressed probabilistically, i.e., as an inference. An intuitive test statistic is

$$t = \frac{\hat{\mu}^{\mathrm{CR}} - \hat{\mu}^{\mathrm{FR}}}{\sqrt{\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{CR}}\right) + \widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{FR}}\right)}}. \tag{3}$$

where the uncertainties associated with the CR and FR map-based estimates of the mean are expressed in terms of *mean square error* (*MSE*) rather than *variance* because the map-based estimators of the means are not necessarily unbiased as a result of possible systematic map error (Cochran 1977, p. 15). The exact forms of the two $\widehat{\mathrm{MSE}}$s are described in the sections that follow. Under the assumption that the data used to construct the FR map are independent of the data used to construct the CR map, no covariance term in the denominator of Eq. (3) is necessary.

A crucial issue is that reference data acquired from any map, regardless of map accuracy, cannot be assumed to be without error. Mowrer and Congalton (2000) characterize

reference data subject to non-negligible error as *imperfect reference data*. In the event of imperfect reference data, compliance with the IPCC good practice guidelines requires incorporation of the effects of this source of uncertainty into $\widehat{\mathrm{MSE}}$s. In particular, the $\widehat{\mathrm{MSE}}$ must incorporate uncertainty due to sampling from the map used as the source of reference data and uncertainty due to the imperfect nature of the sample reference data, i.e., the FR map unit values are subject to error.

The term *hybrid inference* characterizes recently developed methods that combine design-based inferential techniques for assessing the effects of sampling variability and model-based inferential techniques for assessing the effects of imperfect reference data (Fattorini 2012; Corona et al. 2014; McRoberts et al. 2016; Ståhl et al. 2016). Hybrid inference has four key features: (1) A probability sample of population units for which only auxiliary information is available; (2) a prediction technique that uses the auxiliary information to predict the reference data values for the sample units; (3) a design-based estimator of the population parameter using the reference data predictions for the sample units; and (4) estimation of uncertainty using a design-based estimator to accommodate the effects of sampling variability and a model-based estimator to accommodate the effects of uncertainty in the reference data.

For this study, hybrid inference entailed sampling from the FR map, using design-based estimators with the FR sample map reference data to estimate the mean and the component of $\widehat{\mathrm{MSE}}$ due to sampling variability (Sect. 3.2), and using model-based estimators to estimate the component of $\widehat{\mathrm{MSE}}$ due to the effects of the imperfect FR map reference data (Sect. 3.3). The hybrid $\widehat{\mathrm{MSE}}$ is then expressed additively as,

$$\widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right) + \widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right), \tag{4}$$

where the design-based (DB) component, $\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$, estimates the effects of sampling from the FR map and the model-based (MB) component, $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ estimates the effects of uncertainty in the imperfect FR map reference data. $\widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ then replaces $\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ in Eq. (3). The technical objectives in Sects. 3.2 and 3.3 were to address $\hat{\mu}^{\mathrm{FR}}$, $\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ and $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ with $\hat{\mu}^{\mathrm{CR}}$ and $\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ addressed in Sect. 3.4.

3.2 Design-Based Component of Hybrid Inference

Three primary assumptions underlie *design-based inference*, also characterized as *probability-based inference* (Hansen et al. 1983). First, the basis for validity is a probability sample that incorporates some form of randomization. Second, each population unit is assumed to have one and only one possible value, apart from negligible observation or measurement error. Third, the probability of selection for each population unit into the sample is positive and known. Much of the effort for design-based inference involves selecting an appropriate combination of sampling design and corresponding estimator. Familiar sampling designs include simple random, systematic, stratified, multi-phase and multi-stage sampling designs. Estimators corresponding to these designs are generally unbiased or at least approximately unbiased, and uncertainty estimation typically entails comparing observations to their corresponding means or model predictions. All design-based estimators assume that observations or measurements of the response variable have at most negligible uncertainty (Snedecor and Cochran 1967, p. 164).

A relatively simple sampling design used for the current study entailed selecting the central FR map unit within each CR map unit. Cochran (1997, Sect. 8.13) characterizes this design as a *systematic aligned* sampling design. The FR map-based estimate of the mean and the corresponding $\widehat{\mathrm{MSE}}$ were calculated using the simple *expansion* (*Exp*) estimators (Royall and Herson 1973),

$$\hat{\mu}^{\mathrm{FR}} = \frac{1}{N_{\mathrm{CR}}} \sum_{i=1}^{N_{\mathrm{CR}}} \hat{y}_i^{\mathrm{FR}}, \tag{5a}$$

and

$$\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \frac{1}{N_{\mathrm{CR}} \cdot \left(N_{\mathrm{CR}} - 1\right)} \sum_{i=1}^{N_{\mathrm{CR}}} \left(\hat{y}_i^{\mathrm{FR}} - \hat{\mu}^{\mathrm{FR}}\right)^2, \tag{5b}$$

where i indexes the CR map units, $\hat{y}_i^{\mathrm{FR}}$ is the value for the central FR map unit within the ith CR map unit and N_{CR} is the FR sample size. Because the FR map units are considerably smaller than the CR map units, the total number of CR map units, N_{CR}, is much larger than the total number of FR map units; thus, N_{CR} and N_{FR} are not interchangeable. In particular, with the systematic aligned sampling design with one FR sample unit for each CR sample unit, the FR sample size is exactly equal to the total number of CR map units, N_{CR}. Although when used with systematic sample data, $\widehat{\mathrm{MSE}}$ for the Exp estimator may be biased, it tends to be conservatively biased in the sense that $\widehat{\mathrm{MSE}}$s may be slightly too large (Särndal et al. 1992, p. 83). Other sampling designs were considered, but preliminary analyses indicated that the resulting $\widehat{\mathrm{MSE}}$s were larger than for the systematic design. In addition, post-stratified (McRoberts et al. 2013) and model-assisted difference estimators (Ståhl et al. 2016) were also investigated for use with the systematic sample data, but preliminary analyses again indicated there was little to be gained relative to the Exp estimators.

Overall, the advantages of the Exp estimators are that they are simple, intuitive and easy to implement. A disadvantage is that variances are frequently large, particularly for small sample sizes and/or populations with large variability among population unit values. Because $\hat{\mu}^{\mathrm{FR}}$ is an independent estimate of the sub-regional population mean, it can serve as the independent estimate required for comparison to $\hat{\mu}^{\mathrm{CR}}$ using Eq. (3).

3.3 Model-Based Component of Hybrid Inference

Assumptions underlying *model-based inference,* also characterized as *model-dependent inference*, differ considerably from the more familiar design-based inference. First, the basis for validity is correct specification of the model, not a probability sample. Second, model-based inference assumes an entire distribution of possible values for each population unit, not just a single value. Third, randomization occurs via realization of observations from the distributions characterizing the population units selected for the sample, not via the sampling design.

Despite being considered observations without error for the design-based component of hybrid inference (Sect. 3.2), the FR map unit values are model predictions (Sect. 3.2), and the uncertainty of these predictions must be incorporated into the final hybrid $\widehat{\mathrm{MSE}}$. For regression models, the form of the model-based component of $\widehat{\mathrm{MSE}}$ can be formulated as

$$\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \frac{1}{N_{\mathrm{CR}}^2}\sum_{i_1=1}^{N_{\mathrm{CR}}}\sum_{i_2=1}^{N_{\mathrm{CR}}} z'_{i_1}\cdot\widehat{\mathbf{V}}_{\hat{\boldsymbol{\beta}}}\cdot z_{i_2} + \frac{1}{N_{\mathrm{CR}}^2}\sum_{i=1}^{N_{\mathrm{CR}}}\hat{\sigma}_i^2 + \frac{1}{N_{\mathrm{CR}}^2}\sum_{i_1\neq}^{N_{\mathrm{CR}}}\sum_{i_2}^{N_{\mathrm{CR}}}\hat{\sigma}_{i_1}\cdot\hat{\sigma}_{i_2}\cdot\hat{\rho}_{i_1 i_2} \tag{6a}$$

where N_{CR} is again the number of CR map units,

$$z_{ij} = \frac{\partial f\left(\boldsymbol{x}_i;\hat{\boldsymbol{\beta}}\right)}{\partial \beta_j}, \tag{6b}$$

$$\widehat{\mathbf{V}}_{\hat{\mathbf{B}}} = \left(\mathbf{Z}'\cdot\widehat{\boldsymbol{\Sigma}}^{-1}\cdot\mathbf{Z}\right)^{-1} \tag{6c}$$

is the first-order Taylor's series estimator of the covariance matrix for the model parameter estimates, $\widehat{\boldsymbol{\Sigma}}$ is the covariance matrix for the model prediction residuals, $\widehat{\varepsilon}_i$, $\widehat{\sigma}_i^2$ is the estimate of the heteroscedastic residual variance for the FR map from Sect. 2.2.2 and $\widehat{\rho}_{i_1 i_2}$ is the estimate of spatial correlation for the FR map residuals.

The first term of Eq. (6a) is further denoted

$$\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \frac{1}{N_{\mathrm{CR}}^2}\sum_{i_1=1}^{N_{\mathrm{CR}}}\sum_{i_2=1}^{N_{\mathrm{CR}}} z'_{i_1}\cdot\widehat{V}_{\hat{\boldsymbol{\beta}}}\cdot z_{i_2}, \tag{6d}$$

and estimates the effect on model predictions (*pre*) of variability in the sample data used to construct the model, i.e., each model calibration sample would produce different model parameter estimates and, therefore, different model predictions. The computational intensity associated with the double sums in Eq. (6d) can be reduced, using the alternate form,

$$\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \sum_{j_1=1}^{J}\sum_{j_2=1}^{J}\bar{z}_{j_1}\cdot v_{j_1 j_2}\cdot\bar{z}_{j_2}, \tag{6e}$$

where J is the number of model parameters, j_1 and j_2 index the parameters, $v_{j_1 j_2}\in\widehat{\mathbf{V}}_{\hat{\boldsymbol{\beta}}}$, $\overline{z}_j = \frac{1}{N_{\mathrm{CR}}}\sum_{i=1}^{N_{\mathrm{CR}}} z_{ij}$ and z_{ij} is defined by Eq. (6b) (Saarela et al. 2015; McRoberts et al. 2018c). For nonparametric prediction techniques such as Random Forests, Monte Carlo bootstrap procedures can be used to estimate $\widehat{\mathrm{MSE}}_{\mathrm{Pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ (Efron and Tibshirani 1994, pp. 47–48, 113; McRoberts et al. 2018c).

The second term of Eq. (6a) is further denoted

$$\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \frac{1}{N_{\mathrm{CR}}^2}\sum_{i=1}^{N_{\mathrm{CR}}}\widehat{\sigma}_i^2, \tag{6f}$$

and is the result of residual (*res*) variability of observations around their model predictions where $\hat{\sigma}_i^2$ is estimated as described in Sect. 2.2 The third term of Eq. (6a) is denoted

$$\widehat{\mathrm{MSE}}_{\mathrm{spa}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) = \frac{1}{N_{\mathrm{CR}}^2}\sum_{i_1\neq}^{N_{\mathrm{CR}}}\sum_{i_2}^{N_{\mathrm{CR}}}\hat{\sigma}_{i_1}\cdot\hat{\sigma}_{i_2}\cdot\hat{\rho}_{i_1 i_2} \tag{6g}$$

and is the result of spatial (*spa*) correlation among residuals. Spatial correlation, ρ, is often estimated via a correlogram using the model prediction residuals obtained when fitting the

model. However, when plot separation distances exceed the range of spatial correlation which is usually the case for efficient sampling designs, construction of the correlogram for small distances and estimation of the range are not possible. For this study, an estimate of the range of spatial correlation was selected as 200 m, the maximum of values reported in the literature for similar studies (Breidenbach et al. 2008, 2016; McRoberts et al. 2007; Mauro et al. 2017). An exponential correlogram of the form $\rho = \exp(\lambda \cdot d)$ was assumed for which d is distance and $\lambda = \frac{\ln(0.05)}{\nu}$ where ν is the range of spatial correlation defined to be the distance for which $\rho = 0.05$.

3.4 Implementing the Statistical Test

In addition to $\hat{\mu}^{\mathrm{FR}}$ and $\widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right)$, the test statistic of Eq. (3) requires $\hat{\mu}^{\mathrm{CR}}$ and $\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{CR}}\right)$. The CR map-based estimate, $\hat{\mu}^{\mathrm{CR}}$, is easily calculated using the *synthetic* estimator as

$$\hat{\mu}^{\mathrm{CR}} = \frac{1}{N_{\mathrm{CR}}} \sum_{i=1}^{N_{\mathrm{CR}}} \hat{y}_i^{\mathrm{CR}}, \tag{7a}$$

where i indexes CR map units, N_{CR} is the number of CR map units and $\hat{y}_i^{\mathrm{CR}}$ is the value for the ith CR map unit. $\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ is obtained using Eq. (6a), albeit with local meta-data provided by the regional map authors, and replaces $\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ in Eq. (3).

The test corresponding to the first IPCC guideline was that the CR map-based estimate, $\hat{\mu}^{\mathrm{CR}}$, was neither an over- nor under-estimate relative to the true value and took the form of a comparison of $\hat{\mu}^{\mathrm{CR}}$ to the independent FR map-based estimate, $\hat{\mu}^{\mathrm{FR}}$. The exact form of the test depends on how the CR map is considered. If the CR map values and their mean are considered constants without inherent uncertainty apart from being correct or incorrect, then $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right) = 0$ in which case Eq. (3) reduces to

$$t = \frac{\hat{\mu}^{\mathrm{CR}} - \hat{\mu}^{\mathrm{FR}}}{\sqrt{\widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right)}}. \tag{7b}$$

If the CR map values are considered predictions with inherent uncertainty, then $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right) \neq 0$ in which case the test statistic expressed by Eq. (3) must be used.

Because map authors have access to the original data used to construct the map, they can readily calculate $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ for the entire regional map or any local sub-regional portion of it. However, for map users to calculate $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ for only a local sub-region, the regional map authors must provide the meta-data specific to the sub-region of the CR map of interest. For a CR map consisting of N_{CR} units, the number of covariance values necessary to calculate $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ is on the order of N_{CR}^2. Expecting map authors to provide this many values in an easily accessible manner is impractical. A solution would be a method for expressing the local meta-data required for calculating $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ in a summarized form or as functions of the map values. Although such summaries or functions have been proposed for $\widehat{\mathrm{MSE}}_{\mathrm{res}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ and $\widehat{\mathrm{MSE}}_{\mathrm{spa}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ (McRoberts et al. 2018c), such is

not the case for $\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$. Until such summaries or functions are developed, accurate calculation of $\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ for a sub-region of a regional CR map is not possible.

In lieu of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$, two alternatives may be considered. First, the test statistic can be calculated using $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right) = 0$ which yields the same test statistic as Eq. (7a). Second, if the FR map is of greater quality than the CR map, then it is reasonable to assume that $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) \leq \widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$. If so, then substituting $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ for $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ in Eq. (3) yields

$$t = \frac{\hat{\mu}^{\mathrm{CR}} - \hat{\mu}^{\mathrm{FR}}}{\sqrt{\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right) + \widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right)}}. \tag{7c}$$

If $\hat{\mu}^{\mathrm{CR}}$ is not statistically significantly different from $\hat{\mu}^{\mathrm{FR}}$ using either Eqs. (7b) or (7c) to calculate the test statistic, then it will also likely not be statistically significantly different when using Eq. (3), because use of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ rather than $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ in the denominator would presumably produce an even smaller and less statistically significant |t|. However, if $\hat{\mu}^{\mathrm{CR}}$ is statistically significantly different from $\hat{\mu}^{\mathrm{FR}}$ using either Eqs. (7b) or (7c), then the statistical significance of the difference between $\hat{\mu}^{\mathrm{CR}}$ and $\hat{\mu}^{\mathrm{FR}}$ is undetermined because the degree to which use of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ rather than $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ alters $|t|$ and therefore alters the significance is unknown.

4 Results and Discussion

4.1 The Fine Resolution (FR) Map

The model on which the FR map was based was fit to the FIA plot AGB data using weighted nonlinear least squares techniques where the weighting matrix was $\widehat{\boldsymbol{\Sigma}^{-1}}$ (Sect. 2.2). Following selection of mean ALS height as the metric for the first power component of the model, no other metric statistically significantly improved the fit. Quality of fit of the model to the data was assessed using unweighted pseudo-$R^2 = 0.72$ and mean weighted residual of 0.004.

4.2 Estimates

The CR map consisted of 121,236 map units. The synthetic estimate of mean AGB for the CR map was $\hat{\mu}^{\mathrm{CR}} = 53.77$ Mg/ha. The hybrid estimate of the mean using reference data sampled from the FR map was $\hat{\mu}^{\mathrm{FR}} = 51.49$ Mg/ha with standard error, $\mathrm{SE}\left(\hat{\mu}^{\mathrm{FR}}\right) = \sqrt{\widehat{\mathrm{MSE}}^{\mathrm{Hyb}}\left(\hat{\mu}^{\mathrm{FR}}\right)} = 1.34$ Mg/ha. The individual $\widehat{\mathrm{MSE}}$ component contributions to $\mathrm{SE}\left(\hat{\mu}^{\mathrm{FR}}\right)$ expressed as square roots of the respective $\widehat{\mathrm{MSE}}$s were $\sqrt{\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)} = 0.15$ Mg/ha, $\sqrt{\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)} = 1.33$ Mg/ha, $\sqrt{\left(\widehat{\mathrm{MSE}}_{\mathrm{res}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)\right)} = 0.09$ Mg/ha, $\sqrt{\widehat{\mathrm{MSE}}_{\mathrm{spa}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)} < 0.01$ Mg/ha

with $\sqrt{\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)} = \sqrt{\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{pre}}\left(\hat{\mu}^{\mathrm{FR}}\right) + \widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{res}}\left(\hat{\mu}^{\mathrm{FR}}\right) + \widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{spa}}\left(\hat{\mu}^{\mathrm{FR}}\right)} = 1.33$ Mg/ha. Using the Eq. (7b) form of the test statistic, $t = 1.70$, and using the Eq. (7c) form, $t = 1.21$. Assuming a t-distribution for the test statistic and that $|t| < 2.0$ indicates a nonsignificant difference, the hypothesis that $\hat{\mu}^{\mathrm{CR}}$ complies with the first IPCC good practice guideline for the sub-region cannot be rejected, regardless of which form of the test statistic was used and regardless of whether the CR map units were considered to have or not have inherent uncertainty.

Among all $\widehat{\mathrm{MSE}}$ components, $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{pre}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ was dominant. One consequence for this study was that although post-stratified and model-assisted regression estimators would tend to reduce the $\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ component, neither estimator would have had more than a negligible effect on $\mathrm{SE}\left(\hat{\mu}^{\mathrm{FR}}\right)$. A second consequence was that the uncertainty in the FR map reference data, expressed by $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ and dominated by $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{pre}}\left(\hat{\mu}^{\mathrm{FR}}\right)$, was a greater contributor to the uncertainty of $\hat{\mu}^{\mathrm{FR}}$ than was $\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$. The small relative value of $\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ can be attributed to the very large FR map sample size of $N_{\mathrm{CR}} = 121{,}236$. Small relative values of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{res}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ have been reported previously (Ståhl et al. 2016; McRoberts et al. 2018c), as have negligible values of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{spa}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ (McRoberts et al. 2018c). In particular, $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{spa}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ will always be small when the proportion of pairs of sample units separated by distances less than the range of spatial correlation is small relative to the total number of pairs of sample units.

Caution should be exercised when extrapolating these results to situations for which the map accuracies and relative sizes of the CR and FR map units may be substantially different than for this study. Nevertheless, for study areas with very large numbers of CR map units, the systematic aligned sampling design should work well. In addition, $\mathrm{SE}\left(\hat{\mu}^{\mathrm{FR}}\right)$ will likely be dominated by $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{pre}}\left(\hat{\mu}^{\mathrm{FR}}\right)$.

4.3 Additional Considerations

Consideration should be given to the effects of different sources of remotely sensed auxiliary information used to construct the CR and FR maps. For example, for this study, the underlying ALS data used to construct the FR map were filtered to remove returns from man-made objects such as buildings, water towers and utility poles. In addition, because water absorbs ALS pulses, no ALS returns were available for lakes. For the Itasca study area, water area is substantial at approximately 9% of the total county area. Missing ALS metrics for water and man-made objects were set to 0.

As previously noted, reference data acquired from a map, regardless of map accuracy, are subject to error. The hybrid $\widehat{\mathrm{MSE}}$ accommodates the effects of this reference data error via $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$. However, there is no accommodation for the effects of map error on $\hat{\mu}^{\mathrm{FR}}$. Thus, regardless of quality of the FR map, design-based estimators such as Eq. (5a) cannot be assumed to be unbiased when used with reference data subject to error. The effects of this bias on significance levels for tests of hypothesis are generally unknown but likely small.

An assumption underlying the analyses is the availability of meta-data that can be used to calculate at least $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{pre}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ as the dominant component of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ for the

sub-regional portion of the regional CR map. For an FR map constructed using a regression model, $\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ can be readily calculated using the covariance matrix for the model parameter estimates and the map unit values for the model predictor variables. For nonparametric prediction techniques such as Random Forests, Monte Carlo bootstrap techniques can be considered (Efron and Tibshirani 1994, pp. 47–48, 113; McRoberts et al. 2018c). However, the exact form of the resampling for these techniques must mimic the sampling design used to acquire the data used to construct the FR map.

As noted, accurate calculation of $\widehat{\mathrm{MSE}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{CR}}\right)$ requires the meta-data necessary to calculate at least $\widehat{\mathrm{MSE}}_{\mathrm{pre}}^{\mathrm{MB}}\left(\hat{\mu}^{\mathrm{FR}}\right)$ for the local sub-region portion of the CR map. Although the map authors could provide the meta-data, the number of required data elements makes this an impractical option. Alternatively, the map authors could provide the meta-data in a summarized form, but methods for doing so in a simple and straightforward manner are currently not available.

Supplementary clarification on the IPCC good practice guidelines would be beneficial. For example, the first guideline refers to neither over- nor under-estimation, but what is the standard for comparison? For this study, the standard was assumed to be the true value, but true values are seldom known. As a second example, how should uncertainty as expressed in the second guideline be interpreted? Should it be in terms of the deviation between an estimate and the true value, or should it relate to a concept such as precision? As a third example, should the regional map be considered a stand-alone constant product whose map unit values have no inherent uncertainty apart from being correct or incorrect, or should the unit values be considered the mean of an entire distribution of possible values? Resolution of the three issues expressed by these examples is important, if not also crucial, but is beyond the scope of this study.

Finally, accuracy assessment for the entirety of a regional CR biomass map is likely not feasible, if for no other reason than that reference data that span the entire CR map are likely neither available nor feasible to acquire. The best that can be achieved is likely independent local accuracy assessments for each of a set of individual sub-regional sites for which reference data are available or can be acquired (Duncanson et al., in review). Ideally, these sites would be systematically distributed across the CR map such as via the global hexagon-based approach proposed by White et al. (1992). Regardless of how the sub-regional sites are distributed, methods for aggregating their individual accuracy assessments based on potentially quite different site-specific reference data into a comprehensive inference for the entire CR map require attention.

5 Conclusions

Four conclusions were drawn from the study, all based on the assumption that ground reference data cannot be acquired. First, when reference data for assessing the accuracy of a map-based estimate are imperfect, such as when they are obtained from a second map, even if it is of greater quality, hybrid inferential techniques must be used to accommodate both sampling variability and non-negligible errors in the reference data. Second, for this study, the spatially aligned systematic sample consisting of the central fine resolution map unit in each coarse resolution map unit worked well. The design was easy to implement, the expansion estimators were easy to apply, and the $\widehat{\mathrm{MSE}}$

component was dominated by a single component, the model-based prediction component. Third, for this study, the regional, coarse resolution map-based estimate of mean biomass per unit area complied with the first IPCC good practice guideline for the local, Itasca County sub-region. Fourth, future research priorities include developing methods for estimating the model-based component of MSE for only a local portion of a regional map and developing methods for extending inferences from individual sub-regions to the entirety of the regional map.

The methods developed and illustrated for this study are important for tropical forest countries whose ground plot networks are so limited as to preclude plot-based, IPCC compliant estimation, particularly with respect to uncertainty. In addition, they are important for biomass map authors who wish to convey to users general information regarding accuracies of map-based estimates that can be expected for local areas of their regional maps.

References

Baccini A, Goetz SJ, Walker WS, Laporte NT, Sun M, Sulla-Menashe D, Hackler J, Beck PSA, Dubayah R, Freidl MA, Samanta S, Houghton RA (2012) Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat Clim Change 2:182–185

Blackard JA, Finco MV, Helmer EH, Holden GR, Hoppus ML, Jacobs DM, Lister AJ, Moisen GG, Nelson MD, Riemann R, Ruefenacht B, Salajanu D, Weyermann DL, Winterberger KC, Brandeis TJ, Czaplewski R, McRoberts RE, Patterson PL, Tymcio RP (2008) Mapping U.S. forest biomass using national forest inventory data and moderate resolution information. Remote Sens Environ 112:1658–1677

Breidenbach J, Kublin E, McGaughey R, Andersen H-E, Reutebuch S (2008) Mixed-effects models for estimating stand volume by means of small footprint airborne laser scanner data. Photogramm J Finl 21(1):4–15

Breidenbach J, McRoberts RE, Astrup R (2016) Empirical coverage of model-based variance estimators for remote sensing assisted estimation of stand-level timber volume. Remote Sens Environ 173:274–281

Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York, p 428

Corona P, Fattorini L, Franceschi S, Scrinz 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 For Res 44:1303–1311

Duncanson L, Armston J, Disney M, Avitabile V, Barbier N, Calders K, Carte, S, Chave J, Herold M, Crowther T, Falkowski M, Kellner J, Labrière N, Lucas R, MacBean N, McRoberts RE, Meye, V Næsset E, Nickeson JE, Paul KI, Phillips O., Réjou- Méchain M, Román M, Roxburgh S, Saatchi S, Schepashenko D, Scipal K, Siqueira PR, Williams M, Whitehurst A. In review. The importance of global land product validation: towards a standardized protocol for aboveground biomass. Surveys in Geophysics. This issue

Efron B, Tibshirani R (1994) An introduction to the bootstrap. Chapman and Hall/CRC, Boca Raton

Fattorini L (2012) Design-based or model-based inference? The role of hybrid approaches in environmental surveys. In: Fattorini L (ed) Studies in Honor of Claudio Scala, Department of Economics and Statistics. University of Siena, Siena, Italy, pp 173–214

GFOI (2016) Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: methods and guidance from the global forest observations initiative, 2nd edn. Food and agriculture organization, Rome 224 p.https://www.reddcompass.org/download-the-mgdAccessed July 2017

Hansen MH, Madow WG, Tepping BJ (1983) An evaluation of model-dependent and probability-sampling inferences in sample surveys. J Am Stat Assoc 78:776–793

Hansen M, DeFries R, Townshend JR, Carroll M, Dimiceli C, Sohlberg R (2003) 500 m MODIS vegetation continuous fields: tree cover. GLCF, University of Maryland, College Park

Huete A, Didan K, Miura T, Rodriguez E, Gao X, Ferreira L (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens Environ 83:195–213

IPCC (2006) 2006 IPCC guidelines for national greenhouse gas inventories, volume 4: agriculture, forestry and other land use. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K (eds). Published: Institute

for Global Environmental Strategies, Japan. http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html. Accessed February 2018
Mauro F, Monleon VJ, Temesgen H, Ruiz LA (2017) Analysis of spatial correlation in predictive models of forest variables that use LiDAR auxiliary information. Can J For Res 47:788–799
McRoberts RE, Tomppo EO, Finley AO, Heikkinen J (2007) Estimating areal means and variances of forest attributes using the k-nearest neighbors technique and satellite imagery. Remote Sens Environ 111:466–480
McRoberts RE, Hansen MH, Smith WB (2010) United States of America. In: Tomppo E, Gschwantner T, Lawrence M, McRoberts RE (eds) National forest inventories, pathways for common reporting. Springer, Berlin 610 p
McRoberts RE, Næsset E, Gobakken T (2014) Estimation for inaccessible and non-sampled forest areas using model-based inference and remotely sensed auxiliary information. Remote Sens Environ 154:226–233
McRoberts RE, Chen Q, Domke GM, Ståhl G, Saarela S, Westfall JA (2016) Hybrid estimators for mean aboveground carbon per unit area. For Ecol Manag 378:44–56
McRoberts RE, Chen Q, Gormanson DD, Walters BF (2018a) The shelf-life of airborne laser scanning data for enhancing forest inventory inferences. Remote Sens Environ 206:254–259
McRoberts RE, Stehman SV, Liknes GC, Næsset E, Sannier C, Walters BF (2018b) The effects of imperfect reference data on remote sensing-assisted estimators of land cover class proportions. ISPRS J Photogramm Remote Sens 142:292–300
McRoberts RE, Næsset E, Gobakken T, Chirici G, Condés S, Hou Z, Saarela S, Chen Q, Ståhl G, Walters BF (2018c) Assessing components of the model-based mean square error estimator for remote sensing-assisted forest applications. Can J For Res 48:642–649
Mowrer HT, Congalton RG (eds) (2000) Quantifying spatial uncertainty in natural resources: theory and applications for GIS and remote sensing. Sleeping Bear Press, Ann Arbor
Olofsson P, Foody GM, Stehman SV, Woodcock CE (2013) Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sens Environ 129:122–131
Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA (2014) Good practices for estimating area and assessing accuracy of land change. Remote Sens Environ 148:42–57
Pearson TRH, Brown S, Murra L, Sidman G (2017) Greenhouse gas emissions from tropical forest degradation: an underestimated source. Carb Balance Manag 12:3
Pelletier J, Kirby KR, Potvin C (2012) Significance of carbon stock uncertainties on emission reductions from deforestation and forest degradation in developing countries. For Policy Econ 24:3–11
Royall RM, Herson J (1973) Robust estimation in finite populations II. J Am Stat Assoc 68(344):890−893
Saarela S, Schnell S, Grafström A, Tutominen S, Nordkvist K, Hyppä J, Kangas A, Ståhl G (2015) Effects of sample size and model form on the accuracy of model-based estimators of growing stock volume. Can J For Res 45:1524–1534
Saatchi SS, Harris NL, Brown S, Lefsky M, Mitchard ETA, Salas W, Zutta BR, Buermann W, Lewis SL, Hagen S, Petrova S, White L, Silman M, Morel A (2011) Benchmark map of forest carbon stocks in tropical regions across three continents. Proc Natl Acad Sci 108:9899–9904
Särndal C-E, Swensson B, Wretman J (1992) Model assisted survey sampling. Springer, New York, p 694
Snedecor GW, Cochran WG (1967) Statistical methods, 6th edn. The Iowa State University Press, Lowa
Ståhl G, Saarela S, Schnell S, Holm S, Breidenbach J, Healey SP, Patterson PL, Magnussen S, Næsset E, McRoberts RE, Gregoire TG (2016) Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. For Ecosyst 3:5
Stehman SV (2009) Sampling designs for accuracy assessment of land cover. Int J Remote Sens 30(20):5243–5272
Vermote EF, Vermueulen A (1999) Atmospheric correction algorithm: spectral reflectances (MOD09). University of Maryland, College Park
Vogelmann JE, Howard S, Yang L, Larson C, Wylie B, Van Driel N (2001) Completion of the 1990s national land cover data set for conterminous United States from landsat thematic mapper data and ancillary data sources. Photogramm Eng Remote Sens 67:650–661
White D, Kimerling AJ, Overton WS (1992) Cartographic and geometric components of a global sampling design for environmental monitoring. Cartogr Geogr Inf Syst 19(1):5–22

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Ronald E. McRoberts[1] · Erik Næsset[2] · Greg C. Liknes[1] · Qi Chen[3] · Brian F. Walters[1] · Sassan Saatchi[4] · Martin Herold[5]

✉ Ronald E. McRoberts
mcrob001@umn.edu

[1] Northern Research Station, U.S. Forest Service, Saint Paul, MN, USA

[2] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

[3] Department of Geography and Environment, University of Hawai‘i at Mānoa, Honolulu, HI, USA

[4] Joint Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

[5] Laboratory for Geoinformation Science and Remote Sensing, Wageningen University and Research, Wageningen, The Netherlands