biometrics

Quantifying the Model-Related Variability of Biomass Stock and Change Estimates in the Norwegian National Forest Inventory

Johannes Breidenbach, Clara Antón-Fernández, Hans Petersson, Ronald E. McRoberts, and Rasmus Astrup

National Forest Inventories (NFIs) provide estimates of forest parameters for national and regional scales. Many key variables of interest, such as biomass and timber volume, cannot be measured directly in the field. Instead, models are used to predict those variables from measurements of other field variables. Therefore, the uncertainty or variability of NFI estimates results not only from selecting a sample of the population but also from uncertainties in the models used to predict the variables of interest. The aim of this study was to quantify the model-related variability of Norway spruce (*Picea abies* [L.] Karst) biomass stock and change estimates for the Norwegian NFI. The model-related variability of the estimates stems from uncertainty in parameter estimates of biomass models as well as residual variability and was quantified using a Monte Carlo simulation technique. Uncertainties in model parameter estimates, which are often not available for published biomass models, had considerable influence on the model-related variability of biomass stock and change estimates. The assumption that the residual variability is larger than documented for the models and the correlation of within-plot model residuals influenced the model-related variability of biomass stock change estimates much more than estimates of the biomass stock. The larger influence on the stock change resulted from the large influence of harvests on the stock change, although harvests were observed rarely on the NFI sample plots in the 5-year period that was considered. In addition, the temporal correlation between model residuals due to changes in the allometry had considerable influence on the model-related variability of the biomass stock change estimate. The allometry may, however, be assumed to be rather stable over a 5-year period. Because the effects of model-related variability of the biomass stock and change estimates were much smaller than those of the sampling-related variability, efforts to increase the precision of estimates should focus on reducing the sampling variability. If the model-related variability is to be decreased, the focus should be on the tree fractions of living branches as well as stump and roots.

Keywords: uncertainty, variance, mean squared error, biomass stock, biomass change, Monte Carlo simulation, parametric bootstrap

National forest inventories (NFIs) provide estimates of forest parameters for national and regional scales. Because of the vast geographical extent of forests, a full census is impossible. Therefore, NFI results are estimates based on measurements of a sample of the population. The sample usually consists of sample plots distributed over the country. Because only a sample is taken, uncertainty in the form of sampling-related variability is introduced into the estimates. Many variables of interest such as timber volume, biomass, or carbon cannot be measured directly in the field. Therefore, models are used to estimate these variables from tree-level measurements of field variables such as diameter and height. This process introduces uncertainty in the form of model-related variability. Other sources of uncertainty in NFI estimates are measurement errors during fieldwork and calculation errors during data processing for variables that are estimated from measured variables (e.g., Cunia 1965, Intergovernmental Panel on Climate Change 2000).

The uncertainty associated with forest inventory estimates is of central relevance in the further use of inventory results in policy-making and international reporting. For example, an inadequate representation of uncertainties in estimates can result in considerable biases for future projections (Valle 2011, Fortin and Langevin 2012). The uncertainty is also a declaration of quality and reveals the weaknesses of an inventory, which form the basis for identifying areas of further improvement.

The different sources of uncertainty in forest inventories and how they can be considered in estimates have long been studied. For

example, Cunia (1986a, 1986b, 1986c), Gertner (1990), and Gertner and Köhl (1992) analyzed the effect of different error sources under a variety of sampling designs. McRoberts and Westfall (2013) used a Monte Carlo simulation approach to quantify the uncertainty of model-related variability of timber volume estimates. Berger et al. (2013) used a Monte Carlo simulation approach to assess the effects of measurement errors in volume model predictor variables on individual tree volume estimates. Ståhl et al. (2011) introduced a model-based estimator that inherently considered the model- and sampling-related variability. The approach was applied by Ståhl et al. (2013) to quantify the model-related variability of national forest biomass estimates in Sweden and Finland. Whereas Wirth et al. (2004) carried out a meta-analysis with special attention to the uncertainty associated with biomass models, Peltoniemi et al. (2006) and Monni et al. (2007) focused more generally on factors influencing the uncertainty of greenhouse gas inventories.

As part of their greenhouse gas accounting, nations that have signed the United Nations Framework Convention on Climate Change are required to report estimates and uncertainties for the carbon stocks and carbon stock changes in their forests (Intergovernmental Panel on Climate Change 2000). Because tree biomass is an important component of the forest carbon stock, tree biomass uncertainty has recently received renewed attention. Tree biomass can be estimated from NFI data using biomass expansion factors (BEF) or biomass models (Petersson et al. 2012). BEFs are preferred if only aggregated volume measurements are available, whereas biomass models are applicable if single tree variables such as diameter and height are measured systematically (Somogyi et al. 2007). A range of different biomass models exists (for a review, see Jenkins et al. 2003, Zianis et al. 2005, Muukkonen and Mäkipää 2006). For the Nordic countries, the most widely applicable biomass models include those of Marklund (1987, 1988), Petersson and Ståhl (2006), and, more recently, Repola (2009) and Skovsgaard et al. (2011).

National or regional estimates of biomass are usually based on a two-phase survey, similar to the estimation of timber volume (Cunia 1986b). In a first phase (the forest inventory), auxiliary variables that correlate with tree biomass and are easy to measure, such as dbh and height, are measured on a large sample of trees. In a second phase, auxiliary variables and tree biomass are measured in a small number of trees. A regression model is then fit with biomass as the dependent variable and the auxiliary variables (e.g., dbh and height) as independent variables. The biomass model is then used to predict the single tree biomass of the trees measured in the first phase. Usually, the second-phase sample is so costly that it is carried out separately from the first phase, and the biomass models are applied in many subsequent forest inventories (first-phase samples). Therefore, it is not uncommon that the models do not fully cover the geographical range of the applications and are based on trees that were sampled many years before the forest inventory takes place. Sufficient information on the model uncertainty is then frequently missing, with the result that quantification of model-related variability is often accepted as impossible (Cunia 1986b).

In the Norwegian NFI, biomass models developed in Sweden (Marklund 1988, Petersson and Ståhl 2006) are currently used to predict the biomass of sampled trees (Norwegian Climate and Pollution Agency 2010). Models were developed and their parameters were estimated separately for the tree fractions of stem wood, bark, living branches, dead branches, and stump and roots, and must be combined to obtain the total tree biomass. The applied biomass models used are tree species-specific and use dbh and height as independent variables.

In this study, the biomass model-related variability of the biomass stock estimate of Norway spruce (*Picea abies* [L.] H. Karst.) and the stock change estimate between two consecutive inventory cycles were quantified. The model-related variability of the estimates stems from uncertainty in the parameter estimates of the biomass models and the residual variability around the model predictions. The model-related variability of the estimates was approximated using a Monte Carlo simulation method known as the parametric bootstrap (Efron and Tibshirani 1993, p. 53ff). Different properties of the biomass models contribute differently to the uncertainty of the estimate. Therefore, the overall goal was approached based on four research questions:

1. Because the biomass models were fit to Swedish data, it is possible that the model residual variance is larger under Norwegian conditions. Therefore, the influence of the larger residual variance on the model-related variability was assessed.

2. Information on the uncertainty and correlation of the parameter estimates is often missing for existing biomass models. Therefore, the contributions of uncertainty in parameter estimates on model-related variability were analyzed.

3. Trees close to each other tend to grow similarly. Hence, deviations of the actual biomass from the predicted biomass (i.e., the model residuals) tend to be correlated for trees on the same sample plot. Therefore, the effects of correlations among within-plot model residuals on model-related variability were analyzed.

4. Deviations of the actual biomass from the predicted biomass on the same tree at consecutive inventory cycles tend to be correlated. We evaluated how the temporal correlation of these deviations influenced the model-related variability of the biomass stock change.

Material—The Norwegian National Forest Inventory (NNFI)

The NNFI is a permanent sample plot inventory that covers the whole country (Tomter et al. 2010). Except for high mountain areas and some northern parts of the country, the permanent sample plots are located on a 3 × 3-km grid (Landsskogtakseringen 2010). Every year, 20% of the sample plots are visited, which results in 5-year remeasurement cycles. Among other attributes, dbh and tree species of all trees with dbh >5 cm are recorded on circular sample plots with a 8.92-m radius (250 m^2). On plots with 10 trees or fewer, all tree heights are measured using hypsometers. On plots with more than 10 trees, heights are measured from a relascope-selected subsample with a target sample size of 10 trees/plot (Landsskogtakseringen 2010). The heights of the unmeasured trees are estimated using tariff models calibrated at the plot level with the data from measured trees.

The representation factor (also known as design weight or expansion factor) of an observation in the NFI is given by dividing the land area of Norway (including lakes) by the number of potential sample plots on the 3 × 3-km grid independently of their land cover. Because of the NNFI design, the representation factor is 36044.94.

We used spruce trees sampled in the eighth (NNFI8) and ninth inventory cycle (NNFI9) to estimate the spruce biomass stock and

Table 1. NNFI sample tree characteristics.

	NNFI8			NNFI9		
	dbh (cm)	Height (m)	Predicted biomass (kg)	dbh (cm)	Height (m)	Predicted biomass (kg)
Mean	13.25	10.09	108.29	13.44	10.37	112.81
SD	7.43	4.73	162.80	7.54	4.88	168.64
Max.	71.20	33.10	2,286.40	74.00	34.30	2,646.36

Biomass was predicted according to Marklund (1988) and Petersson and Ståhl (2006).

the spruce biomass stock change between the two NNFI cycles. The sample plots in the NNFI8 and NNFI9 were visited between 2000 and 2004 and 2005 and 2009, respectively (Larsson and Hylen 2007). We defined our sample as the 10,384 NNFI plots on the 3 × 3-km grid that were forest according to the NNFI definition in both inventory cycles. Spruce trees were observed on $m = 5,831$ and 6,023 sample plots of the common grid in the NNFI8 and NNFI9, respectively. Harvests between the two inventory cycles were observed on 102 sample plots. The numbers of sample trees in the first phase were $n = 81,868$ and $n = 86,029$ in the NNFI8 and NNFI9, respectively. Of the spruce trees measured in the NNFI8, 75,593 were measured in the NNFI9. Characteristics of the sample trees are given in Table 1. A small proportion of trees (0.11% in the NNFI8 and 0.12% in the NNFI9) were within the extrapolation range with respect to the dbh and height of trees observed in the second-phase sample that was used to fit the biomass models (Marklund 1988, p. 8, Table 3). Nonetheless, the trees were well within the range of data for which the biomass models behave logically (Marklund 1988, p. 18).

Methods
Biomass Models

The biomass models G5, G8, G12, and G20 of Marklund (1988) and case B, category i of Petersson and Ståhl (2006) were used to predict the biomass of the tree fractions of stem wood under bark, bark, living branches including needles, dead branches, and stump with roots down to 2-mm diameter, respectively. These univariate biomass models have the general form

$$\hat{y}_i = e^{\boldsymbol{\beta}^T \boldsymbol{x}_i}, \qquad i = 1, \ldots, n_o \qquad (1)$$

where $\boldsymbol{\beta}$ is the coefficient vector, $\boldsymbol{x}_i$ is the vector of the explanatory variables $\boldsymbol{x}_i = (x_{0i}, x_{1i}, x_{2i}, x_{3i})$ with $x_0 = 1$ (intercept), $x_1 = $ dbh, $x_2 = $ height, and $x_3 = \ln(\text{height})$ and n_o is the number of second-phase observations used by Marklund (1988) and Petersson and Ståhl (2006) to fit the biomass models. The model for stump and roots includes dbh as the only explanatory variable such that $\boldsymbol{x}_i = (x_{0i}, x_{1i})$ in this case. To estimate the linear model coefficients, Marklund (1988) and Petersson and Ståhl (2006) transformed the left- and right-hand sides of models 1 with the natural logarithm. The resulting models are of the general form

$$\ln(y_i) = \boldsymbol{\beta}^T \boldsymbol{x}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2), \qquad (2)$$

where ε_i is the residual and σ^2 is the residual variance. Marklund (1988) and Petersson and Ståhl (2006) assumed the model residuals to be independently and identically normally distributed on the log scale.

The necessary spruce biomass models in Marklund (1988) and Petersson and Ståhl (2006) were refitted using the original data to obtain the parameter covariance matrices $\boldsymbol{\Psi}_f, f = 1, \ldots, p$ for each model f of the $p = 5$ tree fractions. The parameter covariance matrices can be found in Tables A1–A5 in the Appendix. Any dependence among the residuals resulting from spatial correlation was ignored. However, the dependence of residuals between models was assessed using the model residuals. The model residuals of $n_o^{(C)}$ trees with a complete record of all tree fractions can be used to obtain the $p \times p$ residual covariance matrix

$$\boldsymbol{\Sigma} = \text{cov}(\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_f, \ldots, \boldsymbol{\varepsilon}_p) \qquad (3)$$

where $\boldsymbol{\varepsilon}_f = \{\varepsilon_{f, 1}, \ldots, \varepsilon_{f, n_o^{(C)}}\}^T$ is the vector of residuals of model f. However, because Petersson and Ståhl (2006) used a different data set (second-phase sample) than Marklund (1988), $\boldsymbol{\Sigma}$ cannot be directly calculated from the residuals of Petersson and Ståhl (2006) and Marklund (1988) models. Therefore, we followed a two-step procedure to obtain $\boldsymbol{\Sigma}$. First, the correlation matrix

$$\boldsymbol{\Omega} = \text{cor}(\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_f, \ldots, \boldsymbol{\varepsilon}_p) \qquad (4)$$

was calculated by using the residuals of the refitted model G23 for stump and roots by Marklund (1988) instead of the residuals of the model by Petersson and Ståhl (2006).

Second, the residual covariance matrix $\boldsymbol{\Sigma}$ was calculated as

$$\boldsymbol{\Sigma} = \boldsymbol{\Omega}(\boldsymbol{\sigma} \otimes \boldsymbol{\sigma}^T) \qquad (5)$$

where

$$\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_f, \ldots, \sigma_p\}^T \qquad (6)$$

is the vector of model residual SDs and $\otimes$ is the outer product. The residual SD of the model for stump and roots by Petersson and Ståhl (2006) was used in Equation 6. The residual covariance matrix can be found in Table A6 in the Appendix.

Between 505 and 540 trees were used by Marklund (1988) to fit the biomass models for the aboveground fractions, and 316 trees were used to fit the model for stump and roots. There were $n_o^{(C)} = 274$ trees with measurements of all tree fractions that are the basis for $\boldsymbol{\Omega}$.

The residual variance of the biomass models may be greater under Norwegian conditions than for Swedish forests. To simulate a greater residual variance and test the first research question, a new variance-covariance matrix was calculated by multiplying the residual SDs σ_f in Equation 6 by 1.5. This means a 50% greater residual SD than that reported by Marklund (1988) and Petersson and Ståhl (2006) was assumed. The vector of inflated model residual SDs ($\boldsymbol{\sigma}$) was then multiplied with the correlation matrix of the residuals (Equation 4) to obtain $\boldsymbol{\Sigma}$ under the assumption of inflated residual variances.

Model-Related Variability of Biomass Stock and Change Estimates

Parametric Bootstrap of the Biomass Stock

For purposes of focusing specifically on the effect of model-related variability on estimates, NNFI measurements of dbh and both measurements and tariff model estimates of height were assumed to be error-free. The parametric bootstrap was carried out as follows

1. From the distribution $N(\mathbf{0}, \mathbf{\Psi}_f)$ one random sample $\mathbf{e}_f^*$ was drawn for every biomass model f. The vector $\mathbf{e}_f^*$ was added to the original model parameters published by Marklund (1988) and Petersson and Ståhl (2006). Based on these new model parameter estimates and the dbh and height measurements of the NNFI8 and NNFI9,[1] biomass by tree fraction was predicted as $\ln(\bar{\mathbf{y}}_{ij}) = \{\ln(\bar{y}_{1,ij}), \ldots, (\bar{y}_{f,ij}), \ldots, (\bar{y}_{p,ij})\}^T$ for the $j = 1, \ldots, n_i$ trees on the $i = 1, \ldots, m$ sample plots.

2. Because of similar growing conditions, the biomass model residuals for trees on the same plot tend to be correlated (e.g., Repola 2009). The plot-level correlation was not considered in the biomass models (Marklund 1988, Petersson and Ståhl 2006; Equation 2). Therefore, the concept of weighted residuals as described by Lehtonen et al. (2007), was applied. A sample of n random numbers was drawn from the multivariate normal distribution $N(\mathbf{0}, \mathbf{\Sigma})$, resulting in the matrix $\mathbf{\Pi}^{*(T)}$ of the dimensions $n \times p$. Every row in $\mathbf{\Pi}^{*(T)}$ is a p vector of simulated model residuals $\boldsymbol{\varepsilon}_{ij}^{*(T)} = \{\varepsilon_{1,ij}^{*(T)}, \ldots, \varepsilon_{f,ij}^{*(T)}, \ldots, \varepsilon_{p,ij}^{*(T)}\}^T$ for tree j on sample plot i. The superscript (T) indicates the simulated residuals for trees. Analogously, a sample of m random numbers $\mathbf{\Pi}^{*(P)}$ was drawn for each sample plot i. The superscript (P) indicates the simulated residuals for plots. The weighted residuals then result from

$$\varepsilon_{ij}^* = \frac{1}{\sqrt{(1-w)^2 + w^2}}[(1-w) \cdot \varepsilon_{ij}^{*(T)} + w \cdot \varepsilon_i^{*(P)}] \qquad (7)$$

where the weight w is the within-plot correlation.

3. The simulated residuals were added to the biomass predictions on the logarithmic scale. The result was back-transformed to the original scale and multiplied with the representation factor

$$\hat{\mathbf{y}}_{ij}^* = \exp[\ln(\bar{\mathbf{y}}_{ij}) - \boldsymbol{\lambda} + \boldsymbol{\varepsilon}_{ij}^*]\mathbf{r} \qquad (8)$$

where $\boldsymbol{\lambda}$ is the p vector of bias correction factors specific for the respective biomass model[2] and $\mathbf{r}_{ij}$ is the p vector of representation factors.

4. The sum of tree biomass estimates over all n trees is one parametric bootstrap sample of the biomass stock by tree fractions

$$\hat{\mathbf{Y}}^* = \sum_j^n \hat{\mathbf{y}}_j^*$$

where $\hat{\mathbf{Y}}^* = \{\hat{Y}_1^*, \ldots, \hat{Y}_p^*\}$ is a p vector of biomass total estimates by tree fraction. The sum over the tree fractions is the biomass stock denoted $\hat{Y}^*$ resulting from one bootstrap replication.

Steps 1–4 were repeated 2,000 times, resulting in a set of bootstrap replications. The mean of the bootstrap replications $\hat{Y}^*$ is the parametric bootstrap estimate $\hat{Y}^B$ of the biomass stock. The variance of the bootstrap replications $\hat{Y}^*$ is the mean squared error resulting from model uncertainty (MSE_M).

The parametric bootstrap was applied to the spruce trees observed on the sample plots in the NNFI8 and NNFI9 as described in the Material section. Whereas the study of Repola (2009) suggests a within-plot correlation w of 0.5 (for total tree biomass), it was also tested with the values 0 (no correlation on plot level) and 1 (full correlation on plot level; all trees on one plot have the same simulated residual). R software for statistical computing (R Core Team 2012) was used for all analysis.

Parametric Bootstrap of the Biomass Change

The biomass change is given as $\hat{Y}^{B(9)} - \hat{Y}^{B(8)}$, where the superscripts 8 and 9 indicate the bootstrap estimates for the NNFI8 and NNFI9, respectively. It is probable that if the biomass model overestimates (underestimates) the true biomass of a tree in one inventory cycle, the biomass model will overestimate (underestimate) the true biomass of the same tree also in the next inventory cycle. The concept of weighted residuals (Lehtonen et al. 2007) was also used to model this temporal correlation of biomass model residuals. In Equation 7, $\varepsilon_{ij}^{*(T)}$ and $\varepsilon_{ij}^{*(P)}$ are substituted by the within-plot correlation-weighted residuals of the NNFI8 and NNFI9 denoted by $\varepsilon_{ij}^{*(9)}$ and $\varepsilon_{ij}^{*(8)}$, respectively, to model the temporal correlation of biomass model residuals. The new vector of simulated residuals, ε_{ij}^*, is then used in Equation 8. The weight w is given by the temporal correlation. Temporal correlations of 0.5, 0.9, and 1 were simulated.

Sampling-Related Variability of Biomass Stock and Change Estimates

For comparison to the model-related variability, the sampling-related variabilities of the biomass stock and change estimates were also calculated. To focus on the sampling-related variability, the biomass, predicted with the models described above, was assumed to be a measurement without error.

The sample estimate of the biomass stock is given by the ratio estimator (Thompson 2002, Petersson et al. 2012)

$$\hat{Y}^S = \frac{M}{m} \sum_{i=1}^m y_i$$

where M is the population size,[3] m is the number of sample plots, and $y_i = \sum_{j=1}^{n_i} y_{ij}$ is the sum of the biomass predictions y_{ij} for the n_i trees j on plot i. Because the true population size was unknown, M was estimated from the sample by multiplying the number of sample plots by the representation factor. The variance resulting from the estimation of the forest area was ignored to focus on the sampling-related variability. The sample size m was the number of sample plots with observed spruce trees. Because $\hat{Y}^S$ was an estimate of the biomass stock over all tree fractions, y_{ij} was the sum of the biomass predictions by tree fraction per tree.

The introduction of the plot-level biomass is necessary to calculate the variance (or MSE_S) of the estimate because observations for trees measured on the same plot are not independent of each other. An estimate of the variance of the total biomass estimate was given by

$$MSE_S = \widehat{var}(\hat{Y}^S) = M^2 \frac{s^2}{m}$$

Table 2. Biomass stock and change estimates and their associated uncertainties (RMSE) in 1,000,000 Mg.

	NNFI8			NNFI9			Change		
Method	Boot.	Boot.	Original sample	Boot.	Boot.	Original sample	Boot.	Boot.	Original sample
Uncertainty in parameter estimates	Incl.	Excl.		Incl.	Excl.		Incl.	Excl.	
Estimate	319.663	319.520	319.551	349.902	349.730	349.750	30.239	30.210	30.199
RMSE	3.087	0.733	5.073	3.381	0.788	5.476	0.410	0.258	2.507

Parametric bootstrap estimates (Boot.) included (Incl.) and excluded (Excl.) uncertainty in parameter estimates. The residual variance was as reported by Marklund (1988) and Petersson and Ståhl (2006) and the within-plot correlation was set to 0.5. Sample estimates (Original sample) are based on the assumption of a measurement of single tree biomass.

Table 3. Biomass stock and change estimates and their associated uncertainties (RMSE) in 1,000,000 Mg.

	NNFI8		NNFI9		Change	
Method	Boot.	Original sample	Boot.	Original sample	Boot.	Original sample
Estimate	319.640	319.551	349.877	349.750	30.237	30.199
RMSE	3.198	5.073	3.500	5.476	0.509	2.507

Parametric bootstrap estimates (Boot.) included uncertainty in parameter estimates and are based on the assumption of a 50% inflated residual SD compared with values as reported by Marklund (1988) and Petersson and Ståhl (2006). The within-plot correlation was set to 0.5. Sample estimates (Original sample) are based on the assumption of a measurement of single tree biomass.

Table 4. Biomass stock estimates and their associated uncertainties (RMSE) in 1,000,000 Mg.

	NNFI8				NNFI9			
Method	Boot.	Boot.	Boot.	Original sample	Boot.	Boot.	Boot.	Original sample
Within-plot correlation	1	0.5	0		1	0.5	0	
Stock estimate	319.665	319.663	319.654	319.551	349.901	349.902	349.896	349.750
RMSE	3.163	3.087	3.008	5.073	3.464	3.381	3.297	5.476

Parametric bootstrap estimates (Boot.) include uncertainty in parameter estimates and residual variances as reported by Marklund (1988) and Petersson and Ståhl (2006). Within-plot correlations of 1, 0.5, and 0 were simulated. Sample estimates (Original sample) are based on the assumption of a measurement of single tree biomass.

where

$$s^2 = \frac{1}{m-1} \sum_{i=1}^{m} (y_i - \bar{y})^2$$

is the sample variance and $\bar{y}$ is the mean of the y_i's.

In the case of the estimation of the biomass stock change, y_i is the change of the measured biomass on a sample plot between two periods. In our case, this was the observation in the NNFI9 minus the observation in the NNFI8.

Total Variability

The sum of the model- and sampling-related variability is the total variability, $MSE_T = MSE_M + MSE_S$. The root mean squared error (RMSE) is the square root of MSE.

Results
Uncertainty in Parameter Estimates and Residual Variance

If the uncertainty in the model parameter estimates was considered, the model-related variability ($RMSE_M$) of the biomass stock and biomass stock change estimate was considerably larger (1.5 and 4.2 times) than if only residual variability was considered (Table 2). Because ignoring the uncertainty in the parameter estimates resulted in unrealistically small model-related variability of the estimates, model parameter uncertainty is included in all further results.[4]

Inflating the residual variance by 50% compared with the re-ported values (Marklund 1988, Petersson and Ståhl 2006) (termed inflated residual variance in the following), increased the $RMSE_M$ of the biomass stock and biomass stock change by 4 and 24%, respectively (Table 3).

If an inflated residual variance was assumed, the total variability ($RMSE_T$) of the biomass stock estimate was approximately 2% of the estimate in both inventory cycles. The model-related variability of the estimate (MSE_M) accounted for approximately 28% of the total variability (MSE_T). The total variability of the biomass stock change estimate was approximately 8% of the estimate. For the biomass stock change, MSE_M accounted for approximately 4% of the MSE_T. Similar results were obtained if the residual variance was not inflated (Tables 2 and 3).

Within-Plot Correlation

An increased within-plot correlation of the model residuals slightly increased the model-related variability of the biomass stock estimate (Table 4). The influence of the within-plot correlation on the model-related variability of the biomass change estimate was greater than for the stock estimate. The $RMSE_M$ of the stock change increased by approximately 19% if the within-plot correlation was assumed to be 0.5 compared with ignoring the within-plot correlation. The model-related variability further increased by another 15% if the within-plot correlation was assumed to be 1.0 (Table 5). Because of the large proportion of the sampling variability, the total

Table 5. Biomass change estimates and their associated uncertainties (RMSE) in 1,000,000 Mg.

Method	Boot.	Boot.	Boot.	Original sample
Within-plot correlation	1	0.5	0	
Change estimate	30.235	30.239	30.242	30.199
RMSE	0.473	0.410	0.345	2.507

Parametric bootstrap estimates (Boot.) include uncertainty in parameter estimates, a temporal correlation of 0.9, and residual variances as reported by Marklund (1988) and Petersson and Ståhl (2006). Within-plot correlations of 1, 0.5, and 0 were simulated. Sample estimates (Original sample) are based on the assumption of a measurement of single tree biomass.

Table 6. Biomass change estimates and their associated uncertainties (RMSE) in 1,000,000 Mg.

Method	Boot.	Boot.	Boot.	Original sample
Temporal correlation	1	0.9	0.5	
Change estimate	30.240	30.239	30.229	30.199
RMSE	0.401	0.410	0.696	2.507

The temporal correlation was set to 1, 0.9, and 0.5. Parametric bootstrap estimates (Boot.) are based on a within-plot correlation of 0.5 and residual variances as reported by Marklund (1988) and Petersson and Ståhl (2006). Uncertainty in parameter estimates was considered. Sample estimates (Original sample) are based on the assumption of a measurement of single tree biomass.

variability was barely affected by the increased model-related variability resulting from the increased within-plot correlation.

Temporal Correlation

In the above paragraphs, the temporal correlation between the model residuals of a tree that was measured in the NNFI8 and NNFI9 was set to 0.9. Increasing the temporal correlation to 1.0 barely influenced the model-related variability of the estimate. Decreasing the temporal correlation to 0.5 increased the $RMSE_M$ by 70% (Table 6). As for the within-plot correlation, the total variability was barely affected by changes in the temporal correlation due to the large proportion of the sampling-related variability.

Model-Related Variability of Estimates by Tree Fraction

The tree fractions of living branches, stump and roots, and stem wood had the greatest influence on the model-related variability of the biomass stock and change estimates (Figure 1, Table 7). It should be noted that the MSE_M values by tree fraction are not additive because of the correlation between the tree fractions. The MSE_M values were in the same order for the NNFI9 as shown for the NNFI8. The rankings of the MSE_M values according to size were the same for the biomass change as for the stock.

Discussion

The aim of the study was to quantify the model-related variability of biomass stock and change estimates of Norway spruce measured during two consecutive NNFI cycles. With use of a Monte-Carlo simulation, the uncertainties in estimates of the different biomass components were analyzed with respect to how they influenced the model-related variability of biomass stock and change estimates under different assumptions.

Of all the effects analyzed, the uncertainty in parameter estimates had the greatest influence on the model-related variability. If the information on the covariance structure of the model parameters would not have been available and thus been ignored, the model
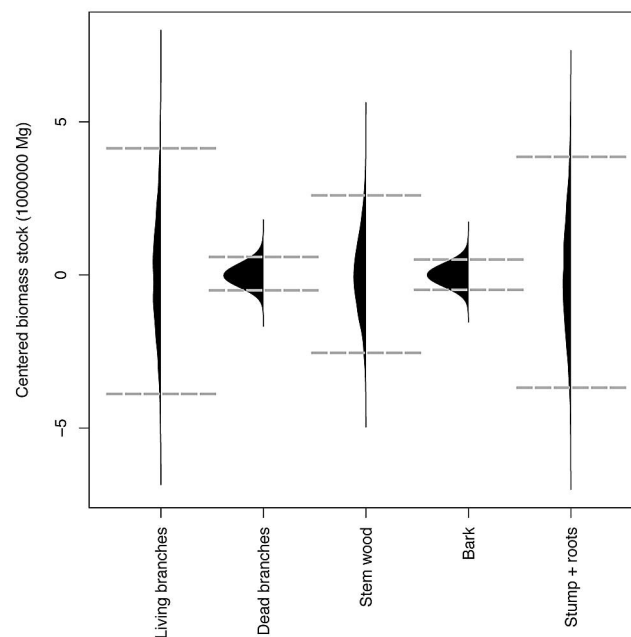
Figure 1. Distribution of the parametric bootstrap samples of the biomass stock of the NNFI8 centered around the mean prediction by tree fraction assuming residual variances as reported by Marklund (1988) and Petersson and Ståhl (2006) and a within-plot correlation of 0.5. Uncertainty in parameter estimates was considered. Dashed lines indicate empirical 95% confidence intervals of the model-related uncertainty in the estimate.

Table 7. Biomass estimates of the NNFI8 by tree fraction and their associated model-related variability in 1,000,000 Mg.

	Living branches	Dead branches	Stem wood	Bark	Stump and roots
Stock estimate	85.197	3.570	133.147	14.697	83.051
RMSE	2.016	0.215	1.248	0.252	1.960

The within-plot correlation was set to 0.5. Residual variances as reported by Marklund (1988) and Petersson and Ståhl (2006) were assumed.

uncertainty would have been considerably underestimated. Unfortunately, information on the covariance structure of model parameter estimates is usually missing in published biomass models. Given the weight of uncertainty in parameter estimates in the model-related variability and the increasing interest in biomass uncertainty, we recommend that future biomass models report the covariance structure of model parameters. To facilitate the analysis of the uncertainty in estimates and forecasts of other variables of interest besides biomass, this recommendation could be extended to forestry models in general.

The model-related variability, as a proportion of the total variability, was considerably greater for the biomass stock than for the change in biomass stock, because the model residuals will be very similar for trees measured in two consecutive forest inventories and thus compensate for each other. Similar results were reported by Ståhl et al. (2013), who used an analytical approach to estimate the model-related variability of national biomass estimates in Sweden and Finland. Whereas Ståhl et al. (2013) considered the aboveground biomass of spruce, pine, and birch, the order of magnitude of the model-related variability of biomass stock and change estimates was similar as in this study.

The results of this study are based on the assumption that Swedish biomass models (Marklund 1988, Petersson and Ståhl 2006) are transferable to Norwegian conditions. This assumption was generally corroborated by two studies that were, however, based on a small, nonrepresentative sample of trees (B.H. Øyen, unpubl. data, 2009; Andreassen 2009). Because it may be assumed that the Swedish biomass models fit the Norwegian data less well than the Swedish data, the residual variance was inflated with respect to Marklund (1988) and Petersson and Ståhl (2006). The choice of a 50% inflated residual variance was, however, arbitrary. The inflated residual variance particularly affected the model-related variability of the biomass stock change estimate. Nonetheless, because the model-related variability of the biomass stock change estimate was small compared with the sampling-related variability, an increase in the residual SD of 50% only marginally increased the total variability.

Increased within-plot correlation resulted in increased model-related variability of the biomass stock change estimate. The same tendency, although less obvious, was observed for the biomass stock. The within-plot correlation describes the similarity of the deviation of the true biomass from the biomass prediction of the model for the trees on one plot. With increased similarity, the residual deviations compensate for each other to a lesser degree and the uncertainty increases. For the same reason, the influence of the within-plot correlation on the model-related variability will increase with a decreasing number of sample plots. Plot-level random-effects models (Repola 2009, Skovsgaard et al. 2011) would be necessary to consider the influence of the within-plot correlation in a more realistic way.

Although the total variability of the estimates was barely affected, it was somewhat surprising that the effects of increased residual variability and within-plot residual correlations on the model-related variability were greater for the biomass stock change estimate than for the biomass stock estimate. The Monte Carlo simulations were therefore repeated for sample plots on which no harvesting occurred between NNFI8 and NNFI9. Barely any influence of increased residual variance and within-plot correlation on the model-related variability of the biomass stock change estimate was visible under this restriction. We can therefore conclude that harvests, although rarely occurring in a 5-year period, have a strong influence on the model-related variability.

The temporal correlation between trees measured in both NNFI cycles had a large effect on the model-related variability of the biomass stock change estimate. Quantifying the actual temporal correlation of the deviation between true biomass and predicted biomass is practically impossible because true biomass is measured using destructive sampling. The difference between true and predicted biomass is related to the allometry of the tree; therefore, an indication of the temporal correlation could be obtained by considering other allometric measurements that are obtained in a nondestructive way. For example, the temporal correlation of the ratio of dbh and height in 8,400 measured trees in the NNFI8 and NNFI9 was 0.97, which suggests that a temporal correlation of 0.9 is a reasonable or perhaps slightly conservative assumption.

Independently of the assumptions made in the bootstrap procedure, the model-related variability was considerably less than the sampling-related variability, especially for the biomass change. Similar results were reported for the estimation of timber volume, given that the tree attributes were measured without systematic errors (Gertner 1990, Gertner and Köhl 1992).

Although by far the most biomass is accumulated in the tree fraction stem wood, the contribution to the model-related variability is greater for living branches as well as stump and roots. This results from the comparatively precise biomass model for stem wood Marklund (1988) and Petersson and Ståhl (2006).

This study focused on the spruce biomass model-related variability of estimates of biomass stock and biomass stock change in Norway. The methodology and approach are, however, generally applicable to other permanent sample plots inventories. Other approaches might be feasible; for example, we could envision an analytical approach to estimate uncertainty for a single biomass equation (e.g., Ståhl et al. 2011, 2013). The Monte Carlo approach applied here is robust, flexible, and intuitive although more computer-intensive than the analytical approach.

Conclusions

The model-related variability of the biomass stock and change estimates of Norway spruce was considerably less than the sampling-related variability of the estimates. This was true for all the different assumptions on the model structure (e.g., within-plot correlation of residuals etc.) made during the bootstrapping procedure. Although the model-related variability of the biomass stock estimate was considerable, the total variability combining model- and sampling-related variability was small compared with the estimate. Whereas the model-related variability of the biomass stock change estimate was small, the total variability was considerable compared with the estimate.

The model-related variabilities of the biomass stock and change estimates were both strongly affected by considering uncertainties in the model parameter estimates. However, assumptions on inflated residual variances and within-plot correlations affected the model-related variability of the biomass stock change estimate much more than the estimate of the biomass stock. This was mainly driven by the large influence of the few plots where harvests occurred. The temporal correlation of model residuals also strongly affected the model-related variability of the biomass stock change estimate but changes in allometry are usually small.

The results suggest that if the objective is reduction in the total uncertainty of biomass or biomass change estimates, then the main emphasis should be to reduce the sampling-related variability. If the model-related variability is to be decreased, the focus should be on the tree fractions of living branches as well as stump and roots. In further studies, investigation into how measurement errors and the uncertainty in the height model for trees without height measurements contribute to the uncertainty of the biomass stock and change estimates given the Norwegian National Forest Inventory design will be performed.

Endnotes

1. This means that the same model parameters were used for trees in the NNFI8 and NNFI9 in one bootstrap repetition.
2. The standard bias correction factor $RMSE^2/2$ was used. It should be noted that a bias correction is not necessary in this case because an individual observation is predicted (including a simulated residual error component), which reconstructs the error structure. Because Marklund's models include already a bias correction factor in the intercept, it had to be subtracted.
3. Often, the concept of land area is used (e.g., Petersson et al. 2012), which is equivalent to the representation here. The population size is the number of units of the same size as a sample plot in the population.
4. The bootstrap estimates and the sample estimate of the total biomass stock differ slightly. Because this bootstrap bias was small compared with the model-related variability, it was ignored (Efron and Tibshirani 1993, p. 128).

Literature Cited

ANDREASSEN, K. 2009. *Comparison of some Nordic biomass equations in Norwegian forest. Preliminary results.* Presented at the Nordic Forest Research Co-operation Committee (SNS) Meeting, Kokkola, Finland June 8–11, 2009.

BERGER, A., T. GSCHWANTNER, R.E. MCROBERTS, AND K. SCHADAUER. 2013. Effects of measurement errors on individual tree stem volume estimates for the Austrian National Forest Inventory. *For. Sci.* In press.

CUNIA, T. 1965. Some theory on reliability of volume estimates in a forest inventory sample. *For. Sci.* 11(1):115–128.

CUNIA, T. 1986a. Construction of tree biomass tables by linear regression technique. In *Estimating tree biomass regressions and their error. Proc. of the Workshop on tree biomass regression functions and their contribution to the error of forest inventory estimates*, Wharton, E., and T. Cunia (eds.). USDA For. Serv., Gen. Tech. Rep. NE-117, Northeastern Forest Experimental Station, Broomall, PA. 10 p.

CUNIA, T. 1986b. Error of forest inventory estimates: Its main components. In *Estimating tree biomass regressions and their error. Proc. of the Workshop on tree biomass regression functions and their contribution to the error of forest inventory estimates*, Wharton, E., and T. Cunia (eds.). USDA For. Serv., Gen. Tech. Rep. NE-117, Northeastern Forest Experimental Station, Broomall, PA. 15 p.

CUNIA, T. 1986c. On the error of forest inventory estimates: Stratified sampling and double sampling for stratification. In *Estimating tree biomass regressions and their error. Proc. of the Workshop on tree biomass regression functions and their contribution to the error of forest inventory estimates*, Wharton, E., and T. Cunia (eds.). USDA For. Serv., Gen. Tech. Rep. NE-117, Northeastern Forest Experimental Station, Broomall, PA. 9 p.

EFRON, B., AND R. TIBSHIRANI. 1993. *An introduction to the bootstrap.* Chapman & Hall/CRC, Boca Raton, FL. 436 p.

FORTIN, M., AND L. LANGEVIN. 2012. Stochastic or deterministic single-tree models: Is there any difference in growth predictions? *Ann. For. Sci.* 69:271–282.

GERTNER, G., AND M. KÖHL. 1992. An assessment of some nonsampling errors in a national survey using an error budget. *For. Sci.* 38(3):525–538.

GERTNER, G. 1990. The sensitivity of measurement error in stand volume estimation. *Can. J. For. Res.* 20(6):800–804.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. 2000. Chapter 6: Quantifying Uncertainties in Practice. In *Good practice guidance and uncertainty management in national greenhouse gas inventories.* Institute for Global Environmental Strategies for the Intergovernmental Panel on Climate Change, Technical Report, Geneva, Switzerland. 34 p.

JENKINS, J., D. CHOJNACKY, L. HEATH, AND R. BIRDSEY. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12–35.

NORWEGIAN CLIMATE AND POLLUTION AGENCY. 2010. *National Inventory Report 2010 Norway.* Norwegian Climate and Pollution Agency, Tech. Rep. TA 2639, Oslo, Norway.

LANDSSKOGTAKSERINGEN. 2010. *Landsskogtakseringens feltinstruks 2010, Håndbok fra Skog og landskap 05/08.* Skog og landskap, Ås, Norway. 128 p.

LARSSON, J.Y., AND G. HYLEN. 2007. *Skogen i Norge. Statistikk over skogforhold og skogressurser i Norge registrert i perioden 2000–2004.* [*Statistics of forest conditions and forest resources in Norway*.]. Viten fra Skog og landskap 1/07. Skog og landskap, Ås, Norway. 91 p.

LEHTONEN, A., E. CIENCIALA, F. TATARINOV, AND R. MÄKIPÄÄ. 2007. Uncertainty estimation of biomass expansion factors for Norway spruce in the Czech Republic. *Ann. For. Sci.* 64:133–140.

MARKLUND, L. 1987. *Biomass functions for Norway spruce in Sweden.* Department of Forest Survey, Rep. 43, Swedish University of Agricultural Sciences, Valsätra-Ultuna, Sweden. 127 p.

MARKLUND, L. 1988. *Biomass functions for pine, spruce and birch in Sweden.* Department of Forest Survey, Rep. 45, Swedish University of Agricultural SciencesValsätra-Ultuna, Sweden. 73 p.

MCROBERTS, R.E., AND J.A. WESTFALL. 2013. The effects of uncertainty in model predictions of individual tree volume on large area volume estimates. *For. Sci.* In press.

MONNI, S., M. PELTONIEMI, T. PALOSUO, A. LEHTONEN, R. MÄKIPÄÄ, AND I. SAVOLAINEN. 2007. Uncertainty of forest carbon stock changes—Implications to the total uncertainty of GHG inventory of Finland. *Climatic Change* 81(3):391–413.

MUUKKONEN, P., AND R. MÄKIPÄÄ. 2006. Biomass equations for European trees: Addendum. *Silva Fenn.* 40(4):763–773.

PELTONIEMI, M., T. PALOSUO, S. MONNI, AND R. MÄKIPÄÄ. 2006. Factors affecting the uncertainty of sinks and stocks of carbon in Finnish forests soils and vegetation. *For. Ecol. Manage.* 232(1–3):75–85.

PETERSSON, H., AND G. STÅHL. 2006. Functions for below-ground biomass of *Pinus sylvestris, Picea abies, Betula pendula*, and *Betula pubescens* in Sweden. *Scand. J. For. Res.* 21(Suppl. 7):84–93.

PETERSSON, H., S. HOLM, G. STÅHL, D. ALGER, J. FRIDMAN, A. LEHTONEN, A. LUNDSTRÖM, AND R. MÄKIPÄÄ. 2012. Individual tree biomass equations or biomass expansion factors for assessment of carbon stock changes in living biomass—A comparative study. *For. Ecol. Manage.* 270:78–84.

R CORE TEAM. 2012. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

REPOLA, J. 2009. Biomass equations for Scots pine and Norway spruce in Finland. *Silva Fenn.* 43(4):625–647.

SKOVSGAARD, J., C. BALD, AND T. NORD-LARSEN. 2011. Functions for biomass and basic density of stem, crown and root system of Norway spruce (*Picea abies* (L.) Karst.) in Denmark. *Scand. J. For. Res.* 26(Suppl. 11):3–20.

SOMOGYI, Z., E. CIENCIALA, R. MÄKIPÄÄ, P. MUUKKONEN, A. LEHTONEN, AND P. WEISS. 2007. Indirect methods of large-scale forest biomass estimation. *Eur. J. For. Res.* 126(2):197–207.

STÅHL, G., S. HOLM, T. GREGOIRE, T. GOBAKKEN, E. NÆSSET, AND R. NELSON. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Can. J. For. Res.* 41(1):96–107.

STÅHL, G., J. HEIKKINEN, H. PETERSSON, J. REPOLA, AND S. HOLM. 2013. Sample based estimation of greenhouse gas emissions from forests—A new approach to account for both sampling and model errors. *For. Sci.* In press.

THOMPSON, S. 2002. *Sampling,* 2nd ed. Wiley-Interscience, Hoboken, NJ. 367 p.

TOMTER, S., G. HYLEN, AND J.E. NILSEN. 2010. National forest inventory report: Norway. In *National forest inventories: Pathways for common reporting*, Tomppo, E., T. Gschwantner, M. Lawrence, and R.E. McRoberts. Springer Verlag, New York. 13 p.

VALLE, D. 2011. Incorrect representation of uncertainty in the modeling of growth leads to biased estimates of future biomass. *Ecol. Appl.* 21(4):1031–1036.

WIRTH, C., J. SCHUMACHER, AND E. SCHULZE. 2004. Generic biomass functions for Norway spruce in Central Europe—A meta-analysis approach toward prediction and uncertainty estimation. *Tree Physiol.* 24(2):121–139.

ZIANIS, D., P. MUUKKONEN, R. MÄKIPÄÄ, AND M. MENCUCCINI. 2005. *Biomass and stem volume equations for tree species in Europe.* Silva Fennica Monographs, vol. 4. Finnish Society of Forest Science, Finnish Forest Research Institute, Vantaa, Finland. 63 p.

Appendix

Table A1. Parameter covariance matrix (Ψ_β) of the biomass model for stem wood (G5) by Marklund (1988).

	Intercept	dbh	Height	Log(height)
Intercept	0.00187			
dbh	0.00063	0.02804		
Height	0.00011	−0.00008	0.00001	
Log(height)	−0.00149	−0.00566	−0.00008	0.00226

Table A2. Parameter covariance matrix (Ψ_β) of the biomass model for stem bark (G8) by Marklund (1988).

	Intercept	dbh	Height	Log(height)
Intercept	0.00612			
dbh	0.00238	0.09170		
Height	0.00035	−0.00034	0.00003	
Log(height)	−0.00491	−0.01770	−0.00027	0.00709

Table A3. Parameter covariance matrix (Ψ_β) of the biomass model for living branches (G12) by Marklund (1988).

	Intercept	dbh	Height	Log(height)
Intercept	0.00454			
dbh	0.00418	0.16322		
Height	0.00027	−0.00030	0.00004	
Log(height)	−0.00428	−0.03535	−0.00027	0.01096

Table A4. Parameter covariance matrix (Ψ_β) of the biomass model for dead branches (G20) by Marklund (1988).

	Intercept	dbh	Height	Log(height)
Intercept	0.03943			
dbh	0.04184	1.36314		
Height	0.00215	−0.00762	0.00037	
Log(height)	−0.03575	−0.22849	−0.00153	0.06607

Table A5. Parameter covariance matrix (Ψ_β) of the biomass model for stump and roots by Petersson and Ståhl (2006).

	Intercept	dbh
Intercept	0.00230	
dbh	−0.00438	0.00963

Table A6. Residual covariance matrix Σ.

	Res. G12	Res. G20	Res. G5	Res. G8	Res. PS0.06
Res. G12	0.13976				
Res. G20	−0.10997	1.12792			
Res. G5	−0.00065	0.05229	0.02366		
Res. G8	0.01360	0.05564	0.02072	0.07749	
Res. PS0.06	0.03414	0.00251	0.01133	0.02912	0.10438

G5, G8, G12, and G20 refer to biomass models by Marklund (1988). PS0.06 refers to the biomass model by Petersson and Ståhl (2006). Res., residual.