Scandinavian Journal of Forest Research

ISSN: 0282-7581 (Print) 1651-1891 (Online) Journal homepage: http://www.tandfonline.com/loi/sfor20

Advances and emerging issues in national forest inventories

Ronald E. McRoberts , Erkki O. Tomppo & Erik Næsset

To cite this article: Ronald E. McRoberts , Erkki O. Tomppo & Erik Næsset (2010) Advances and emerging issues in national forest inventories, Scandinavian Journal of Forest Research, 25:4, 368-381, DOI: 10.1080/02827581.2010.496739

To link to this article: http://dx.doi.org/10.1080/02827581.2010.496739

Published online: 28 Jul 2010.

Scandinavian Journal of Forest Research, 2010; 25: 368–381

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Advances and emerging issues in national forest inventories

RONALD E. McROBERTS[1], ERKKI O. TOMPPO[2] & ERIK NÆSSET[3]

[1]*Northern Research Station, US Department of Agriculture Forest Service, 1992 Folwell Avenue, St Paul, Minnesota, MN 55108, USA,* [2]*Finnish Forest Research Institute, Helsinki, Finland, and* [3]*Norwegian University of Life Sciences, Ås, Norway*

Abstract

National forest inventories (NFIs) have a long history, although their current major features date only to the early years of the twentieth century. Recent issues such as concern over the effects of acid deposition, biodiversity, forest sustainability, increased demand for forest data, international reporting requirements and climate change have led to the expansion of NFIs to include more variables, greater diversity in sampling protocols and a generally more holistic approach. This review focuses on six selected topics: (1) a brief historical review; (2) a summary of common structural features of NFIs; (3) a brief review of international reporting requirements using NFI data with an emphasis on approaches to harmonized estimation; (4) an overview of inventory estimation methods that can be enhanced with remotely sensed data; (5) an overview of nearest neighbors prediction and estimation techniques; and (6) a brief overview of several emerging issues including carbon inventories in developing countries and use of lidar data. Although general inventory principles will remain unchanged, sampling designs, plot configurations and measurement protocols will require modification before they can be applied in countries with tropical forests. Technological advances, particularly in the use of remotely sensed data, including lidar data, have led to greater inventory efficiencies, better maps and accurate estimation for small areas.

Keywords: *estimation, harmonization, interference, k-nearest neighbor, lidar, remote sensing.*

Introduction

Strategic national forest inventories (NFIs) are conducted by at least 40 countries representing 2.4 billion hectares of forest, more than half the forested area of the earth (Tomppo et al., 2010*a*, Preface). Thus, a comprehensive review of the history, advances and emerging issues for the entire international forest inventory community is an impossible task for a relatively short journal article. As a result, this review focuses on only selected areas: (1) a brief historical review with emphasis on recent issues that have shaped current approaches to implementing NFIs; (2) a summary of common, structural features of NFIs, albeit with a large diversity of operational implementations; (3) a brief review of international reporting requirements using NFI data with emphasis on approaches to harmonized estimation; (4) an overview of inventory estimation methods that can be enhanced with remotely sensed data; (5) an overview of nearest neighbors prediction methods; and (6) a brief overview of several emerging issues.

History

The history of forest inventories traces back to at least the end of the Middle Ages, when intensive use of forest resources led to wood shortages and forced users to begin forest planning (Loetsch & Haller, 1973; Gabler & Schadauer, 2007). These inventories focused on forest area and crude estimates of what is now characterized as growing stock volume. In addition, they were usually local in scope, they emphasized assessments of timber resources for specific purposes, and they were often conducted by timber users (Loetsch & Haller, 1973; Davis et al., 2001). However, such inventories were soon acknowledged to be inadequate for compiling national

Correspondence: R. E. McRoberts, Northern Research Station, US Department of Agriculture Forest Service, 1992 Folwell Avenue, St Paul, Minnesota, MN 55108, USA. E-mail: rmcroberts@fs.fed.us

(Received 12 March 2010; accepted 24 May 2010)

ISSN 0282-7581 print/ISSN 1651-1891 online © 2010 Taylor & Francis
DOI: 10.1080/02827581.2010.496739

forest information for purposes of formulating national forest policy; thus, NFIs were initiated. NFIs have different histories in different countries, but some type of forest information has been collected in both European and North American countries since the nineteenth century (von Berg, 1858; LaBau et al., 2007). Early NFIs were very much commodity driven with emphasis on acquiring information on forest area, volume and increment of growing stock, and the amount of timber. In addition, they addressed issues of age, size and species structures of forests, silvicultural status, and necessary cutting and silvicultural regimes (Ilvessalo, 1927). Their primary purpose was to provide information for forestry authorities, timber users and planners who developed national forest policies.

Multiple issues have emerged in recent decades to cause re-evaluation of the traditional role of forest inventories. First, as a result of declines in sensitive forest ecosystems in the 1980s, forest health monitoring programs were developed, some of which were integrated into NFIs. Second, the role of forests in providing non-wood goods and services such as wildlife habitat, recreational opportunities and contributions to water quality has received increased attention in recent years, particularly in urbanized societies. Third, human-induced habitat losses and tropical deforestation have accelerated the rate of species extinctions and led to international regimes such as the Convention on Biological Diversity (CBD, 2009) which, in turn, led to forest biodiversity monitoring programs in the early and mid-1990s. Fourth, processes such as the Ministerial Conference on the Protection of Forests in Europe (MCPFE, 2009) and the Montréal Process (2005) were established with goals of assessing forest sustainability criteria and indicators. Fifth, the combined effects of the use of fossil fuels, deforestation in the tropics and farming have produced increased levels of atmospheric carbon dioxide (CO_2) and other greenhouse gases. In response, many countries joined the United Nations Framework Convention on Climate Change (UNFCCC) to consider actions to reduce these effects and to cope with the expected global temperature increases (UNFCCC, 2009). Of particular importance is that forest is a land-use category whose resources can be managed to produce a positive effect on the greenhouse gas balance (Cienciala et al., 2008). The combined effects of all these issues have cast a new light on the role of forests and created new expectations for both forestry and forest inventories.

Importantly, the changing roles of forests have substantially altered demands for forest information. Because NFIs are the primary source of forest information for all these purposes, the scope of NFIs has broadened accordingly, resulting in the introduction of a wide variety of new variables and corresponding design modifications. Today, NFIs reflect a wider and more holistic perspective on global forest resources, their management and uses. This perspective is expressed via seven broad thematic elements of sustainable forest management: (1) extent of forest resources and their contribution to the global carbon cycle; (2) forest health and vitality; (3) forest biological diversity; (4) productive functions of forests; (5) protective functions of forests; (6) socioeconomic functions of forests; and (7) legal, policy and institutional frameworks related to forests (Tomppo et al., 2010*a*).

Structural features of national forest inventories

Two NFI approaches are used, one based on data acquired using statistical sampling, and the other based on data aggregated from stand-level management inventories. Although the evidence is now clear that the latter approach tends to underestimate forest area and growing stock volume (Tomppo et al., 2001; Adermann, 2010), systematic forest assessments based on statistical sampling methods began only in the twentieth century. The first sample-based NFIs were initiated in the Nordic countries: 1919 in Norway, 1921 in Finland and 1923 in Sweden. In the United States of America sample-based inventories were initiated in 1928. Today, the NFIs of most countries are sample based. A brief summary of the sampling design, plot configuration and inventory variable features of NFIs based on the detailed results reported by Tomppo et al. (2010*a*, Chapter 2) follows.

Sampling designs and plot configurations

NFIs have exhibited a progressive evolution towards statistical sampling techniques, with the majority of countries now using probability sampling designs (Tomppo et al., 2010*a*, Table 2.3). With these designs, each population unit has a positive, known probability of selection into the sample, and a randomization procedure is used to select the particular units for the sample. Probability sampling designs are of multiple forms. Most NFIs use systematic sampling based on two-dimensional grids whose starting points are randomly selected. Grid spacings range from 0.5 to 20 km, although most are in the range 1.4–4.0 km. Other countries use tessellated sampling designs in which large areas are tessellated with regular polygons and plot locations are randomly selected within the polygons. In addition, stratified sampling is often used when different sampling intensities are desired for different land uses and land covers, different climatic and topographical regions, or regions of different interest or priority. Strata are most often defined using maps constructed using a combination of political, biophysical and remotely sensed data, and forest attribute predictions such as proportion forest cover, forest type and volume.

NFIs now generally use detached sample plots of which all, or at least some, are permanent to facilitate estimation of change and trends as required for sustainability assessments. In addition, the plots are often configured in clusters for purposes of logistical efficiency. The number of plots per cluster is often four, although it may be as large as 18. The shapes and sizes of sample plots vary enormously among NFIs. The inventories of nearly all countries with boreal and temperate forests use circular plots with multiple concentric components for which smaller trees are measured only on the smaller components. However, square and rectangular plots are also used, and for some variables transect and Bitterlich sampling are used.

The diversity of sampling approaches and plot sizes and shapes has increased recently to accommodate the need for more diverse forest information. For example, transects may be established within large circular plots for assessing deadwood, and soil pits may be used to assess soil carbon. Thus, NFIs whose sole objectives previously related to estimation of timber resources have given way to NFIs that encompass a host of outputs including ground vegetation, deadwood, biodiversity and soil information, in addition to timber information.

Inventory cycles are typically either 5 years or 10 years. Traditionally, inventory cycles were achieved by conducting complete periodic inventories over 1–2 years and then repeating the inventories 5–10 years later. Increasingly, however, inventory cycles in Europe and North America are achieved by measuring 10–20% of plots each year. If the 10–20% of plots measured annually are systematically distributed over entire regions, then current estimates can be obtained in any year and the effects of catastrophic events such as hurricanes, wind and ice storms, and pest and disease outbreaks can be rapidly assessed.

Inventory variables

The two most common NFI variables for which population parameters are estimated are forest area and growing stock volume (Tomppo et al., 2010*a*, Tables 2.1 and 2.2) Priority is given to estimates of the current status of these parameters and change in these parameters between inventories. All countries assess forest area, and most base their definitions of forest on percentage canopy cover and minimum patch area, with many also using minimum area width and minimum tree height at maturity. Canopy cover thresholds range from 10% to 50%, with the smaller threshold being the most common. Methods for assessing forest area fall into four categories: (1) assessment of ground plots or ground sampling points; (2) interpretation of plot locations using aerial photographs; (3) counting forest and non-forest points on systematic grids superimposed on aerial photographs; and (4) assessment of maps. Assessment of ground plots is the most common method, is increasing in frequency of use and is probably the most accurate method.

National definitions of growing stock volume differ considerably with respect to multiple components. Nearly all countries base growing stock volume on tree diameter over bark. However, beyond this criterion, definitions of growing stock are quite diverse. For example, the height at which diameter is measured, often characterized as breast height, varies from 1.3 to 1.5 m, the minimum tree diameter at breast height (dbh) varies from 0.0 to 12.7 cm, and the minimum top height diameter varies from 0.0 to 10.1 cm for trees used to estimate volume. In addition, countries are split between those that do and those that do not include stump volume, branches and recently dead trees in growing stock volume. Most countries estimate volume for productive forest land, but others also estimate it for poorly productive forest land and other wooded land. Although volumes for these land classes are usually estimated separately, the definitions for the classes vary considerably.

International reporting and harmonization

Numerous international agreements require that countries report national estimates of parameters related to forest area, growing stock volume and their change. Parties to the UNFCCC (1992) are required to produce annual reports of greenhouse gas emissions and removals by sources and sinks. The CBD (2009) requires that countries identify and monitor components of biological diversity for purposes of conservation and sustainable use. Both the Montréal Process (2005) and the MCPFE (2009) require that member countries report on sustainability and biodiversity indicators.

Extensive and comprehensive data on the status of forests available for reporting under international agreements are provided by NFIs. Although NFIs share a primary objective of conducting forest resource assessments to assess the sustainable yield and management of forests, the previous section clearly documents that they do not assess common sets of variables or use common definitions of variables. Further, inventory sampling designs, plot configurations, measurement protocols and analytical methods vary considerably among countries. An effect of these differences is a lack of comparability in estimates submitted for international reporting by individual countries using their own definitions, thresholds and measurement protocols. Traub et al. (1997) reported that the pan-European estimate of forest area decreased by 6% when the United Kingdom's definition of forest was used and increased by 3% when Luxembourg's definition was used. Cienciala et al. (2008) reported that when using a definition proposed by the Food and Agriculture Organization (FAO) of the United Nations rather than the Finnish national definition, Finland's forest area estimate increased by 10.6%, its growing stock volume estimate increased by 2.7% and its annual volume increment estimate increased by 1.8%. If a minimum dbh of 10.4 cm were used for Finnish forests instead of 0.0 cm, the volume estimate would decrease by 14%, the volume increment estimate would decrease by 25% and the carbon sink estimate would decrease by 26%.

Comparability of inventory estimates submitted for international reporting may be achieved using two approaches, standardization and harmonization. However, because NFI features such as sampling designs and plot configurations for individual countries have been developed over time to accommodate their unique topographies, climates, forest types and commercial interests, standardization is often not a viable option. Harmonization acknowledges that individual countries have developed the unique features of their NFIs for specific purposes and are justified in their desire to maintain them and, therefore, focuses on developing methods for producing comparable estimates despite the lack of standardization.

The issue of harmonization has received increased international attention in recent years by the Intergovernmental Panel on Climate Change (IPCC), the Global Forest Resources Assessments (FRA) conducted by FAO, and the 2000 Temperate and Boreal Resource Assessment (TBFRA), as a part of FRA 2000. The most comprehensive effort directed toward harmonization of NFI estimates was conducted by COST Action E43, *Harmonization of National Forest Inventories in Europe: Techniques for Common Reporting* (COST, 2009; COST Action E43, 2008; Tomppo et al., 2010*a*). The approach to harmonization proposed by COST Action E43 entailed two steps: (1) construct common reference definitions; and (2) construct bridges that facilitate estimation in accordance with the reference definition using data collected using national definitions.

Vidal et al. (2008) described the process used by COST Action E43 to develop a reference definition for forest land. The most commonly used variables as indicated by the proportions of European countries that use them and the areas these countries represent are minimum area, minimum tree crown cover, minimum area width and minimum tree height (Table I). Thus, the proposed reference definition for forest land is based on these four variables. A reference definition must not only identify the relevant variables but also specify thresholds for these variables. For a tree crown cover threshold of 10%, data are currently available for European countries representing more than 65% of the total area and nearly 75% of their total forest area (Table II). Thus, the proposed reference definition for forest land includes a tree crown cover threshold of 10%. Similar analyses were used to select a threshold of 0.5 ha for minimum area, 5 m for minimum tree height at maturity and 20 m for minimum area width. Thus, the COST Action E43 reference definition for forest land was (Tomppo et al., 2010*a*):

> Forest is land spanning more than 0.5 ha with trees higher than 5 metres and with tree crown cover of at least 10%, or able to satisfy these thresholds in situ. For tree rows or shelterbelts, a minimum width of 20 m is required. It does not include land that is predominantly agricultural or urban land use.

This definition is very similar to the 2005 FRA definition (FAO, 2005).

Bridges for producing estimates in accordance with reference definitions using data collected using national definitions can take multiple forms. Crucial factors affecting construction of bridges are the variables and corresponding thresholds in the national definitions under which data are acquired. The nature of these data is the primary factor that distinguishes three kinds of bridge: *reductive*,

Table I. Variables used in national definitions of forest land.[a]

	Proportion[b]		
Variable	Countries	Area of countries	Forest area of countries
Minimum area	0.96	0.99	0.99
Minimum tree crown cover	0.81	0.69	0.56
Minimum width	0.74	0.66	0.54
Minimum tree height	0.59	0.47	0.42

Note: [a]McRoberts et al. (2009); [b]proportions for 27 European countries participating in COST Action E43.

Table II. Tree crown cover threshold used in national definitions of forest land.[a]

Tree crown cover threshold	Cumulative proportion[b] Country area	Forest area
0.05	0.12	0.12
0.10	0.67	0.73
0.20	0.79	0.81
0.30	0.84	0.87
0.50	0.94	0.95
Not used	0.06	0.05

Note: [a]McRoberts et al. (2009); [b] proportions for 27 European countries participating in COST Action E43.

expansive and *neutral* bridges. Reductive bridges are used when the national definition is broader in scope than the reference definition, in which case some of the national data can be simply excluded. When the scopes of the reference and national definitions are the same, neutral bridges are appropriate. Expansive bridges are necessary when the scope of the reference definition is broader than that of the national definition, in which case data are missing for estimation based on the reference definition and must be supplied via prediction, imputation or other methods. Examples of bridges for harmonization purposes are found in McRoberts et al. (2009) and Tomppo et al. (2010*a*).

Using remotely sensed data to enhance inventory sampling and estimation

Although remotely sensed data are increasingly used to enhance inventories, they cannot completely replace ground sample data. The expansion of inventories to address issues such as sustainability and biodiversity requires information on variables such as deadwood, lichens and soil carbon. In addition, remotely sensed data are not currently sufficient for producing species-level estimates, particularly in tropical countries such as Brazil, which is estimated to have more than 11,000 tree species that reach diameters of 10 cm and more than 3000 species of which there are more than a million individuals (Hubbell et al., 2008). Nevertheless, the statistical sophistication of methods related to the use of remotely sensed data has increased substantially in recent years (McRoberts & Tomppo, 2006). The two most important conclusions, as described and discussed in the following sections, are that remotely sensed data have now been incorporated into operational forest inventories, and that estimates obtained from remote sensing-based maps can now be expressed in forms similar to sample-based estimates.

Sampling

The efficiencies of inventory sampling designs have been evaluated using maps constructed using satellite imagery that depict forest attributes such as volume by tree species. A map can be considered a model of forests and land use which can then be used to assess measurement and travel costs using Geographical Information System (GIS) techniques. This approach has been used operationally to plan sampling designs since 1992 for the eight inventories of northern Finland conducted as part of the Finnish NFI and subsequently for the entire country. Standard errors of population parameters were estimated using map predictions based on satellite image and field data, as well as sampling simulations (Tomppo, 2009). In the very north of Finland, where the road network is sparse and travel costs are large, a double sampling for stratification design was used. A similar approach focusing on optimal sample allocation (Cochran, 1977) was used in a collaborative effort among the Tanzanian government, FAO and the Finnish Forest Research Institute (Tomppo et al., 2010*b*) to plan the sampling design for the new Tanzanian NFI. Advantages of this approach are that greater sampling intensity can be allocated to areas with greater variation in forest variables and resources can be allocated using sound statistical principles.

Expansion factors

When using plot observations as the basis for estimates of inventory variables for subpopulations such as municipalities, counties and provinces or states, the per unit area observation for each plot must be multiplied or expanded by the area the plot represents to obtain an estimate of the total for the subpopulation. For example, if 100 trees are observed on a 0.5 ha plot, then the observation converts to a per unit area observation of 200 trees ha^{-1}. For a sampling intensity of one plot per 1000 ha, the plot expansion factor is 1000, in which case the total number of trees represented by the plot observation is 200,000 trees. More accurate area estimates may be obtained when plot expansion factors are derived from landscape features. For example, if the plot whose observation converts to 200 trees ha^{-1} is the only plot on a particular soil type and there are 1500 ha of this soil type in the population, then a plot expansion factor of 1500 would produce a better estimate of the total number of trees than would the 1000 expansion factor.

The Finnish NFI has developed an innovative method for estimating expansion factors using remotely sensed data that can be used to improve small area estimates (Tomppo & Halme, 2004). Whereas NFI plot data are routinely used to estimate population parameters for areas on the order of 100,000s of hectares, a typical small area such as a municipality is approximately 10,000 ha and lacks sufficient numbers of plots to produce acceptably precise estimates using only plot data (Tomppo, 1996; Tomppo et al., 2008). The approach relies on a synthetic estimator that uses plot data for a larger region such as a Landsat scene to produce estimates for a smaller area such as a municipality. For application to the small area, the expansion factor, c_j, for the jth plot in the larger region is calculated using the k-nearest neighbors (kNN) technique as:

$$c_j = \sum_{i=1}^{N} w_{j,i},$$

where $w_{j,i}$ is a measure of the similarity of the ith pixel in the small area to the pixel containing the center of the jth plot in the larger region. Thus, the plot expansion factor is a function of the proportion of pixels in the small area that are similar to the pixel containing the center of the plot in the larger area (Tomppo, 1996).

Stratified estimation

NFIs that use probability-based sampling designs typically also use approaches to inference that are characterized as design based. Properties of design-based estimators derive from the probabilities of selection of population units into the sample. Design-based inference is based on two assumptions: a probability sample and a constant value, apart from measurement error, for each population unit. Estimators for population parameters are derived to correspond to the sampling designs and are typically unbiased, although the estimate obtained with any particular sample may deviate considerably from the true value of the population parameter.

Stratified estimators are a subclass of design-based estimators that are necessary for cross-strata estimation when sampling methods and intensities differ by strata. Stratified estimation can also be used to increase the precision of estimates and thereby enhance inferences, regardless of whether stratified sampling is used. The essence of stratified estimation is to assign sampled population units to homogeneous classes or strata, calculate within stratum sample means and variances, and then calculate a weighted average of the within stratum estimates where the weights are proportional to the stratum sizes. If the stratification is effective, the stratified variance estimate will be less than the variance estimate based on an assumption of simple random sampling. Stratified estimates are calculated using estimators described by Cochran (1977).

Historically, stratified estimation for inventory applications was implemented in northern Finland, the USA and elsewhere using a modification of a double sampling for stratification approach (Bickford, 1952; Poso, 1972; Poso & Kujala, 1978). In the first phase, an initial, dense, systematic grid of photoplots was laid over aerial photographs of an inventory area. The photoplots were interpreted with respect to the land use and the size, density and species of trees on forest land and grouped into homogeneous classes or strata. The relative areal extent of each stratum was estimated as the proportion of first phase photoplots assigned to the stratum. The second phase sample was drawn from the photoplots in the strata using either systematic or random sampling. Field plots were established with centers at the centers of the second phase photoplots and were measured by field crews. The data for the two phases were combined using stratified estimators (Cochran, 1977).

In recent years, satellite imagery has replaced aerial photography as a source of information for constructing strata in many countries (McRoberts et al., 2002*a*, *b*, 2006; Nilsson et al., 2005). With these approaches, the satellite imagery is often classified with respect to selected forest attributes, and the strata are either the classes or aggregations of the classes. The weight for each stratum is easily calculated using GIS techniques as the proportion of pixels with centers in the stratum. McRoberts et al. (2006) proposed an approach to stratification in which strata are based on maps of classes of proportion forest cover; e.g. 0.000–0.099, 0.100–0.499, 0.500–0.899 and 0.900–1.000. A similar approach is now in operational use in the USA. The utility of stratifications is often assessed using relative efficiency, defined as:

$$\mathrm{RE} = \frac{\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{SRS}})}{\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{Str}})},$$

where $\hat{\mu}_{\mathrm{SRS}}$ is the estimate obtained under the assumption of simple random sampling and $\hat{\mu}_{\mathrm{Str}}$ is the estimate obtained using the stratified estimator. For forest area and volume in the north central USA, $\mathrm{RE} = 5.87$ and $\mathrm{RE} = 2.71$, respectively, were obtained. RE is equivalent to the factor by which sample sizes would have to be increased to achieve the same precision using simple random sampling estimators without stratification as is achieved using the stratified estimator. Thus, $\mathrm{RE} = 2.0$ indicates that in the absence of stratification, the sample size would have to be doubled to achieve the same precision, a prohibitively expensive proposition for most NFIs.

Model-assisted estimation

Model-assisted estimators are a second subclass of design-based estimators in which a model is used to describe the population, particularly the point scatter (Särndal et al., 1992). Relative to stratified estimators, model-assisted estimators make greater use of the same ancillary data because they use the complete range of the data rather than grouping them into classes. Although model-assisted estimators use model predictions for population units, they rely on the probability sample for validity. Model-assisted estimators can take multiple forms including the familiar ratio, regression and difference estimators.

Särndal et al. (1992) defines the model-assisted difference estimator which, for equal probability samples, can be expressed as:

$$\hat{\mu}_{\mathrm{Dif}} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i - \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i),$$

with variance that can be approximated as:

$$\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{Dif}}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2,$$

where y_i is an observation, $\hat{y}_i$ is a model prediction, n is the sample size, and N is the population size. The primary advantage of model-assisted estimators, in general, is that they capitalize on the relationship between the sample observations and their model predictions to reduce the variance of the population parameter estimate. A particularly appealing feature of the difference estimator is that any model or prediction approach may be used to obtain the population unit predictions. Despite their flexibility and the reductions in variances they produce, model-assisted estimators suffer the effects of small sample sizes when used for small area estimation.

Recent applications of model-assisted techniques include Opsomer et al. (2007), who used a model-assisted approach with non-parametric regression to estimate multiple forest attributes in the western USA from forest inventory and Landsat data. In addition, Baffetta et al. (2009) estimated timber volume in Italy using the model-assisted difference estimator with forest inventory and Landsat data and the *k*NN technique; Andersen et al. (2009) used a model-assisted approach with a regression model to estimate biomass in Alaska, USA, from forest inventory and lidar (light detection and ranging) data; and McRoberts (2010) used the model-assisted difference estimator with a logistic regression model to estimate proportion forest in the USA from forest inventory and Landsat data.

Model-based estimation

The assumptions underlying model-based inference differ considerably from those underlying probability-based inference. First, the observation for a population unit is a random variable whose value is considered a realization from a distribution of possible values, rather than a constant as is the case for probability-based inference. Second, the basis for a model-based inference is the model, not the probabilistic nature of the sample as is the case for probability-based inference. McRoberts (2010) provides a comparison of design-based and model-based approaches.

With model-based inference, a model,

$$y_i = f(\mathbf{X}_i; \boldsymbol{\beta}) + \varepsilon_i,$$

is fit to the sample data, where $\mathbf{X}_i$ is the vector of ancillary variables for the ith population unit, $\boldsymbol{\beta}$ is a vector of parameters to be estimated, and ε_i is the random residual. The model prediction, $f(\mathbf{X}_i; \hat{\boldsymbol{\beta}})$, for the ith population unit serves as an estimate of the mean of the distribution of possible values for the unit; i.e. $\hat{\mu}_i = f(\mathbf{X}_i; \hat{\boldsymbol{\beta}})$. Model-based estimators for the population parameter, μ, are:

$$\hat{\mu}_{\mathrm{Mod}} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mu}_i$$

and

$$\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{Mod}}) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\hat{\mathrm{Cov}}(\hat{\mu}_i, \hat{\mu}_j),$$

where calculation of $\hat{\mathrm{Cov}}(\hat{\mu}_i, \hat{\mu}_j)$ depends on the particular approach to prediction used. The primary challenges with model-based approaches are computational intensity, lack of fit assessment and variance estimation.

McRoberts et al. (2006, 2007, 2010) used model-based approaches with the *k*NN technique with Landsat data to estimate proportion forest area, basal area, volume and stem density. Both Nelson et al. (2004) and Ståhl et al. (2010) used model-based approaches with lidar data to estimate aboveground biomass.

Nearest neighbors techniques

Within the international forest inventory community, nearest neighbors techniques have emerged as a class of useful and intuitive methods for predicting forest attributes from climatic, topographic, remotely sensed and other ancillary variables. With nearest neighbors techniques, population unit predictions are calculated as linear combinations of observations for the population units in the sample that are nearest or most similar in a space of ancillary variables to the unit requiring a prediction. Nearest neighbors techniques are appealing for multiple reasons: (1) they can be used for both univariate and multivariate prediction; (2) they are non-parametric in the sense that no assumptions regarding the distributions of response or ancillary variables are necessary; (3) they are synthetic in the sense that they can readily use information external to the geographical area for which an estimate is sought; (4) they are useful for map construction and small area estimation; (5) they can be used with a wide variety of data sets; and (6) when used to classify remote sensing data, they lend themselves better to statistical analysis than many other classification techniques.

The primary forest inventory applications of nearest neighbors techniques are in three areas: (1) imputing missing values for forest inventory and monitoring databases (Eskelson et al., 2009); (2) enhancing design-based inference; and (3) mapping, small area estimation and model-based inference. The Finnish NFI has used the *k*NN technique operationally with satellite image and NFI field plot data for small area estimation since 1990.

More recently, nearest neighbors techniques have gained considerable popularity and utility for use with forest inventory and satellite image data to estimate forest attribute response variables. Applications have truly been international, with published reports for Austria (Koukal et al., 2007), Canada (LeMay & Temesgen, 2005), China (Tomppo et al., 2001), Costa Rico (Thessler et al., 2008), Finland (Tomppo & Halme, 2004; Tomppo, 2006), Ireland (McInerney & Nieuwenhuis, 2009), Italy (Chirici et al., 2008; Baffetta et al., 2009), Japan (Kajisa et al., 2008), New Zealand (Tomppo et al., 1999), Norway (Gjertsen, 2007), Siberia (Fuchs et al., 2009), Sweden (Nilsson, 1997; Holmström & Fransson, 2003; Reese et al., 2003), the USA (Moeur & Stage, 1995; McRoberts et al., 2007) and Zambia (Maltamo & Eerikäinen, 2001). A bibliography of published peer-reviewed paper on nearest neighbors is available at: http://blue.for.msu.edu/NAFIS/biblio.html.

For notational purposes, let $\mathbf{Y}$ denote a possibly multivariate vector of response variables with observations for a sample of size n from a finite population of size N, and let $\mathbf{X}$ denote a vector of ancillary variables with observations for all population units. In the terminology of nearest neighbors techniques, the set of population units for which observations of both response and ancillary variables are available is designated the reference set; the set of population units for which predictions of response variables are desired is designated the target set; and the space defined by the ancillary variables, $\mathbf{X}$, is designated the feature space. All elements of both the reference and target sets are assumed to have a complete set of observations for all feature space variables. For continuous response variables, the *k*NN prediction for the *i*th target set element is:

$$\tilde{y}_i = \frac{1}{W_i} \sum_{j=1}^{k} w_{ij} y_j^i,$$

where $\{y_j^i, j = 1, 2, \ldots, k\}$ is the set of response variable observations for the k reference set elements that are nearest to the ith target set element in feature space with respect to a distance metric, d, and w_{ij} is the weight assigned to the jth nearest neighbor with $W_i = \Sigma_{j=1}^{k} w_{ij}$. For categorical variables the predicted class of the ith target set element is the most heavily weighted class among the k nearest neighbors. Frequent choices for the weights are $w_{ij} = d_{ij}^t$, where $t \in [-2, 0]$. Many distance measures may be expressed in matrix form as:

$$d_{ij} = (\mathbf{X}_i - \mathbf{X}_j)' \mathbf{M} (\mathbf{X}_i - \mathbf{X}_j),$$

where i denotes a target set element for which a prediction is sought, j denotes a reference set element, $\mathbf{X}_i$ and $\mathbf{X}_j$ are the vectors of observations of feature space variables for the ith and jth elements, respectively, and $\mathbf{M}$ is a square matrix. Popular choices for $\mathbf{M}$ include the identity matrix which results in Euclidean distance, a non-identity diagonal matrix which results in weighted Euclidean distance, the inverse of the covariance matrix of the feature space variables which results in Mahalanobis distance (Kendall & Buckland, 1982). Other choices for $\mathbf{M}$ are based on canonical correlation analysis (Moeur & Stage, 1995; Maltamo & Eerikäinen, 2001; Temesgen et al., 2003; LeMay et al., 2008) and canonical correspondence analysis (Ohmann & Gregory, 2002; Ohmann et al., 2007; Pierce et al., 2009). Chirici et al. (2008) and Hudak et al. (2008) evaluated additional distance metrics.

Recent investigations of nearest neighbors techniques have shifted from simple descriptions of applications to more foundational work on efficiency and inference. McRoberts (2009) reported diagnostic tools for evaluating and enhancing nearest neighbors prediction of continuous, univariate response variables. Finley et al. (2006) and Finley and McRoberts (2008) investigated enhanced search algorithms for identifying nearest neighbors. Tomppo and Halme (2004) and Tomppo et al. (2009) used a genetic algorithm approach to optimize the weights for ancillary variables in the distance metric. However, the full potential of nearest neighbors techniques cannot be realized until they can be used for construction of valid statistical inferences, a process that requires variance estimates. Kim and Tomppo (2006) reported a method for estimating the uncertainty of predictions using a variogram approach; McRoberts et al. (2007) derived a *k*NN variance estimator for areal means from the conceptual assumptions underlying *k*NN estimation; and Magnussen et al. (2010*b*) investigated an estimator of mean square error.

Nearest neighbors techniques have been used with satellite imagery to support and enhance most inventory estimation methods. McRoberts et al. (2002*a*) used the *k*NN technique to construct strata for use with stratified estimation. The Finnish NFI uses the *k*NN technique operationally with satellite imagery to calculate expansion factors (Tomppo et al., 2008). Baffetta et al. (2009) described an application of the *k*NN technique for use with the model-assisted difference estimator. McRoberts et al. (2007) used the *k*NN technique for a model-based approach to inference for small area applications.

Emerging issues

Climate change

The United Nations-REDD (Reducing Emissions from Deforestation and forest Degradation) program is a collaborative partnership among FAO, the United Nations Development Program and the United Nations Environment Program (UN-REDD, 2009). UN-REDD was launched in September 2008 to create financial value for carbon stored in forests, to offer incentives for developing countries to reduce forest carbon emissions and to invest in low-carbon paths to sustainable development. The UN-REDD program supports countries in the development of cost-effective and robust measurement, reporting and verification (MRV) systems based on good science and available technology.

Because of the similarity between MRV programs and NFIs, some REDD countries are initiating new NFIs to support their MRVs. However, factors unique to these developing, often tropical, countries may require NFI features that deviate considerably from features of NFIs countries with boreal and temperate forests. First, although forest information requirements and expectations for developing countries are great, national financial resources are often limited. In addition, whereas information produced by NFIs in developed countries is used for strategic, national policy making, in developing countries forest policy decisions are often made locally, meaning that information for smaller areas is necessary. Thus, optimization of sampling designs using ancillary data such as satellite imagery is crucial, although satellite image-based classification may be more difficult in tropical countries because of the large number of species. Second, the combination of limited financial resources and large proportions of forest lands that are remote and difficult to access means that the sampling designs will be likely to feature strata with considerably different sampling intensities. Further, selection of the sampling intensities within strata may not be a function of desired precision as much as a function of an acceptable distribution of resources. Third, because of limited financial resources and limited access to remote regions, greater reliance on remotely sensed data may be expected. However, for tropical countries, acquisition of cloud-free imagery will be a challenge. In addition, methods for assessing the accuracy of satellite image-based maps and deriving estimates and inferences from them are necessary (McRoberts, 2010). Fourth, whereas radii as great as 12 m for circular NFI plots may be appropriate for less dense boreal and temperate forests, they may be too large for more dense tropical forests. For example, stem densities in tropical forests may make it impossible to determine accurate distances from plot centers to trees on the periphery of large radii plots. As a result, a larger number of smaller circular plots may be necessary or, as recommended by FAO, narrow, rectangular plots may be necessary (Saket et al., 2002). However, the latter plots are not without challenges such as dealing with four corner control points versus a single plot center control point for circular plots. In addition, the greater perimeter of rectangular plots for the same plot area means more boundary trees will have to be assessed in terms of whether they are or are not on the plot. Finally, facilitation of harmonized international reporting should be considered when defining forest and growing stock volume and when selecting minimum thresholds for variables used in these definitions such as patch size, crown cover, height at maturity, breast height, diameter at breast height and top height diameter. In particular, thresholds such as patch size of 0.1 ha and crown cover of 10% for use in the definition of boreal and temperate forest may be inappropriate or difficult to accommodate for tropical forests.

Lidar

Most remotely sensed data used by NFIs have been from passive sensors such as Landsat and SPOT that respond only to reflected sunlight. Data are now becoming widely available from active systems such as lidar that measure the elapsed time between emission of a short duration pulse of light and detection of the reflected pulse back at the sensor. The elapsed time is converted to distance which, for forestry applications, provides information that can be used to assess forest canopy height and vertical structure.

For inventory applications, data from small-footprint, discrete-return lidar systems with a ground footprint size of approximately 0.1–2.0 m have been the most intensively investigated. Such systems typically record two to five discrete echoes of each emitted pulse, often with one or more echoes reflected from the canopy and a last echo reflected from the ground. The first task with lidar data is usually to determine the terrain surface and then to estimate the relative height of the canopy echoes that are reflections from vertical positions within and on the outer surface of the forest canopy. When aggregated over a sample plot or forest stand, the collection of heights estimated from canopy echoes resembles the distribution of biological material in various vertical height layers. This information can be combined with field observations to construct statistical models for predicting height, growing stock volume, total biomass and other biophysical parameters (Nelson et al., 1988; Næsset 1997*a*, *b*).

Small-footprint applications date back to approximately 1980 (Solodukhin et al., 1977; Nelson et al., 1984) and focus on estimating cross-sectional areas and heights of individual tree canopies using profiling systems. Lidar data collected from profiling systems are only collected along a narrow strip underneath the aircraft and are used to produce two-dimensional profiles of heights that can be used to predict biophysical parameters (Maclean & Krabill, 1986; Nelson et al., 1988).

With the introduction of small-footprint scanning lidar systems, also known as airborne laser scanning (ALS), for forest research in approximately 1995, lidar-based height information became available as three-dimensional, wall-to-wall data. Such data can be used either to segment and characterize individual tree crowns (Hyyppä & Inkinen, 1999; Persson et al., 2002) or to estimate areal means for tree height, tree stem volume, total biomass, tree size distribution and other biophysical parameters (Næsset, 1997*a*, 1997*b*; Gobakken & Næsset, 2004). With the areal approach, field sample plot data and the lidar data are combined to construct models that are then used to predict biophysical parameters of interest for individual cells of a grid that covers the entire area of interest (Næsset, 2002). Cell sizes are equal to the field plot sizes, and cell predictions can be aggregated to produce estimates for stands or areal units of interest. Whereas the single-tree approach requires lidar measurements at densities greater than 2–5 pulses m^{-2}, the areal approach can provide accurate estimates with pulse densities of only approximately 0.5–1.0 pulses m^{-2} (Gobakken & Næsset, 2008; Magnussen et al., 2010*a*). Non-parametric methods such as the *k*NN technique have been used successfully for both tree-level (Maltamo et al., 2009) and areal applications (Maltamo et al., 2006; Packalén & Maltamo, 2007).

To date, most successful lidar-based applications have been for management inventories, for which they have been shown to be more cost efficient than conventional methods that use a combination of aerial photography and/or field data (Eid et al., 2004). Fully operational or semi-operational management inventories using mainly low-pulse lidar data have been conducted in the Nordic countries, North America, southern Europe and Australia in support of commercial thinning and management (e.g. Næsset, 2004, 2007).

For operational use in NFIs, airborne lidar data may play different roles including partial replacement of tree-level plot measurements or as ancillary data for estimation of parameters such as total biomass over larger areas. Full integration of lidar into NFIs currently faces multiple challenges: (1) tree heights are always underestimated, although corrections can be made; (2) errors are introduced as the result of failure to detect suppressed trees or prediction of non-existing trees; (3) tree species can be difficult to identify; and (4) estimates must be provided by timber assortments such as saw timber and pulp wood, not just total tree stem volume. For the Nordic countries, where many of the lidar studies have been conducted, progress has been made on species identification. Holmgren et al. (2008) showed that fusion of aerial imagery and lidar data can improve species classification for individual trees, and Packalén and Maltamo (2007) showed that the same fusion can be used to estimate species specific parameters in areal estimation. Further, dense lidar data without supporting aerial imagery can produce accurate species predictions for individual trees (Vauhkonen et al., 2009). Breidenbach et al. (2010) proposed methods for statistically correcting errors occurring with individual tree crown segmentation algorithms. However, additional progress is necessary for regions that have more species than in the Nordic countries and that have extensive mixed species and uneven-aged stands. For management inventories, timber assortments have been better estimated with lidar than with conventional methods using field assessments (Peuhkurinen et al., 2007; Packalén & Maltamo, 2008), but positive results must be demonstrated for NFI applications.

Finally, wall-to-wall airborne lidar coverage is costly if estimates are sought only for large regions. Thus, applications using samples of lidar data collected along strips, rather than for entire areas, are under development (Andersen et al., 2009; Gregoire et al., 2010; Ståhl et al., 2010). However, efficient methods for producing valid statistical areal inferences for inventory population parameters based on lidar data require further investigation. A more extended discussion of the use of lidar for forestry applications is presented in McRoberts et al. (2010).

Summary

NFIs have had a long history, although their major features such as statistical sampling and estimation date only to the early years of the twentieth century. In addition, NFIs have been at least partially reshaped in recent years as a result of concern for issues such as acid deposition, biodiversity and forest sustainability, increased demand for and use of forest data, and international reporting requirements. Technological advances, particularly the widespread availability of remotely sensed data, have led to greater inventory efficiencies, map construction and accurate estimation for small areas. Other technological advances such as lidar that are now on the horizon have the potential to lead to even greater efficiencies and even more useful spatial products and estimates. Finally, the current concern for global climate change and recognition of the ecological necessity of preserving tropical forests has added a new impetus for initiating NFIs in developing, southern hemisphere countries. Unique challenges will emerge; in particular, although general inventory principles will remain unchanged, the sampling designs, plot configurations and measurement protocols that have become familiar to inventory experts in northern hemisphere countries with mostly boreal and temperate forests will require modification before they can be applied in countries with tropical forests.

Acknowledgements

The authors acknowledge the valuable comments on the lidar section received from Johannes Breidenbach and Hans Ole Ørka of the Norwegian University of Life Sciences and Matti Maltamo of the University of Joensuu. No external funding was received.

References

Adermann, V. (2010). Country report for Estonia. In E. Tomppo, T. Gschwantner, M. Lawrence, & R. E. McRoberts (Eds.), *National forest inventories: Pathways for common reporting* (pp. 171–184). Heidelberg: Springer.

Andersen, H.-E., Barrett, T., Winterberger, K., Strunk, J. & Temesgen, H. (2009). Estimating forest biomass on the western lowlands of the Kenai Peninsula of Alaska using airborne lidar and field plot data in a model-assisted sampling design. *Proceedings of the IUFRO Division 4 Conference, Extending forest inventory and monitoring over time and space*, May 19–22, 2009, Quebec City, Canada. Retrieved February 2010, from http://blue.for.msu.edu/meeting/proceed.php

Baffetta, F., Fattorini, L., Franceschi, S. & Corona, P. (2009). Design-based approach to *k*-nearest neighbours technique for coupling field and remotely sensed data in forest surveys. *Remote Sensing of Environment*, *11*, 463–475.

Bickford, C. A. (1952). The sampling design used in the forest survey of the northeast. *Journal of Forestry*, *50*, 290–293.

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T. & Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment*, *114*, 911–924.

CBD (2009). *The Convention on Biological Diversity.* Retrieved January 2010, from http://www.cbd.int/

Chirici, G., Barbati, A., Coronia, P., Marchetti, M., Travaglini, D., Maselli, F. & Bertini, R. (2008). Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Remote Sensing of Environment*, *112*, 2688–2700.

Cienciala, E., Tomppo, E., Snorrason, A., Broadmeadow, M., Colin, A., Dunger, K., et al. (2008). Preparing emission reporting from forests: Use of national forest inventories in European countries. *Silva Fennica*, *42*, 73–88.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: Wiley.

COST. (2009). European Cooperation in Science and Technology. Retrieved January 2010, from http://www.cost.esf.org/

COST Action E43. (2008). *Harmonisation of National Forest Inventories in Europe: Techniques for Common Reporting*. Retrieved February 2010, from http://www.metla.fi/eu/cost/e43/

Davis, L. S., Johnson, K. N., Bettinger, P. & Howard, T. E. (2001). *Forest management to sustain ecological, economic, and social values* (4th ed.). Illinois: Waveland Press.

Eid, T., Gobakken, T. & Næsset, E. (2004). Comparing stand inventories based on photo interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research*, *19*, 512–523.

Eskelson, B. N. I., Temesgen, H., LeMay, V., Barrett, T. N., Crookston, N. L. & Hudak, A. T. (2009). The role of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research*, *24*, 235–246.

FAO (2005). *Global Forest Resources Assessment 2005. Progress towards sustainable forest management* (FAO Forestry Paper 147). Rome: Food and Agriculture Organization. Retrieved May 2010, from http://www.fao.org

Finley, A. O. & McRoberts, R. E. (2008). Efficient k-nearest neighbor searches for multi-source forest attribute mapping. *Remote Sensing of Environment*, *112*, 2203–2211.

Finley, A. O., McRoberts, R. E. & Ek, A. R. (2006). Applying an efficient *k*-nearest neighbor search to forest attribute imputation. *Forest Science*, *52*, 130–135.

Fuchs, H., Magdon, P., Kleinn, C. & Flessa, H. (2009). Estimating aboveground carbon in a catchment of the Siberian tundra: combining satellite imagery and forest inventory. *Remote Sensing of Environment*, *113*, 518–531.

Gabler, K. & Schadauer, K. (2007). Some approaches and designs of sample-based national forest inventories. *Austrian Journal of Forest Science*, *124*, 105–133.

Gjertsen, A. K. (2007). Accuracy of forest mapping based on Landsat TM data and a *k*NN-based method. *Remote Sensing of Environment*, *110*, 420–430.

Gobakken, T. & Næsset, E. (2004). Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research*, *19*, 529–542.

Gobakken, T. & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, *38*, 1095–1109.

Gregoire, T. G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R. & Holm, S. (2010). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark county, Norway. *Canadian Journal of Forest Research* (in press).

Holmgren, J., Persson, Å. & Söderman, U. (2008). Species identification of individual trees by combining high resolution LIDAR data with multi-spectral images. *International Journal of Remote Sensing*, *29*, 1537–1552.

Holmström, H. & Fransson, J. E. S. (2003). Combining remotely sensed optical and radar data in *k*-NN estimation of forest variables. *Forest Science*, *4*, 409–418.

Hubbell, S. P., He, F., Condit, R., Borda-de-Agua, L., Kellner, J. & ter Steege, H. (2008). How many tree species are there in the Amazon and how many of them will go extinct? *Proceedings of the National Academy of Sciences*, *105*, 11498–11504.

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E. & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment, 112, 2232–2245. [Corrigendum. *Remote Sensing of Environment*, *113*, 289–290.]

Hyyppä, J. & Inkinen, M. (1999). Detecting and estimating attributes for single trees using laser scanner. *Photogrammetric Journal of Finland*, *16*, 27–42.

Ilvessalo, Y. (1927). The Forests of Suomi (Finland). Results of the general survey of the forests of the country carried out during the years 1921–1924. *Communicationes Ex Instituto Quaestionum Forestalium Finlandiae*. Editae 11.

Kajisa, T., Murakami, T., Mizoue, N., Kitahara, F. & Yoshida, S. (2008). Estimation of stand volumes using the k-nearest neighbors method in Kyushu, Japan. *Journal of Forest Research*, *13*, 249–254.

Kendall, M. G. & Buckland, W. R. (1982). *A dictionary of statistical terms* (4th ed.). New York: Longman Group.

Kim, H.-J. & Tomppo, E. (2006). Model-based prediction error uncertainty estimation for *k*-nn method. *Remote Sensing of Environment*, *104*, 257–263.

Koukal, T., Suppan, F. & Schneider, W. (2007). The impact of relative radiometric calibration on the accuracy of *k*NN-predictions of forest attributes. *Remote Sensing of Environment*, *110*, 431–437.

LaBau, V. J., Bones, J. T., Kingsley, N. P., Lund, H. G. & Smith, W. B. (2007). *A history of the forest survey in the United States: 1830–2004*. (FS-877). Washington, DC: Forest Service, US Department of Agriculture.

LeMay, V. M., Maedel, J. & Coops, N. C. (2008). Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Remote Sensing of Environment*, *112*, 2578–2591.

LeMay, V. & Temesgen, H. (2005). Comparison of nearest neighbor methods for estimating basal area and stems per ha using aerial auxiliary variables. *Forest Science*, *51*, 109–119.

Loetsch, F. & Haller, K. E. (1973). *Forest inventory*, Vol. I. München: BLV Verlagsgesellschaft.

Maclean, G. A. & Krabill, W. B. (1986). Gross-merchantable timber volume estimation using airborne lidar system. *Canadian Journal of Remote Sensing*, *12*, 7–18.

McInerney, D. & Nieuwenhuis, M. (2009). A comparative analysis of *k*-NN and decision tree methods for the Irish National Forest Inventory. *International Journal of Remote Sensing*, *30*, 4937–4956.

McRoberts, R. E. (2009). Diagnostic tools for nearest neighbors techniques when used with satellite imagery. *Remote Sensing of Environment*, *113*, 489–499.

McRoberts, R. E. (2010). Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sensing of Environment*, *114*, 1017–1025.

McRoberts, R. E., Cohen, W., Stehman, S., Næsset, E. & Tomppo, E. (2010). Remote sensing-based maps of forest attributes. *Scandinavian Journal of Forest Research*, *25*, 340–367.

McRoberts, R. E., Holden, G. R., Nelson, M. D., Liknes, G. C. & Gormanson, D. D. (2006). Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the USDA Forest Service. *Canadian Journal of Forest Research*, *36*, 2968–2980.

McRoberts, R. E., Nelson, M. D. & Wendt, D. G. (2002*a*). Stratified estimation of forest area using satellite imagery, inventory data, and the *k*-nearest neighbors technique. *Remote Sensing of Environment*, *82*, 457–468.

McRoberts, R. E. & Tomppo, E. O. (2007). Remote sensing support for forest inventories. *Remote Sensing of Environment*, *110*, 412–419.

McRoberts, R. E., Tomppo, E. O., Finley, A. O. & Heikkinen, J. (2007). Estimating areal means and variances of forest attributes using the *k*-nearest neighbors technique and satellite imagery. *Remote Sensing of Environment*, *111*, 466–480.

McRoberts, R. E., Tomppo, E. O., Schadauer, K., Vidal, C., Ståhl, G., Chirici, G., et al. (2009). Harmonizing national forest inventories. *Journal of Forestry*, *107*, 179–187.

McRoberts, R. E., Wendt, D. G., Nelson, M. D. & Hansen, M. D. (2002*b*). Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, *81*, 36–44.

Magnussen, S., Næsset, E. & Gobakken, T. (2010*a*). Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sensing of Environment*, *114*, 700–712.

Magnussen, S., Tomppo, E. & McRoberts, R. E. (2010*b*). A model-assisted k-nearest neighbour approach to remove extrapolation bias. *Scandinavian Journal of Forest Research*, *25*, 174–184.

Maltamo, M. & Eerikäinen, K. (2001). The most similar neighbour reference in the yield prediction of Pinus kesiya stands in Zambia. *Silva Fennica*, *35*, 437–451.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A. & Kangas, J. (2006). Non-parametric estimation of stem volume using laser scanning, aerial photography and stand register data. *Canadian Journal of Forest Research*, *36*, 426–436.

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P. & Tokola, T. (2009). Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. *Silva Fennica*, *43*, 507–521.

MCPFE (2009). *Ministerial Conference on the Protection of Forests in Europe*. Retrieved January 2010, from http://www.mcpfe.org/
Moeur, M. & Stage, A. R. (1995). Most similar neighbor--An improved sampling inference procedure for natural resource planning. *Forest Science*, *41*, 337–359.
Montréal Process (2005). *Montréal Process*. Retrieved January 2010, from http://www.rinya.maff.go.jp/mpci/
Næsset, E. (1997*a*). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*, 49–56.
Næsset, E. (1997*b*). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, *61*, 246–253.
Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*, 88–99.
Næsset, E. (2004). Accuracy of forest inventory using airborne laser-scanning: evaluating the first Nordic full-scale operational project. *Scandinavian Journal of Forest Research*, *19*, 554–557.
Næsset, E. (2007). Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, *22*, 433–442.
Nelson, R., Krabill, W. & Maclean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment.*, *15*, 201–212.
Nelson, R., Krabill, W. & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*, 247–267.
Nelson, R., Short, A. & Valenti, M. (2004). Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. *Scandinavian Journal of Forest Research*, *19*, 500–511.
Nilsson, M. (1997). *Estimation of forest variables using satellite image data and airborne lidar*. PhD thesis, Department of Forest Resource Management and Geomatics Swedish University of Agricultural Sciences. *Acta Universitatis Agriculturae Sueciae Silvestria*, 17.
Nilsson, M., Holm, S., Reese, H., Wallerman, J. & Engberg, J. (2005). Improved forest statistics from the Swedish National Forest Inventory by combining field data and optical satellite data using post-stratification. In H. Olsson (Ed.),, *Proceedings of ForestSat 2005*, Borås, Sweden, May 31–June, 3 2005. Jönköping: National Board of Forestry, Sweden (pp. 22–26).
Ohmann, J. L. & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research*, *32*, 725–741.
Ohmann, J. L., Gregory, M. J. & Spies, T. A. (2007). Influence of environment, disturbance, and ownership on forest vegetation of coastal Oregon. *Ecological Applications*, *17*, 18–33.
Opsomer, J. D., Breidt, F. J., Moisen, G. G. & Kauermann, G. (2007). Model-assisted estimation of forest resources with generalized additive models. *Journal of the American Statistical Association*, *102*, 400–408.
Packalén, P. & Maltamo, M. (2007). The *k*-MSN method in the prediction of species specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sensing of Environment*, *109*, 328–341.
Packalén, P. & Maltamo, M. (2008). The estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Canadian Journal of Forest Research*, *38*, 1750–1760.
Persson, Å., Holmgren, J. & Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, *68*, 925–932.
Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J. & Packalén, P. (2007). Preharvest measurement of marked stands using airborne laser scanning. *Forest Science*, *53*, 653–661.
Pierce, K. B., Jr, Ohmann, J. L., Wimberly, M. C., Gregory, M. J. & Fried, J. S. (2009). Mapping wildland fuels and forest structure for land management: A comparison of nearest neighbor imputation and other methods. *Canadian Journal of Forest Research*, *39*, 1901–1916.
Poso, S. (1972). A method for combining photo and field samples in forest inventory. *Communicationes Instituti Forestalis Fenniae*, *76*, 1–133.
Poso, S. & Kujala, M. (1978). A method for national forest inventory in northern Finland. *Communications Instituti Forestalis Fenniae*, *93*, 1–54.
Reese, H., Nilsson, M., Granqvist Pahlén, T., Hagner, O., Joyce, S. & Tingelöf, U. (2003). Countrywide estimates of forest variables using satellite data and field data from the National Forest Inventory. *Ambio*, *32*, 542–548.
Saket, M., Altrell, D., Branthomme, A. & Vuorinen, P. (2002). *FAO's approach to support national forest assessments for country capacity building*. Rome: Forestry Department, FAO.
Särndal, C.-E., Swensson, B. & Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer-Verlag.
Solodukhin, V., Zukov, A. & Mazugin, I. (1977). Possibilities of laser aerial photography for forest profiling. *Lesnoe Khozyaisto (Forest Management)*, *10*, 53–58. (In Russian.)
Ståhl, G., Holm, S., Gregoire, T. G., Gobakken, Næsset, E. & Nelson, R. (2010). Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research* (in press).
Temesgen, H., LeMay, V. M., Froese, K. L. & Marshall, P. L. (2003). Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management*, *177*, 277–285.
Thessler, S., Sesnie, S., Bendaña, Z. S. R., Ruokolainen, K., Tomppo, E. & Finegan, B. (2008). Using *k*-nn and discriminant analyses to classify rain forest types in a Landsat TM image over northern Costa Rica. *Remote Sensing of Environment*, *112*, 2485–2494.
Tomppo, E. (1996). Multi-source National Forest Inventory of Finland. In R. Päivinen, J. Vanclay & S. Miina (Eds.), *New thrusts in forest inventory. Proceedings of the Subject Group S4.02-00, Forest Resource Inventory and Monitoring, and Subject Group S4.12-00, Remote Sensing Technology*, Vol. I. IUFRO XX World Congress, Tampere, Finland, August 6–12, 1995. EFI. *EFI Proceedings*, 7, 27–41.
Tomppo, E. (2006). The Finnish multi-source national forest inventory—Small area estimation and map production. In A. Kangas & M. Maltamo (Eds.), *Forest inventory—Methodology and applications* (pp. 195–224). Dordrecht: Springer.
Tomppo, E. (2009). The Finnish National Forest Inventory. In R. E. McRoberts, G. A Reams, P. A. Van Deusen & W. H. McWilliams (Eds.), *Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium*, Monterey, California, October 16–19, 2006 (USDA Forest Service Gen. Tech. Rep. WO-79, pp. 39–46).
Tomppo, E., Goulding, C. & Katila, M. (1999). Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. *Scandinavian Journal of Forest Research*, *14*, 182–192.
Tomppo, E., Gschwantner, T., Lawrence, M. & McRoberts, R. E. (2010*a*). *National forest inventories: Pathways for common reporting*. Dordrecht: Springer.
Tomppo, E., Haakana, M., Katila, M. & Peräsaari, J. (2008). *Multi-source national forest inventory*. Dordrecht: Springer.

Tomppo, E. & Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of *k*-NN estimation: a genetic algorithm approach. *Remote Sensing of Environment*, *92*, 1–20.

Tomppo, E., Korhonen, K. T., Heikkinen, J. & Yli-Kojola, H. (2001). Multisource inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. *Silva Fennica*, *35*, 309–328.

Tomppo, E., Malimbwi, R., Katila, M., Chamuya, R., Mäkisara, K., Peräsaari, J., et al. (2010*b*). *Sampling design for NAFORMA–Forest Monitoring and Assessment of Tanzania*. Presentation. Workshop on NAFORMA Sampling Design, November 30–December 1, 2009, Dar es Salaam, Tanzania.

Tomppo, E. O., Gagliano, C., De Natale, F., Katila, M. & McRoberts, R. E. (2009). Predicting categorical forest variables using an improved *k*-nearest neighbour estimator and Landsat imagery. *Remote Sensing of Environment*, *113*, 500–517.

Traub, B., Köhl, M. & Päivinen, R. (1997). Simulation study. In *Study on European forestry information and communication system: Report on forest inventory and survey systems*. Luxembourg: European Commission.

UNFCCC (1992). *United Nations Framework Convention on Climate Change*. FCCC/INFORMAL-MAL/84-GE. 05-62220(E) 200705. Retrieved February 2010, from www.unfccc.int/resource/docs/convkp/conveng.pdf

UNFCCC (2009). *United Nations Framework Convention on Climate Change*. Retrieved February 2010, from http://unfccc.int/2860.php

UN REDD (2009). UN REDD Programme. Retrieved February 2010, from http://www.fao.org/climatechange/51379/en/

Vauhkonen, J., Tokola, T., Packalen, P. & Maltamo, M. (2009). Identification of Scandinavian commercial species of individual trees from very high density ALS data using alpha shape metrics. *Forest Science*, *55*, 37–47.

Vidal, C., Lanz, A., Tomppo, E., Schadauer, K., Gschwantner, T., di Cosmo, L. & Robert, N. (2008). Establishing forest inventory reference definitions for forest and growing stock: A study towards common reporting. *Silva Fennica*, *42*, 247–266.

von Berg, E. (1858). Kertomus Suomenmaan metsistä 1858 [A description of Finnish forests in 1858]. In Edmund von Berg: Kertomus Suomenmaan metsistä 1858, sekä Kuvia suuresta muutoksesta [A description of Finnish forests in 1858 and pictures of the great change]. In M. Leikola (Ed.), (1995), *Kustannusosakeyhtiö Metsälehti, Gummerus Kirjapaino Oy.* Jyväskylä, Finland. (In Finnish.)