Remote Sensing of Environment 211 (2018) 146–153

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Lidar supported estimators of wood volume and aboveground biomass from the Danish national forest inventory (2012–2016)

Steen Magnussen[a,*], Thomas Nord-Larsen[b], Torben Riis-Nielsen[b]

[a] Natural Resources Canada, Canadian Forest Service, Canada
[b] University of Copenhagen, Faculty of Science, Denmark

ARTICLE INFO

Keywords:
Mean ratio
Model assisted estimation
Model-calibration
Multiple imputations
Interpenetrating panel design
Standard error

ABSTRACT

National forest inventories (NFI) provide estimates of forest resources at the national and regional level but are also increasingly used as basis for mapping forest resources based on remotely sensed data. Such maps procure local estimates of forest resources but may also improve precision of national and regional estimates. Supported by a countrywide airborne laser scanning (circa 2014) and a national land-use map (circa 2014), direct (DI), model-assisted (MA), and model calibrated (MC) estimates of wood volume (V) and aboveground biomass (AGB) densities in forest areas derived from the Danish NFI (2012–2016) are presented. Nonlinear models with three LiDAR metrics are used to predict V and AGB in forested areas. According to these models, the predicted values of V and AGB in sample plots missed in the field inventory was lower than in those visited in the field; we therefore opted for estimation with multiple (stochastic) imputations. MA estimates for the country suggested a 2% lower level of both V and AGB densities with errors 45% lower than estimated errors in DI results. National MC estimates were close to the DI estimates with an error approximately 40% lower than errors in DI estimates yet 5% greater than the MA estimates of error. Multiple imputations had the strongest impact on DI estimates, but only a weak impact on MA and MC results.

1. Introduction

National forest inventories (NFI) are typically planned and designed to provide national and regional estimates of important forest resource attributes with a desired precision (Mandallaz, 2008, ch. 10; Ranneby et al., 1987; Tomppo, 2006). Contemporary NFIs employ a large number of field plots where attribute values of interest are either measured directly or derived from existing models of attribute associations (for examples, see Vidal et al., 2016). A systematic distribution of field plot locations across the sampling frame with a temporal and spatially balanced measurement cycle is now a common feature of many NFIs (for example, Schreuder et al., 2000).

Auxiliary variables, obtained by remote sensing (satellites, aircrafts, and drones) and correlated with attributes of interest, not only provide model-based forest resource maps of interest to regional and local forest management and land-use planners (Corona et al., 2014b; Nilsson et al., 2016; Nothdurft et al., 2009), but may also improve the precision of national and regional estimates (Magnussen et al., 2013; McRoberts et al., 2006; Saarela et al., 2015). A prime example is the growing use of LiDAR metrics of canopy heights and canopy density derived from airborne laser scanner (ALS) data in support of national, regional, and enterprise forest inventories (Lindgren et al., 2015; Maltamo et al., 2009; Melville et al., 2015b; Næsset, 2014).

Since LiDAR registers the height of actively growing vegetation surfaces at specific points in time, it becomes important to synchronize the acquisition of ALS and field measurements of heights in order to optimize the predictive power of LiDAR metrics (Massey and Mandallaz, 2015; McRoberts et al., 2016). Everything else equal, this is easier to accomplish at the regional level, or for a small country like Denmark with a surface area $A_{\mathrm{DK}} = 43{,}098\ \mathrm{km}^2$ (Nord-Larsen and Riis-Nielsen, 2010; Nord-Larsen and Schumacher, 2012; Wulder et al., 2012).

We investigate possible improvements in the precision of national and regional estimates of volume and total aboveground live tree biomass achievable through a combined use of field data from the NFI plots and a census of LiDAR based predictions of these attributes. The benchmarks for the comparisons are estimates derived exclusively from the field data collected in sample plots (direct estimates or DI for short). Direct estimates are compared to model-assisted (MA) and model-calibrated (MC) results where we have exploited LiDAR derived model-based predictions (Särndal, 2007; Wu and Sitter, 2001).

Nonresponse is the norm in every national and regional inventory

* Corresponding author at: Pacific Forestry Center, 506 West Burnside Rd., Victoria, BC V8Z 1M5, Canada.
E-mail address: Steen.magnussen@canada.gc (S. Magnussen).

https://doi.org/10.1016/j.rse.2018.04.015
Received 10 November 2017; Received in revised form 31 January 2018; Accepted 8 April 2018

(Patterson et al., 2012). Denied access, dangerous terrain, inclement weather, budgetary constraints, and several other issues can prevent fieldwork. We demonstrate multiple (stochastic) imputations (Carpenter and Kenward, 2013; McRoberts, 2001; Van Deusen, 1997) as an option for dealing with non-response that is made easier by available model-based predictions of missing values (McRoberts, 2003) and dedicated software solutions.

2. Materials and methods

2.1. The NFI sampling design

The Danish National Forest Inventory (NFI) is a continuous inventory, with partial replacement of sample plots located on a 2 × 2-km grid covering the entire country. The sampling frame is the land surface of Denmark (Nord-Larsen and Johansen, 2016). Approximately one-third of the sample plots are permanent and are re-measured in every cycle of the NFI, whereas two-thirds are temporary with a location selected at random within a particular 2 × 2-km grid cell and measurement cycle. The sample plots are distributed, geographically, into five non-overlapping and spatially balanced interpenetrating panels. Each year within a 5-year cycle, a different panel is measured. Each sample plot is composed of four circular subplots with a radius of 15 m and located in the corners of a 200 × 200 m square. An overview on the number of NFI sample plots and subplots and their measurement schedule is in Table 1. The design included $n = 15{,}137$ sample plots. However, subplots with a center in open water were dropped from the design leaving us with 96% of the plots with a full complement of four subplots, and 4% of the plots approximately equally split to plots with 1, 2 or 3 subplots. The number of subplots measured each year is between 19 and 21% of the full complement.

Based on interpretation of aerial photos, each subplot belongs to one of three categories: (0) unlikely to contain forest or other wooded land cover, (1) likely to contain forest, and (2) likely to contain other wooded land. To avoid errors of omissions, a class 0 means an absence of vegetation within a 15–20 m buffer zone around a subplot potentially in category (1) or (2). Further, we took great care to detect recent afforestation. All plots with at least one subplot in category (1) or (2) as well as those observed to have a forest cover in the most recent rotation of measurements are visited in the field. In this study, the aerial photo-interpretation is not used beyond this purpose.

In the current inventory cycle, a forest cover was observed in 7185 subplots and consequently these units provided information about forest V and AGB. By definition, V and AGB is zero in subplots unlikely to contain forest and not visited in the field.

When a subplot included different land-use classes or different forest stands, the area occupied by different classes (stands) was calculated geometrically. The lower limit for a registration of a forest area was 2% of the subplot area. The area in forest in the ith subplot (a_{fi}) served to compute the density of V and AGB in forested areas.

Under the assumption of a simple random sampling without replacement (SRS) of plots, the sample inclusion probability of the ith sample plot ($i = 1, \ldots, n$) is $n \times m_i \times a_{SP} \times {A_{DK}}^{-1}$ where m_i is the number of subplots in the ith plot, a_{SP} is the area of a subplot, i.e. a circle with a radius of 15 m viz. 706.86 m^2. The joint inclusion probability π_{ij} of the ith and jth sample plot becomes $(n \times m_i \times a_{SP} \times {A_{DK}}^{-1}) \times (n-1) \times m_j \times a_{SP} \times (A_{DK} - m_i \times a_{SP})^{-1}$.

Denmark has five administrative regions and we derive both a national and regional estimates. The region membership of a sample plot was determined by the majority of subplot memberships.

Table 1
The number of subplots and subplots with at least 2% forest area by year of measurement (YEAR).

Year	Subplots	Subplots with forest
2012	8617	1443
2013	8630	1511
2014	8590	1401
2015	8590	1415
2016	8572	1415

2.2. Calculating volume and biomass in a plot

The data collected in a DNFI subplot and the compilation of V and AGB is detailed in (Nord-Larsen and Johansen, 2016). Briefly, tree stems are measured for diameter (*dbh*, mm) in three concentric circular plots according to size. Only trees with a DBH ≥ 400 mm are recorded on the entire subplot. Tree height (*h*, dm) was measured on a subset of 2 to 6 trees to allow a localized regression of *h* on *dbh* (Sloboda et al., 1993). For plots with no measurements of *h*, a set of generalized regression models were used (Johannsen, 2002). The total wood volume was then estimated with existing volume functions (Madsen, 1985, 1987; Madsen and Heusèer, 1993) and a scaling by the area supporting the tree data. The sum of wood volumes in the subplots yields the volume for a sample plot.

AGB values were obtained from a combination of: individual tree biomass functions (Nord-Larsen and Nielsen, 2015; Skovsgaard et al., 2011; Skovsgaard and Nord-Larsen, 2012); volume functions paired with biomass expansion factors (Skovsgaard and Nord-Larsen, 2012); and species specific wood density values (Moltesen, 1985). Total aboveground biomass in a sample plot is the area weighted sum of individual tree biomass in the subplots Note for notational convenience we use the V and AGB when we refer to volume and biomass per unit forest area. Qualifiers such as "mean ratio" and "density" are left out when context allows it.

2.3. Auxiliary data

An airborne laser scanner (ALS) survey of Denmark (2014–2015) provided a suite of LiDAR metrics used in connection with field observations from the NFI to fit models for predicting V and AGB. The ALS data were captured with a Riegel LMS-680i scanner in a fixed-wing aircraft flying at an average altitude of 680 m above ground level and at speed of approximately 240 km hr^{-1}. Within a single flight line, the average point density of first-return echoes was 4.6 per m^2 with a mean footprint size of 21 cm. Over forested areas, the actual point density of first returns was 7.5 per m^2 and 14.9 per m^2 for all return classes (Table 2). Only data with scan angles $< 30°$ were accepted. Up to five pulses and their intensity were recorded for each outgoing pulse. Close to 90% of the survey was done during the leaf-off period, but two smaller campaigns were completed in the summers of 2014 and 2015 during leaf-on conditions. The vertical accuracy of registered vegetation heights was 5 cm, while the horizontal accuracy of was 15 cm. Additional details are in (Nord-Larsen et al., 2017).

The number of growing seasoned lapsed between capture of the laser scanning data and a measurement of a subplot varied from −2 to +2 years with an approximately equal number of subplots to years skipped. The scanning covered 2441 subplots with field measurements completed within one growing season of the capture of the LiDAR metrics. Only these subplots were used to derive models for the association between wood volume and aboveground biomass as dependent variables and LiDAR metrics as explanatory variables.

After linking the laser scanning data to the 2441 subplots, a backward selection method starting with a 'full' nonlinear model with six LiDAR metrics stopped at a more parsimonious model with just three predictors. The resulting models were (Nord-Larsen et al., 2017, Table 3):

$$\begin{aligned} \widehat{V} &= 15.7277 \times Dz_{mean12}^{1.2254} \times Dz_{p95_{12}}^{-0.0138} \times IR_1^{0.9049} \\ \widehat{AGB} &= 7.5237 \times Dz_{mean12}^{1.2462} \times Dz_{p95_{12}}^{-0.0006} \times IR_1^{0.8453} \end{aligned} \tag{1}$$