Remote Sensing of Environment xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Large-area hybrid estimation of aboveground biomass in interior Alaska using airborne laser scanning data

Liviu T. Ene[a], Terje Gobakken[a,*], Hans-Erik Andersen[b], Erik Næsset[a], Bruce D. Cook[c], Douglas C. Morton[c], Chad Babcock[d], Ross Nelson[e]

[a] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Ås, Norway
[b] USDA Forest Service, Pacific Northwest Research Station, Seattle, WA 98195, USA
[c] Code 618/Biospheric Science Laboratory, NASA-Goddard Space Flight Center, Greenbelt, MD 20771, USA
[d] School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195, USA
[e] Code 618/Biospheric Science Laboratory–retired, NASA-Goddard Space Flight Center, Greenbelt, MD 20771, USA

A R T I C L E I N F O

Keywords:
Airborne laser scanning
Hybrid estimation
Biomass
Alaska

A B S T R A C T

Use of data from airborne laser scanning (ALS) is a well-established practice for enhancing the accuracy of forest inventories in combination with ground-based observations. For regular monitoring of large areas, wall-to-wall ALS data is economically prohibitive. However, when data are acquired in a strip-sampling mode, ALS can support the estimation of forest resources on large areas such as counties, states, and nations. This study investigated the combined use of ALS data acquired along parallel strips and co-located field observations of live and standing dead trees to estimate aboveground biomass (AGB) at regional and sub-regional levels. The study area comprised the Tanana Inventory Unit (TIU) (138,566 km^2), located in interior Alaska, and four sub-regions (106–7282 km^2) within the TIU. The field data consisted of 1676 ground observations from six independent field campaigns. Ninety-four plots came from the Forest Inventory and Analysis (FIA) program of the USDA Forest Service. The FIA plots, each consisting of four 1/60 ha subplots, were acquired using a probabilistic sampling design. The remaining ground observations were collected during field inventories across TIU, and consisted of 335 independent field plots and 257 clusters containing up to four plots each. The estimation procedure followed a two-phase sampling design grounded within a hybrid inferential framework, which combines design-based estimation for the first phase and model-based inference in the second phase. Post-stratified estimation by land-cover class indicated that the hybrid estimation obtained 11–55% better precision compared to direct (design-based) estimation based solely on field observations. When the field sample sizes were reduced to 25%, the standard errors for the hybrid estimation increased by 3 to 5 pp. Direct and hybrid estimates were compared for two sub-regions, and the hybrid estimated AGBs were 6.4% and 13.3% larger than the direct estimates. On both these study sub-regions, the 95% confidence interval of the hybrid estimated mean AGB included the direct estimate of the mean and vice versa.

1. Introduction

Large-area forest inventories carried out by National Forest Inventory (NFI) programs across entire countries and regions require unbiased estimators for the attributes of interest for national and international statistical reporting (Lawrence et al., 2010). These unbiased estimators require field observations collected following probabilistic sampling designs and estimation methods grounded in a design-based inferential framework (Lawrence et al., 2010). However, due to a limited number of field plots (due to cost constraints), direct estimation methods based solely on ground plots do not provide precise estimates at sub-regional spatial scales, or by land-use or land-cover classes (Fahey et al., 2010; Gonzales et al., 2010). The need for accurate and timely information on forest resources at different geographic scales has motivated research that employs remotely sensed information to enhance traditional estimation methods for forest attributes such as forest area, volume, biomass, and carbon stocks. Such estimates are relevant not only for reporting purposes, but also to support sustainable forest ecosystem management practices (McRoberts et al., 2010a).

Airborne laser scanning (ALS) has proven to be an excellent source

* Corresponding author.
E-mail addresses: Liviu.ene@wsl.ch (L.T. Ene), Terje.Gobakken@nmbu.no (T. Gobakken), handersen@fs.fed.us (H.-E. Andersen), Erik.Naesset@nmbu.no (E. Næsset), Bruce.Cook@nasa.gov (B.D. Cook), Douglas.Morton@nasa.gov (D.C. Morton), babcoc76@uw.edu (C. Babcock).

http://dx.doi.org/10.1016/j.rse.2017.09.027
Received 15 January 2017; Received in revised form 12 September 2017; Accepted 20 September 2017
0034-4257/

of auxiliary information for increasing the precision of forest inventory estimates (Zolkos et al., 2013; Næsset, 2014). Currently, estimating quantitative forest structure attributes (volume, biomass, etc.) using ALS data is considered superior to other approaches using remotely sensed data produced by either active or passive sensors (Zolkos et al., 2013). The strengths of ALS have been recognized in support of inventories for stand-level management, NFIs, and MRV (Measuring, Reporting, and Verification) for e.g. the REDD+ mechanism (Reducing Emissions from Deforestation and Forest Degradation in Developing Countries) under the United Nations Framework Convention on Climate Change (GOFC-GOLD, 2014).

However, the high accuracy of ALS-assisted inventories comes at a high cost, which renders this approach unfeasible for acquiring wall-to-wall data over large areas (Næsset et al., 2009; Mitchard et al., 2012). Alternatively, ALS data acquired on a small portion of the population of interest can be integrated in a sampling framework, allowing precise estimation at much smaller inventory costs.

Forestry applications of airborne lasers in a sampling mode were pioneered using profiling systems. For instance, the Portable Airborne Laser System (PALS; Nelson et al., 2003) was applied to estimate forest biomass for areas up to 5000 km^2 (Nelson et al., 2004; Nelson et al., 2008) using a line intersect sampling estimation framework. Subsequent work with PALS in Hedmark, Norway (27,000 km^2; Nelson et al., 2012) considered other inferential frameworks such as the model-based technique (Ståhl et al., 2011) employed in the current study and a design-based and model-assisted approach (Gregoire et al., 2011).

With the advent of commercially available ALS systems, large-area, ALS-assisted forest inventory applications have focused primarily on area-based estimation methods that couple plot-level estimates of forest attributes (like volume and biomass) with variables extracted from co-located ALS measurements using various statistical modelling approaches. Vauhkonen et al. (2014) provide a thorough review of the development of ALS forest applications. Double-sampling applications using ALS measurements were discussed by Parker and Evans (2004), Andersen and Breidenbach (2007), Corona and Fattorini (2008), Stephens et al. (2012), and Ene et al. (2016a). Further technological developments, especially the 2003 launch of the NASA's ICESat Geosciences Laser Altimetry System (GLAS) lidar, inspired a three-phase sampling strategy for large-area biomass estimation in boreal forests that combined space-based laser observations with PALS profiling laser data and ground measurements (Boudreau et al., 2008; Nelson et al., 2009; Neigh et al., 2013; Margolis et al., 2015). Subsequent work with three-phase designs continues with the use of an ALS system in place of the airborne profiler (Nelson et al., 2017) to tie satellite laser measurements to ground estimates of biomass in the continental United States and Mexico. Recent work has also included Bayesian model-based techniques (e.g. Kauranne et al., 2017).

Using ALS as a strip sampling tool where the remotely-sensed information is collected along swaths (corridors of ALS data) has been suggested by Næsset (2005) and Gobakken et al. (2006), and suitable statistical estimators were developed by Gregoire et al. (2011), Ståhl et al. (2011), Ene et al. (2016a), and Ringvall et al. (2016). Empirical studies performed in boreal forests (Andersen et al., 2009; Andersen et al., 2011; Ene et al., 2013; Gobakken et al., 2012) and miombo woodlands in Africa (Ene et al., 2016b) indicated that the ALS-assisted inventories can be cost efficient, with little loss of precision, even when data with only partial coverage are used (Ene et al., 2016a).

Inventories using ALS in a sampling mode require that the first-phase sample (i.e., the ALS sample) is acquired using a probabilistic sampling design, which allows us to expand the sample estimates using the known inclusion probabilities for each ALS strip in the sample. The second-phase sampling (i.e., the field inventory data collection) provides the ground observations required for building a predictive model, and can follow either a probabilistic design (Gregoire et al., 2011; Ståhl et al., 2011; Gobakken et al., 2012) or a non-probabilistic sampling design. However, when the observations are purposively sampled, the selection mechanism should not depend on the response values (ignorable sampling), such that the design can be ignored for inference (see Thompson, 2012, p. 136, Corona et al., 2014 and references therein). Moreover, an external model, i.e., a model fitted to data provided by another inventory (Mandallaz, 2008, p. 71) can also be used if field data collection is not feasible, as long as the estimated variance-covariance matrix of the model parameter estimates is available (Ståhl et al., 2011, 2016).

When both first- and second-phase sampling follow a probabilistic design, the estimation can be considered approximately unbiased under the design-based inferential framework, regardless of the quality of the predictive model (Särndal et al., 1992). However, when using a non-probabilistic second-phase sample or an external model, the inference relies heavily on the quality of the model, and the unbiasedness is no longer guaranteed. Recently, this inferential framework that combines design and model-based inferential approaches has been termed “hybrid inference” (Corona et al., 2014; Ståhl et al., 2016). For an insightful comparison of the model-based, model-assisted, and hybrid inferential frameworks, see Gregoire (1998), Corona et al. (2014), and Ståhl et al. (2016).

Using ALS as a strip-sampling tool for estimating forest biomass in remote forest regions across Alaska has previously been considered to be a cost-efficient solution for providing accurate estimates needed for decision-making by communities and agencies across the Kenai Peninsula (Andersen et al., 2009) and upper Tanana Valley (Andersen et al., 2011). With very high costs to establish new field plots in these remote regions, using ALS in a sampling framework can provide a means for accurate estimation of forest resources across large extents, even when the field sample size is severely limited. The FIA program of the USDA Forest Service Pacific Northwest Research Station, with primary responsibility for monitoring and assessment of the forest resources of Alaska, is investigating more cost-effective inventory techniques that draw on a combination of sparse field measurements and remotely-sensed data to produce information required for decision making at community, sub-regional, and regional levels. Thus, ALS sampling can be integrated with the FIA system to address problems related to cost and logistics to provide high-quality estimates at regional and sub-regional levels.

Using ALS in a strip-sampling framework in Alaska poses particular challenges to estimation. For example, a systematic selection of ALS strips hinders the unbiased estimation of sampling variability. Analysts frequently assume that systematically-acquired flight lines can be handled as a random sample, and this assumption can lead to inflated variance estimates where the degree of inflation depends on the presence and strength of geographic trends associated with the parameter being estimated. For instance, if systematic east-west flight lines are flown on a study area where there is a pronounced forest biomass gradient from north to south, sampling variance can be significantly overestimated if ALS data collected systematically are assumed to be a random sample (Särndal et al., 1992; Ene et al., 2012).

The overall study aim was to broaden the applicability of large-area, ALS-assisted inventories by relaxing the requirement of using a probability field sample as assumed by Andersen et al. (2009, 2011) and Gobakken et al. (2012). Considering a hybrid inferential framework, post-stratified estimation of per-hectare aboveground biomass (AGB) by land-cover classes was pursued at regional and sub-regional levels by incorporating ground inventory data from different campaigns and agencies acquired using a combination of probabilistic and non-probabilistic sampling, and auxiliary variables derived from a probabilistic ALS sample. The efficiency of the ALS-assisted estimation relative to direct estimation solely based on the ground sample was assessed considering different field sampling intensities.

2. Material

The material consisted of ground data provided by field inventories