Forest Ecology and Management 331 (2014) 12–18

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Using a remote sensing-based, percent tree cover map to enhance forest inventory estimation

Ronald E. McRoberts *, Greg C. Liknes, Grant M. Domke

Northern Research Station, U.S. Forest Service, 1992 Folwell Avenue, Saint Paul, MN 55108, USA

ARTICLE INFO

Article history:
Received 10 May 2014
Received in revised form 24 July 2014
Accepted 27 July 2014
Available online 20 August 2014

Keywords:
Stratified estimator
Model-assisted regression estimator
Precision
Forest inventory
Auxiliary information

ABSTRACT

For most national forest inventories, the variables of primary interest to users are forest area and growing stock volume. The precision of estimates of parameters related to these variables can be increased using remotely sensed auxiliary variables, often in combination with stratified estimators. However, acquisition and processing of large amounts of remotely sensed data can be costly and laborious, and stratified estimation requires construction of strata and satisfaction of within-stratum sample size constraints. An alternative to both challenges is to use an existing remote sensing-based, spatial product with the model-assisted estimators. The latter estimators use continuous auxiliary information directly rather than their aggregation into strata and are not subject to such severe sample size constraints. The objective of the study was to compare estimates of mean proportion forest area and mean growing stock volume per unit area obtained using both stratified and model assisted estimators with a remote sensing-based percent tree canopy cover map as auxiliary information. For a study area in Minnesota, USA, the primary conclusion was that estimates obtained with both sets of estimators were acceptably precise, but that the model-assisted estimators were easier to implement and facilitated aggregation of estimates from smaller sub-areas to estimates for larger areas.

Published by Elsevier B.V.

1. Introduction

National forest inventories (NFI) typically report estimates of parameters related to forest area and growing stock volume and their changes using unbiased probability-based (design-based) estimators. The estimates are used for multiple purposes including strategic planning (USDA-FS, 2012) and reporting for an increasing number of international agreements such as the Global Forest Resources Assessment (FAO, 2010) and Annex 1 of the United Nations Framework Convention on Climate Change (UNFCCC, 2006). However, for important inventory parameters related to forest area and volume, limited sample sizes inhibit these estimators from producing sufficiently precise estimates unless the estimation process is enhanced using auxiliary information. Remote sensing-based thematic maps are increasingly used as auxiliary information to address this challenge.

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service conducts the NFI of the United States of America (USA) and has conducted extensive research on using remotely sensed data to enhance inventory estimates via stratified estimation. (Hansen and Wendt, 2000; McRoberts et al., 2002a, b, 2006, 2012; Liknes et al., 2004, 2009; Nelson et al., 2005; Westfall et al., 2011). However, for large areas such as states, provinces, and regions, the labor and costs associated with acquiring and processing large amounts of remotely sensed data inhibit this practice. For example, 10–20 Landsat scenes are required to cover individual states in the Midwestern region of the USA. An alternative is to use a readily available, remote sensing-based, spatial, thematic product such as the National Land Cover Database (NLCD) (Vogelmann et al., 2001; Homer et al., 2004, 2007). The NLCD is a 30-m $\times$ 30-m, multi-class, land cover dataset that has been widely used as a source of auxiliary information for multiple purposes. McRoberts et al. (2002a) aggregated the thematic classes of the 1992 NLCD to forest and non-forest and then constructed four related strata. These strata, when used with stratified estimators (Section 3.2.2), reduced variances of estimates of mean proportion forest area by factors as great as 3.2 for four Midwestern states. This approach was operationally implemented for at least some of the regional FIA programs in the USA. McRoberts et al. (2006) later showed that stratifications based on estimates of pixel-level probabilities of forest cover reduced variances of estimated mean proportion forest by factors as great as 5.9 and variances of estimates of mean growing stock volume per unit area by factors as great as 2.5. The similarity between the probability of forest cover

* Corresponding author. Tel.: +1 651 649 5174.
E-mail address: rmcroberts@fs.fed.us (R.E. McRoberts).

http://dx.doi.org/10.1016/j.foreco.2014.07.025
0378-1127/Published by Elsevier B.V.

and the percent tree canopy cover layer for the 2001 NLCD facilitated operational implementation of the latter approach to stratification.

One disadvantage of stratified estimation is that the full utility of continuous auxiliary data such as percent tree canopy cover is not realized when the data are aggregated into a small number of strata. In addition, when using continuous auxiliary information for stratification, strata boundaries must be selected, and sufficient numbers of observations per stratum must be ensured (Westfall et al., 2011). The model-assisted regression estimators are alternative estimators that more fully utilize continuous auxiliary data. These estimators calculate an initial estimate as the aggregation of estimates for individual population units and then adjust the initial estimate using differences between unit-level estimates and observations for a probability sample (Section 3.2.3). The increased availability of remotely sensed satellite and lidar data and products based on them has increased the appeal of model-assisted estimators for forest inventory applications (Baffetta et al., 2009; Gregoire et al., 2011; McRoberts, 2010, 2011; McRoberts and Walters, 2012; McRoberts et al., 2013a, b; Næsset et al., 2011, 2013a, b; Vibrans et al., 2013; Sannier et al., 2014). With these estimators, selection of strata boundaries is not necessary because the auxiliary data are used in their continuous form. In addition, although overall minimum samples sizes must be accommodated, satisfaction of a sample size criterion for each of multiple strata is not necessary.

For inventory estimation, the ultimate analytical objective is a statistical inference in the form of a confidence interval calculated as $\hat{\mu} \pm t_{1-\alpha} \cdot \sqrt{V\hat{a}r(\hat{\mu})}$ where $\hat{\mu}$ is the estimate of a mean, $V\hat{a}r(\hat{\mu})$ is an estimate of the variance of the estimated mean, and t corresponds to the confidence level. The primary objective of the study was to compare estimates of mean proportion forest area and mean forest growing stock volume per unit area using continuous NLCD tree canopy cover data as auxiliary information with three sets of statistical estimators: (1) the simple random sampling estimators, (2) the stratified estimators, and (3) the model-assisted regression estimators. A particular underlying objective was to determine if the model-assisted regression estimators circumvent the disadvantages of the stratified estimators without introducing additional disadvantages.

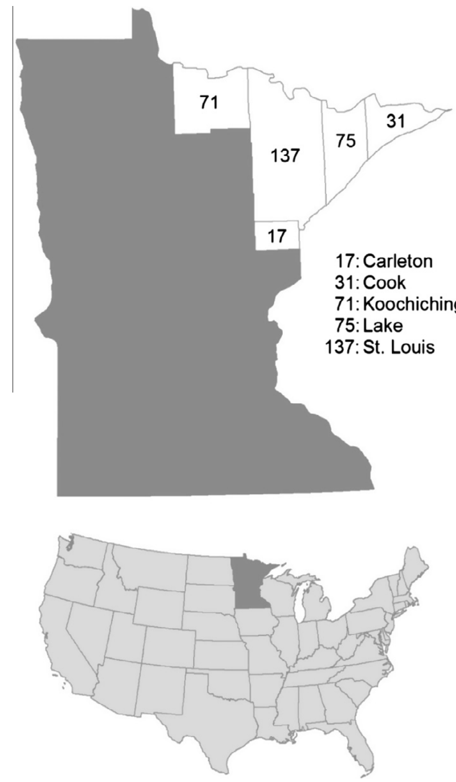

Fig. 1. Study area consisting of Minnesota FIA Unit 1 and five subordinate counties by name and U.S. Federal Information Processing Standard (FIPS) codes.

2. Data

2.1. Forest inventory field data

The study area was Minnesota FIA Inventory Unit 1 and the five counties included within the Unit (Fig. 1). Land use for the study area consists of forest land dominated by aspen-birch and spruce-fir associations, agriculture, wetlands, and water (Miles et al., 2011). The FIA program samples without replacement and has established field plot centers in permanent locations using a quasi-systematic sampling design that is regarded as producing an equal probability sample (McRoberts et al., 2010). Each FIA plot consists of four 7.32-m (24-ft) radius circular subplots that are configured as a central subplot and three peripheral subplots with centers located at distances of 36.58 m (120 ft) and azimuths of 0°, 120°, and 240° from the center of the central subplot. Field crews visually estimate the proportion of each subplot that satisfies the FIA definition of forest land: minimum area of 0.4 ha (1.0 ac); minimum canopy cover of 10%; stand width, measured as external crown-to-crown distance, of at least 36.6 m (120 ft); and forest land use. Field crews also observe species and measure diameter at-breast-height (dbh) (1.37 m, 4.5 ft) and height for all trees with dbh of at least 12.7 cm (5 in.). Volumes for individual trees are estimated using statistical models (Woodall et al., 2011), aggregated at plot-level, expressed as volume per unit area, and for inventory estimation purposes are considered to be observations without error (McRoberts and Westfall, 2014). For this study, data were used for 2681 FIA plots measured between 2008 and 2012. At the time the plots were measured, the sampling intensity in the study area was approximately one plot per 1200 ha. Two variables were considered: proportion forest area (FOR) and growing stock volume per unit area (VOL, m^3/ha).

2.2. Auxiliary data

The NLCD is a 30-m × 30-m, multi-class, Landsat-based land cover product of the Multi-Resolution Land Characteristics Consortium, a collaboration of multiple agencies of the U.S. Government (Homer et al., 2004, 2007). The NLCD is national in scope and provides spatial data for thematic classes such as urban, agriculture, and forest and separately for percent tree canopy cover. The latter data reflect land cover, not land use, and were predicted from Landsat 7 ETM + images and high resolution reference data using regression trees (Huang et al., 2001). For this study, the 2001 NLCD percent tree canopy cover (PCT) product was used as auxiliary information for both stratified and model-assisted estimation. For each FIA plot, FOR and VOL were associated with PCT for the NLCD map unit containing the plot center. Preliminary investigations indicated only negligible benefits accrued from using mean PCT for the 3 × 3 block of pixels centered on the map unit containing the plot center.

3. Methods

3.1. Assumptions

All three estimators rely on the same three underlying assumptions: (1) a finite population, U, consisting of N units in the form of