Remote Sensing of Environment 130 (2013) 87–95

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Using satellite image-based maps and ground inventory data to estimate the area of the remaining Atlantic forest in the Brazilian state of Santa Catarina

Alexander C. Vibrans [a], Ronald E. McRoberts [b,*], Paolo Moser [a], Adilson L. Nicoletti [a]

[a] *Universidade Regional de Blumenau, Blumenau, Santa Catarina, Brazil*
[b] *Northern Research Station, U.S. Forest Service, Saint Paul, MN USA*

ARTICLE INFO

Article history:
Received 6 July 2012
Received in revised form 24 October 2012
Accepted 27 October 2012
Available online xxxx

Keywords:
Confidence interval
Model-assisted regression estimator
IPCC good practice guidance
Accuracy assessment
Santa Catarina Forest and Floristic Inventory

ABSTRACT

Estimation of large area forest attributes, such as area of forest cover, from remote sensing-based maps is challenging because of image processing, logistical, and data acquisition constraints. In addition, techniques for estimating and compensating for misclassification and estimating uncertainty are often unfamiliar. Forest area for the state of Santa Catarina in southern Brazil was estimated from each of four satellite image-based land cover maps, and an independent estimate was obtained using observations of forest/non-forest for more than 1000 points assessed as part of the Santa Catarina Forest and Floristic Inventory. The latter data were also used as an accuracy assessment sample for evaluating the four maps. The map analyses consisted of identifying classification errors, constructing error matrices, calculating associated accuracy measures, estimating bias, and constructing 95% confidence intervals for proportion forest estimates using a model-assisted regression estimator. Overall accuracies for the maps ranged from 0.876 to 0.929. The standard errors of the estimates were all smaller than the standard error of the simple random sampling estimate by factors ranging from approximately1.23 to approximately 1.69. The model-assisted regression estimator lends itself to easy implementation for adjusting for estimated classification bias and for constructing confidence intervals.

Published by Elsevier Inc.

1. Introduction

Forest ecosystems are among the most biologically rich and genetically diverse terrestrial ecosystems on earth (Dinerstein et al., 1995; Holdridge, 1947, 1967). Further, these lands provide habitat for 70% of known animal and plant species (Matthews et al., 2000), contribute almost half the terrestrial net primary biomass production (Groombridge & Jenkins, 2002), and provide vital economic, social, and environmental benefits.

1.1. Carbon accounting

Forest ecosystems also play a vital role in the global greenhouse gas (GHG) balance. Conversion of forest to other land uses accounts for as much as 25% of anthropogenic GHG emissions (Achard et al., 2002; Gullison et al., 2007), but the forestry sector is also the only one of the five sectors identified by the United Nations Framework Convention on Climate Change that has the potential for removal of GHG emissions from the atmosphere. Carbon accounting assesses the scale of GHG emissions from the forestry sector relative to other sectors and is typically conducted using one of two primary approaches. With the stock-difference approach, commonly used by countries with established national forest inventories (NFI), annual emissions are estimated as the mean annual difference in carbon stocks between two points in time. With the gain–loss approach, the net balance of additions to and removals from a carbon pool is estimated as the product of the rate of land use area changes, called activity data, and the responses of carbon stocks for particular land use changes, called emission factors. For developing countries with remote and inaccessible forests, the gain–loss approach can be used as a component of a national measurement, reporting, and verification (MRV) system.

Giardin (2010) notes that MRV systems typically include ground-based components for estimating emission factors and remote sensing-based components for estimating activity data for forest area and forest area change. The GOFC-GOLD Sourcebook (2012, Chapter 2) emphasizes the role of satellite remote sensing as an important source of data for estimating area changes, and the Good Practice Guidance (GPG) of the Intergovernmental Panel on Climate Change (IPCC) asserts that estimates, “should be accurate in the sense that they are neither over- nor underestimated as far as can be judged, and that uncertainties are reduced as far as practicable” (Penman et al., 2003, Section 5.2.1). Two practical statistical implications for remote sensing-based assessments are clear: (1) bias in statistical estimators of activity data resulting from misclassification of remotely sensed data should be estimated, and (2) uncertainty in remote sensing-based estimates of activity data should also be estimated. Apart from estimation of bias, the accuracy of estimates in

* Corresponding author. Tel.: +1 651 649 5174; fax: +1 651 649 5285.
E-mail address: rmcroberts@fs.fed.us (R.E. McRoberts).

0034-4257/$ – see front matter. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.rse.2012.10.023

the sense of over- or underestimation as per the IPCC GPG cannot be judged, and uncertainty cannot be reduced unless it is first estimated.

1.2. Remote sensing-based approaches for estimating area of land use and land use change

For use with an MRV, two primary remote sensing-based approaches for estimating the area of land use change are appropriate. Direct classification entails classification of change from ground observations of change and two sets of remotely sensed data that have been merged into a single dataset (Hayes & Cohen, 2007). In this case, forest area change can be estimated directly from the change map. Post-classification entails comparison of two classifications that are constructed separately using remotely sensed data from two different dates (Coppin et al., 2004; McRoberts & Walters, 2012). In this case, forest area change can be estimated by comparing the two independent estimates of forest area. Regardless of whether the direct or post-classification approach is used, estimates of bias should be calculated and subtracted from the estimate obtained from the maps and uncertainty in the form of variances should be estimated.

Although the remote sensing community generally prefers the direct classification approach, post-classification may often be the only alternative. For example, an initial assessment of forest/non-forest change may require comparison of estimates of forest area obtained from a current forest/non-forest map and an historical, baseline forest/non-forest map of different resolution (Penman et al., 2003, Section 2.4.4.1). For this application, direct classification is rarely feasible, leaving post-classification as the only alternative. Thus, methods for estimating the bias and uncertainty of forest area estimates obtained from baseline forest/non-forest maps are a necessary prelude to post-classification estimation of forest area change in accordance with the IPCC Good Practice Guidance.

1.3. Challenges for tropical regions

Remote sensing-based estimation of forest area and forest area change in tropical regions incurs both technical and scientific challenges including the diversity of definitions of “forest”, the large number of land use forms and anthropic vegetation types in the tropics (Steininger, 2000), lack of adequate remote sensing data and rural cadastral information in many regions, and lack of personnel qualified to process remote sensing data for large geographic regions. For many years tropical forest cover mapping was dominated by the question of how to detect deforestation of primary forest areas (Tucker & Townshend, 2000), while quantification of forest recovery, secondary forest formations, exploited primary forests and land use mosaics was not sufficiently analyzed. From both silvicultural and ecological perspectives, tropical secondary forests have historically attracted considerably less attention as a research objective than primary forests (Corlett, 1995; Finnegan, 1996). However, in the context of climate change and Reduction of Emissions from Deforestation and forest Degradation (REDD) discussions, the importance of secondary forests has been highlighted as a potential carbon sink (Fehse et al., 2002; Olschewskia & Benítez, 2005).

The reliability of remote sensing-based classifications of tropical secondary forests is inhibited by two factors, the complexity of these forests and the inefficiency of automated digital processing methods. First, most authors acknowledge that disturbed natural vegetation formations form the vast majority of remnants in the Brazilian Atlantic forest (Oliveira-Filho & Fontes, 2000). In Santa Catarina, these secondary forests are characterized by structurally simplified forest types and early successional stages. Further, the distinctions between well-developed, mature forests in the sense of Veloso et al. (1991) and other woody formations of tree and shrub species, often with more than one type of land use (agrisilviculture, silvipastoral) and including permanent crops such as coffee, tea and banana plantations are continuous, not categorical. These phenomena inhibit accurate classification of land use classes with many commonly used remote sensing techniques.

Second, although automated digital image processing methods are generally preferable to other methods, they may be less efficient for very large tropical areas due to complicating factors such as the sizes of the areas which are on the order of hundreds of thousands of square kilometers, the large numbers of image scenes that must be combined, the large number of forest formations in a variety of environmental conditions, and selection of areas for acquiring training data. Further, techniques that have been used vary considerably with respect to factors such as sources, resolutions, and transformations of remotely sensed data and parametric, non-parametric, and segment-based classification techniques (Carvalho & Scolforo, 2008; Oliveira et al., 2010). As expected, the complexity of the forests, the diversity of data sources, and variety of classification techniques inevitably lead to differences among maps of the same region.

Ribeiro et al. (2009) reviewed the literature and existing surveys for the Brazilian Atlantic forest, analyzed patterns of fragmentation, examined the conflicting estimates of the extent and distribution of the remaining Atlantic forest, and adjusted earlier estimates of forest cover for the entire biome from a range of 7 to 8% to a range of 11.6 to 16%. Differences among estimates are attributed to the inclusion of secondary formations in more recent surveys and remnants smaller than 100 ha which account for 32–40% of the total remaining forest area. Ribeiro et al. (2009) compiled the results by biogeographic subregions (Silva & Casteleti, 2003) but did not consider the political divisions in the biome. The overall result is an urgent need for data and methods to support rigorous statistical comparisons of forest/non-forest maps with respect to the baseline forest area estimates that may be obtained from them.

1.4. Objectives

Completion of the Santa Catarina Forest and Floristic Inventory (IFFSC) presents an unprecedented opportunity to conduct statistically rigorous comparisons of remote sensing-based forest/non-forest maps. For use as an accuracy assessment dataset, the IFFSC data satisfy important criteria: independence from training data, adequate sampling intensity, and broad geographical coverage. The objectives of the study were threefold: (1) to document methods for assessing the accuracy of forest/non-forest maps, for estimating parameters characterizing the populations depicted by the maps, and for comparing estimates for different maps; (2) to assess the accuracy of four satellite image-based land use maps for the state of Santa Catarina using the IFFSC ground data as an accuracy assessment dataset; and (3) to compare estimates of proportion forest cover obtained from the four maps.

2. Data

The study area was defined as the southern Brazilian state of Santa Catarina, located between latitudes 26° and 29° S and between longitudes 48° and 53° W and with area of 95,346 km^2 (Fig. 1). Three phytogeographic subdivisions established by Klein (1978) and Veloso et al. (1991) were used: dense ombrophilous forests (DEN), mixed ombrophilous forests with Araucaria (MIX) and deciduous forests (DEC). Vector files and related data for the four surveys were kindly provided by the responsible institutions (Table 1).

2.1. Remote sensing-based forest/non-forest maps

Four satellite image-based maps of forest cover for Santa Catarina have been constructed since 2005: (1) a survey of forest remnants of