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 12 December 2012

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

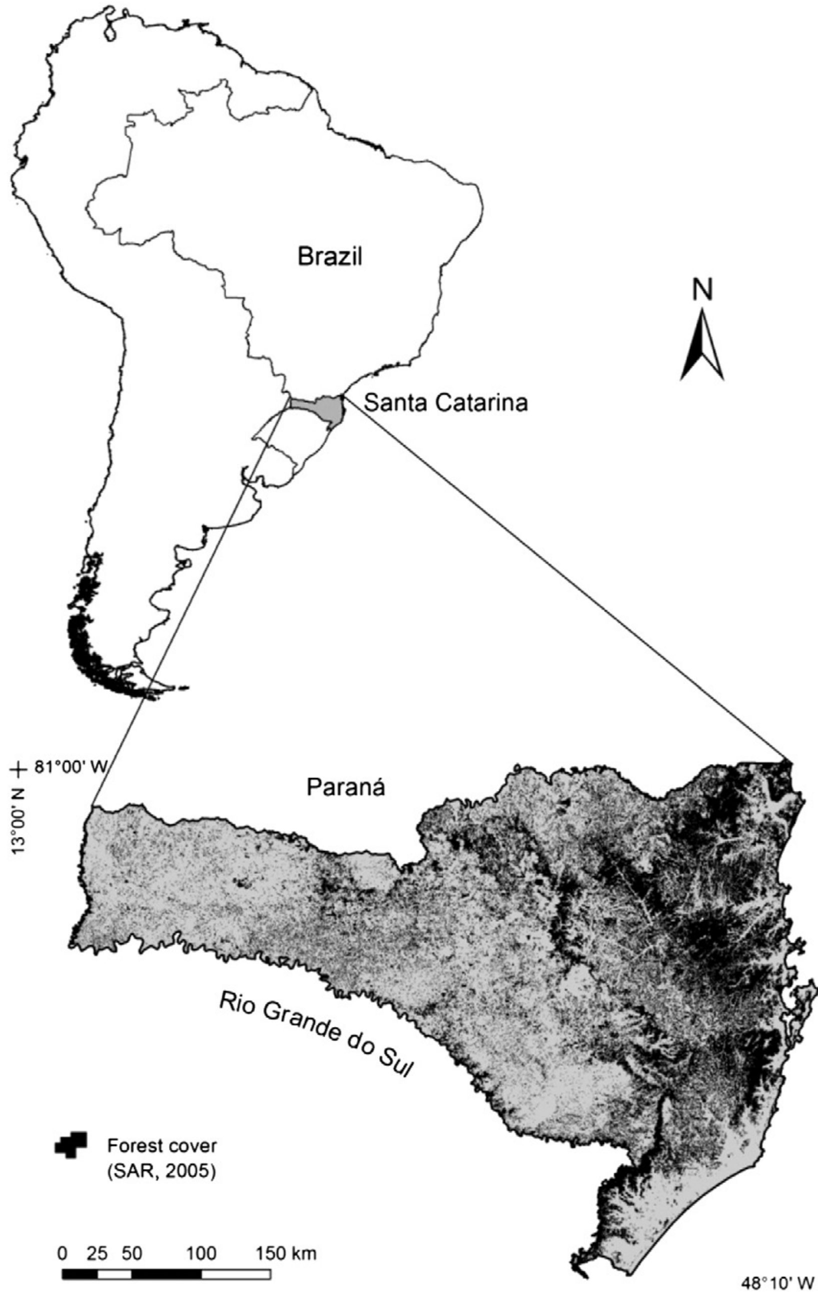

Fig. 1. State of Santa Catarina with black indicating forest cover as predicted from the LCF/SAR map (SAR, 2005).

Santa Catarina (Levantamento da Cobertura Florestal da Secretaria de Agricultura), later designated LCF/SAR (SAR, 2005); (2) a survey of the native vegetation of the Atlantic forest designated PROBIO (Cruz & Vicens, 2007); (3) an atlas of the Atlantic forest remnants, designated Atlas 2008 (Fundação, 2009); and (4) a general thematic map of the state of Santa Catarina (Projeto de Proteção da Mata Atlântica), designated PPMA (Geoambiente, 2008). Comparison of estimates of the area of forest fragments as proportions of the state's territory revealed large discrepancies, ranging between 0.22 and 0.41. The urgent necessity of a critical appraisal of these results motivated the present study.

Thematic classes for the four maps were similar but not always directly comparable: LCF/SAR classes — agriculture, pastures, forest, plantation forests, mangroves, restinga (coastal forest and open shrubby vegetation) (Scarano, 2009), forest regrowth, water bodies, dunes, urban areas, and clouds; PROBIO classes — forest remnants (which includes all forest, forest enclaves and ecotones, non-forest remnants which includes the enclaves without forests and vegetation refuges) and pioneer formations (with marine, fluvial or limnic influence); Atlas 2008 classes — forest remnants (primary and secondary) of DEN, MIX, DEC and savannas, restinga, including the lowland forests or "restinga higrófila", mangroves and dunes; and PPMA classes — agriculture, pastures and grasslands, pioneer forest formations, forests in primary, intermediate or advanced successional stages, plantation forests, mining areas, urban areas, and water bodies.

2.2. Santa Catarina Forest and Floristic Inventory (IFFSC)

The IFFSC data consist of ground plot observations of forest cover, structure, and composition obtained between 2007 and 2010 for 1074 equal probability sample points located at the intersections of a 10-km × 10-km grid. Of the 1074 points, 444 were in areas predicted

Table 1
Characteristics of the land use and forest cover maps for Santa Catarina.

Map attribute	Map			
	LCF/SAR (2005)	PROBIO (2007)	Atlas 2008 (2009)	PPMA (2010)
Executor	State Secretary of Agriculture	Federal Ministry of Environment	NGO	State Environment Agency
Satellite, sensor and year of image capturing	Landsat-5 TM, Landsat-7 ETM+ 2003/2004+	Landat-7 ETM+ (2001–2003), SRTM (2000)	CBERS-2 CCD (2005) Landsat-5 TM (2005–2008)	SPOT-4 2005
Scale of classification/ presentation	1:50,000/ 1: 50,000	1:250,000/ 1:250,000	1:50,000/ 1:50,000	1:50,000/ 1:50,000
Spatial Resolution	30-m × 30-m	30-m × 30-m	30-m × 30-m	20-m × 20-m (10-m × 10-m pan)
Minimal mapping area (nominal)	10 ha	40 ha	5(3) ha	2.5 ha
Method of interpretation/ classification	Visual	Object-based classification	Visual	Non-supervised classification (ISOSEG)
Number of control points for validation	Not informed	8000	Not informed	Not informed
Average accuracy	Not informed	0.86	Not informed	0.90
Definition of forest	Forests (mangrove and restinga excluded)	Forests of 3 types: DEN, MIX, DEC and ecotons (mangrove and restinga excluded)	Primary and secondary forest formations (mangrove and restinga excluded)	Primary forests and forests at medium and advanced successional stages (no initial successional stage forests)

to have forest cover according to at least one of the two forest cover maps available when the inventory was conducted. The LCF/SAR map indicated forest at 354 points, and the PPMA map indicated forest for 356 sample points. Between the two maps, 444 sample points were classified as forest, but field crews found that only 298 of the points actually had forest cover. For the 1074 IFFSC sample points, overall proportional agreement among the four maps was 0.762 of which 0.608 was for forest and 0.154 was for non-forest. Overall proportional agreement between pairs of maps ranged from 0.856 to 0.888 of which 0.614 to 0.715 was for forest and 0.165 to 0.275 was for non-forest. A sample plot, in the form of a cluster of four crosswise 1000-m^2 subplots (20-m × 50-m) was installed at each of the 298 points to collect field data on forest composition and structure (Vibrans et al., 2010). Floristic, structural and physiognomic information was acquired for all plots and was used to assess the accuracy of the maps (Vibrans et al., 2010, 2012).

The sampling points were located in the field using global positioning system (GPS) receivers that were calibrated at the geodesic station SAT 91858 IBGE located on the campus of the Regional University of Blumenau. Average observed error was 4.09 m, whereas the equipment itself indicates an accuracy of 5–25 m when operating in autonomous mode, even under closed forest canopies. This accuracy depends on the number and position of satellites tuned, atmospheric conditions, signal noise and other factors. The error read from the equipment during the field work ranged from 1.1 to 13.0 m with an average of 5.32 m. Accuracies of this order of magnitude are considered suitable for a mapping scale of 1:50,000.

3. Methods

3.1. Initial data processing

The vector file containing the sample point locations was overlaid on each of the four thematic maps and coregistered with reference to WGS 84 Datum. For each sample point, the land use class from the map was determined. For further analyses, land use classes were aggregated to forest and non-forest. Because early-stage forest for the PPMA map and intermediate-stage forest for the LRF/SAR map were only vaguely defined, and because neither represented more than 0.05% of the total area, these two classes were excluded from the forest class. Also because of methodological difficulties with the definition and classification of pioneer vegetation under marine, fluvial or wetland influence, including mangroves, restinga formations, dunes and other wetlands, these classes as mapped by PROBIO and LCF/SAR were also excluded from the forest class. Thus, a conservative estimate of forest cover would be obtained by eliminating possible sources of overestimation. All analyses were based on the two classes, forest and non-forest.

3.2. Simple random sampling estimation

Forest area was estimated using only the observations from the IFFSC ground sample under the assumption of simple random sampling (SRS). Because forest area is simply the product of total area and proportion forest, and because total area is known, estimation focused on proportion forest. The SRS estimator of proportion forest, $\hat{p}_{SRS}$, is simply the proportion of sampling points observed to have forest cover,

$$\hat{p}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} y_i, \tag{1}$$

where n is the total number of sample points, i indexes the sample, and

$$y_i = \begin{cases} 0 & \text{if non} - \text{forest land cover is observed} \\ 1 & \text{if forest land cover is observed} \end{cases}. \tag{2}$$

The variance of $\hat{p}_{SRS}$ is estimated as,

$$\begin{aligned} V\hat{a}r(\hat{p}_{SRS}) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{p}_{SRS})^2 \\ &= \frac{\hat{p}_{SRS}(1-\hat{p}_{SRS})}{n}, \end{aligned} \tag{3}$$

and a (1-α)% error confidence interval can be constructed as,

$$\hat{p}_{SRS} \pm t_{1-\frac{\alpha}{2}} \cdot \mathrm{SE}(\hat{p}_{SRS}), \tag{4}$$

where $t_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ percentile of Student's t-distribution and SE denotes standard error and is calculated as $SE(\hat{p}_{SRS}) = \sqrt{V\hat{a}r(\hat{p}_{SRS})}$. When a systematic sample is used, as in this case, variances may be slightly overestimated relative to estimates based on a simple random sample (Särndal et al., 1992). The primary advantage of the SRS estimator is that it is intuitive and unbiased, but the disadvantage is that variances may be large.

3.3. Map accuracy

The accuracies of the four thematic maps were evaluated using error or confusion matrices. An error matrix (Table 2) is used to compare numbers or proportions of observations and predictions by class and provides the necessary information for commonly used measures of accuracy: *overall accuracy* (OA), which is the proportion of observations correctly classified; *user's accuracy* (UA), which is the ratio of the number of correct predictions and the total number of predictions for a class; *producer's accuracy* (PA), which is the ratio of the number of correct predictions and the total number of observations for a class; *omission error rate* which is the ratio of the number of incorrect predictions and the total number of predictions for a class; and *commission error rate*, which is the ratio of the number of incorrect predictions and the number of observations for a class. Although data acquired using probability samples are preferable for assessing the accuracy of remote sensing-based classifications, such accuracy assessment data are very difficult, if not infeasible, to acquire for large areas of tropical forest. However, a concerted effort was made to acquire such data for this study.

Table 2
Error matrix where the subscripts 0, 1, and + denote 'non-forest', 'forest', and the sum over both classes, respectively; OA denotes overall accuracy; and n is the total sample size.

Ground truth (IFFSC)	Thematic map		Total	Producer's accuracy
	Non-forest	Forest		
Non-forest	n_{00}	n_{01}	n_{0+}	$\frac{n_{00}}{n_{0+}}$
Forest	n_{10}	n_{11}	n_{1+}	$\frac{n_{11}}{n_{1+}}$
Total	n_{+0}	n_{+1}	n	
User's accuracy	$\frac{n_{00}}{n_{+0}}$	$\frac{n_{11}}{n_{+1}}$		$OA = \frac{n_{00}+n_{11}}{n}$

3.4. Model-assisted estimation

Inferences in the form of confidence intervals for the map-based estimates of proportion forest were constructed using the model-assisted regression (MAR) estimator described by McRoberts and Walters (2012). This approach produces estimates from maps that are adjusted for classification errors and estimates of variances that can be used to construct confidence intervals for the proportion forest estimates.

The MAR estimator relies on both map predictions and accuracy assessment sample observations (Särndal et al., 1992). As an analogue to Eq. (2), define,

$$\hat{y}_i = \begin{cases} 0 & \text{if the map prediction of land cover is non} - \text{forest} \\ 1 & \text{if the map prediction of land cover is forest} \end{cases}. \tag{5}$$

An initial estimator of proportion forest is,

$$\hat{p}_{initial} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i, \tag{6}$$

where N is population size in terms of number of map pixels. The bias in this estimator resulting from classification errors can be estimated as,

$$\begin{aligned} \hat{Bias}(\hat{p}_{initial}) &= \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \\ &= \frac{n_{01}-n_{10}}{n}, \end{aligned} \tag{7}$$

where n_{01} and n_{10} are obtained from the error matrix (Table 2). The MAR estimator is then defined as,

$$\begin{aligned} \hat{p}_{MAR} &= \hat{p}_{initial} - \hat{Bias}(\hat{p}_{initial}) \\ &= \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i - \frac{n_{01}-n_{10}}{n}. \end{aligned} \tag{8}$$

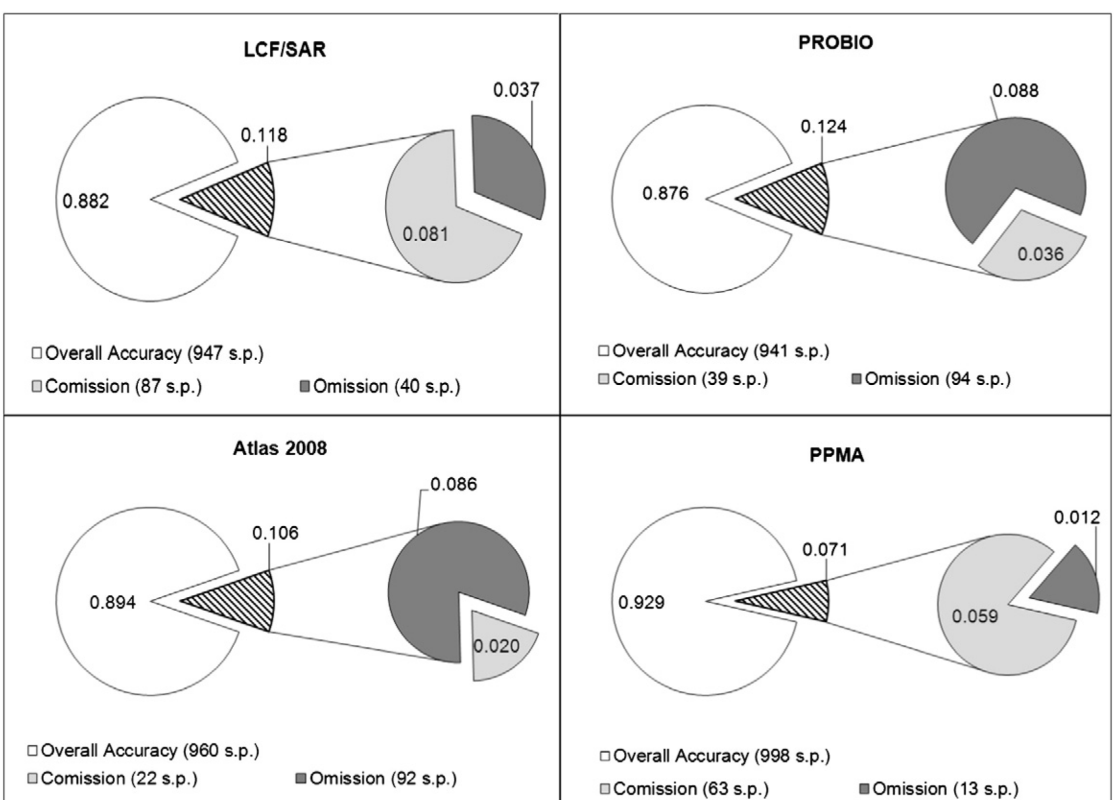

Fig. 2. Accuracies from error matrices for the four forest cover maps evaluated with IFFSC ground data, where s.p. denotes sample point.

Table 3
Plot-based, simple random sampling, forest cover estimates for Santa Catarina and for each phytogeographic region, where n is the sample size, and n_F is the number of sampling points with observed forest cover.

Estimate	Santa Catarina	Mixed ombrophilous forests (MIX)	Dense rain forests (DEN)	Deciduous forests (DEC)
Total area (km^2)	95,346	56,395	31,281	7671
Total area without pioneer formations (km^2)[a]	94,922	55,971	31,281	7671
n	1074	618	364	92
n_F	298	136	147	15
$\hat{p}_{SRS}$	0.277	0.220	0.404	0.163
$SE(\hat{p}_{SRS})$ [b]	0.014	0.017	0.026	0.039
Estimated forest area (km^2)	26,338	12,317	12,633	1251
Lower 95% CI limit (km^2)	23,744	10,452	11,024	660
Upper 95% CI limit (km^2)	28,932	14,183	14,242	1842
Forest cover 95% CI (proportion)	0.251–0.305	0.187–0.253	0.352–0.455	0.086–0.240

[a] With fluvial, limnic and marine influence (mangroves, restinga, dunes).
[b] SE denotes standard error and is calculated as the square root of variance.

McRoberts and Walters (2012) showed that subject to sample size and independence assumptions, the variance of $\hat{p}_{MAR}$ can be estimated as,

$$V\hat{a}r(\hat{p}_{\mathrm{MAR}}) = \frac{1}{\mathrm{n}-1}\left[(1-OA) - B\hat{i}as^2(\hat{p}_{initial})\right], \tag{9}$$

where $OA = \frac{n_{00}+n_{11}}{n}$ is obtained from the error matrix (Table 2). Variance estimates calculated using the MAR estimator may be smaller than for the SRS estimator if the correlation between the accuracy assessment observations and the map class predictions, as reflected in OA, is large. A (1-α)% confidence interval can be constructed in the same manner as for the SRS estimator, Eq. (4).

4. Results and discussion

4.1. Simple random sampling estimates

Application of the SRS estimators using only the IFFSC field data resulted in an estimate of forest cover for Santa Catarina, excluding areas covered by pioneer formations under fluvial, marine or limnic influence, of 26,338 km^2 with standard error of 2594 km^2 and 95% confidence interval of 23,744 km^2–28,932 km^2 (Table 3). The equivalent estimate of proportion forest cover was 0.277 with 95% confidence interval of 0.251–0.305.

4.2. Map accuracy

Accuracies for the four maps were assessed using information in the error matrices constructed using the IFFSC ground data (Fig. 2). For the LCF/SAR map, OA = 0.882; for the PROBIO map, OA = 0.876; for the Atlas 2008 map, OA = 0.894; and for the PPMA map, OA = 0.929 (Table 5). UAs for forest ranged from 0.754 to 0.906 with UA = 0.906 for the Atlas 2008 map being the greatest and UA = 0.754 for the LCF/SAR map being the least. PAs for forest ranged from 0.691 to 0.958 with PA = 0.958 for PPMA map being the greatest and PA < 0.700 for the Atlas 2008 and PROBIO maps being the least. For non-forest, PAs and UAs for all four maps were greater than 0.85.

Forest commission errors are defined as incorrect map predictions of the forest class for pixels the IFFSC ground crews observed as non-forest, and forest omission errors are defined as incorrect map predictions of the forest class relative to IFFSC ground crew observations. Errors may be caused not only by classification errors, but also by land use changes such as deforestation that occur between the image acquisition and ground observation dates, errors in the rectification of the plot locations and map pixels, and discrepancies between the plot and pixel sizes. The latter two issues pertain to whether plots and pixels represent the same area on the ground. The complete error matrix for the forest and non-forest classes for the LCF/SAR map is reported in Table 4.

Table 4
Error matrix for LCF/SAR map.

Ground observations (IFFSC)	Thematic map LCF/SAR			
	Non-forest	Forest	Total	Producer's accuracy
Non-forest	680	87	767	0.887
Forest	40	267	307	0.870
Total	720	354	1074	
User's Accuracy	0.944	0.754		Overall accuracy (OA): 0.882

4.3. Model-assisted estimation

Bias estimates were calculated using Eq. (7) and data from the error matrices. For the LCF/SAR and PPMA maps, commission errors exceeded omission errors, meaning that bias estimates were positive and that $\hat{p}_{initial}$ overestimated the proportion of forest cover by 0.044 and 0.047, respectively. For the PROBIO and Atlas 2008 maps, omission errors exceeded commission errors, meaning that bias estimates were negative and that $\hat{p}_{initial}$ underestimated the proportion of forest cover by 0.051 and 0.065, respectively. Thus, although the four maps had similar OAs, ranging from 0.876 to 0.929, they had dissimilar bias estimates, ranging from -0.065 to 0.047. These examples illustrate that similar OAs do not necessarily indicate similar bias estimates. OA is based simply on the proportion of correctly classified pixels, whereas bias estimates reflect the degree to which omission and commission errors fail to compensate for each other. Thus, a map with a large OA may also have a large bias estimate, and a map with a small OA may also have a small bias estimate.

The bias estimates calculated using Eq. (7) were used to adjust $\hat{p}_{initial}$ to obtain $\hat{p}_{\mathrm{MAR}}$ as per Eq. (8) (Fig. 2). For the Atlas 2008 map, the 92 omission errors for the 304 forested areas are reflected in the substantial bias estimate of −0.065 which causes the change from $\hat{p}_{initial} = 0.224$ to $\hat{p}_{MAR} = 0.289$. For the PPMA map, the 63 commission errors for the 358 points mapped as forest but found to have no forest cover in the field resulted in a bias estimate of 0.047 which leads to a reduction of forest cover estimates from $\hat{p}_{initial} = 0.415$ to $\hat{p}_{\mathrm{MAR}} = 0.368$. Information from the error matrices was also used to calculate variances and standard errors using Eq. (9) and confidence intervals using Eq. (4). The results of the analyses for the four forest cover maps are reported in Table 5, and estimates of forest cover for the state and for individual forest types are reported in Table 6.

4.4. Comparisons

The results of comparing the maps depended on the comparison criteria (Table 5). The MAR estimates of proportion forest for the four maps ranged from 0.289 for the Atlas 2008 map to 0.368 for the PPMA map. Differences between MAR estimates for all pairs of maps were statistically significantly different from zero ($P<0.05$) except for the difference between the estimates for the LCF/SAR and PROBIO maps. The four maps had similar OAs ranging from OA = 0.876 to OA = 0.929, all of which satisfy the accuracy criterion of 0.85 proposed by Anderson et al. (1976). However, estimates of bias resulting from classification errors varied widely, from −0.065 for the Atlas 2008 map to 0.047 for the PPMA map. Finally, standard errors, which reflect a combination of both OAs and bias estimates as per Eq. (9), were similar, ranging from 0.008 to 0.011.

Among the four maps, the PPMA map had the greatest OA, smallest absolute bias estimate, and smallest standard error of the estimate of proportion forest. Thus, for many applications, the PPMA map may be regarded as preferable. This result can possibly be attributed to several

Table 5
Estimates of forest cover for Santa Catarina using simple random (SRS) and model-assisted (MAR) estimation approaches.

Estimate	SRS	Map			
		LCF/SAR	PROBIO	Atlas 2008	PPMA
Mapped forest area (km^2)	–	35,499	25,680	21,341	39,531
OA	–	0.882	0.876	0.894	0.929
$\hat{p}_{initial}$	–	0.372	0.269	0.224	0.415
$\hat{Bias}(\hat{p}_{initial})$	–	0.044	−0.051	−0.065	0.047
$\hat{p}$ [a]	0.277	0.328	0.320	0.289	0.368
SE($\hat{p}$) [a,b]	0.014	0.010	0.011	0.010	0.008
95% CI ($\alpha \approx 0.05$)	0.251–0.305	0.307–0.349	0.299–0.341	0.270–0.309	0.352–0.384

[a] For SRS, $\hat{p} = \hat{p}_{SRS}$; for maps $\hat{p} = \hat{p}_{MAR}$.
[b] SE denotes standard error and is calculated as the square root of variance.

factors: (1) differences between the dates of the inventory plot observations and the remotely sensed data were small relative to the other maps; (2) the spatial resolution of the imagery was finest for the PPMA map; and (3) and the minimum mapping unit area was smallest for the PPMA map. However, for some applications other maps may be preferable. For example, UA for the forest class was 0.82 for the PPMA map but 0.91 for the Atlas 2008. Thus, if the objective is to locate forest land on the ground, the Atlas 2008 map may be preferable.

The SRS estimate of proportion forest cover for Santa Catarina is $\hat{p}_{SRS} = 0.277$ which is smaller than, but in the same general range as the MAR estimates. The standard errors of the MAR estimates were all smaller than the standard error of the SRS estimate by factors ranging from approximately1.23 to approximately 1.69. These results can be primarily attributed to the relatively large OAs which result from large correlations between map predictions and accuracy assessment observations. This result indicates that the maps not only produced fairly accurate ($0.876 \leq OA \leq 0.929$) border-to-border spatial depictions of the locations of forest and non-forest land cover, which the IFFSC plot observations cannot accomplish, but also contributed to producing more precise estimates of proportion forest cover than were obtained using only the plot data.

Neither imagery sources, classification methods, nor minimal mapping units provide definitive explanations for differences among estimates, particularly between the smaller, more conservative (pessimistic) estimates from the PROBIO and Atlas 2008 maps and the larger, more liberal (optimistic) estimates from the LCF/SAR and PPMA maps. Therefore, other factors must be considered. First, subtropical forest cover in the study region varies considerably with respect to species composition, age, exploitation history, density, basal area, and amount of biomass (Liebsch et al., 2008). Although nearly all forests have been or still are quite intensively exploited for timber and non-timber products, they have not been consistently managed following specific management guidelines, thus increasing their heterogeneity in terms of species composition and stand conditions. Second, the previously noted difficulties associated with constructing accurate forest cover maps for very large geographic areas (Section 1) may have contributed to inconsistencies. In particular, the validity of map accuracy assessments deteriorates when sufficient ground sample data are not acquired for dates comparable to those of the remotely sensed data on which the map is based. In addition, the Atlas 2008 and PROBIO maps were produced by nationwide projects that likely lacked intensive field support in Santa Catarina. Third, specific vegetation types that vary with latitude, altitude and distance from coasts and vary relative to agricultural traditions and land use patterns result in regionally unique natural and semi-domesticated vegetation mosaics (Groeneveld et al., 2009). Examples include extensive cattle grazing in thinned native forests, exploitation by both small and large landowners of mate tea trees (*Ilex paraguariensis*) which causes canopy openings, and shifting cultivation practices that sometimes leave fallow lands more than 10 years to regenerate tree cover (Liebsch et al., 2008; Teixeira et al., 2009). The resulting landscape patterns have unique spectral and spatial characteristics that are difficult to classify correctly with medium resolution imagery. The areal extent of these patterns could easily have been underestimated by the Atlas 2008 and PROBIO maps and overestimated by LCF/SAR and PPMA maps. Fourth, none of the studies established a clear a priori definition of forest. Ribeiro et al. (2009) defined forest as arborous vegetation having regenerated at least 15 years in the past and with canopy height of at least 10 m because these conditions could be detected with medium resolution imagery. This definition was confirmed by the IFFSC field work which found that 97.9% of all sampled forest remnants mapped as forest by medium-resolution imagery had field-measured canopy height of at least 10 m and basal area of at least 10 $m^2 \cdot ha^{-1}$ (Vibrans et al., 2011, 2012). Thus, forest regrowth with canopy height and basal area less than these limits must be assessed by means other than classification of medium resolution imagery.

Attention should be drawn to the fact that the more conservative forest cover estimates were obtained from the Atlas 2008 map, which was constructed by a non-governmental organization (NGO), and the PROBIO map, which was constructed by a university under the direction of the Federal Ministry of Environment. Conversely, the more liberal estimates were obtained from the LCF/SAR and PPMA maps which were constructed by external, commercial enterprises on behalf of the regional government. A more cautious, "safety first" approach in the selection of classification thresholds and accuracy standards and in image processing, interpretation and classification steps by the Atlas 2008 and PROBIO studies could have produced more omission than commission errors (Fig. 2). Exclusion of all doubtful image features from the forest class (Section 3.1), without clear field evidence, could have been motivated by either conscious or unconscious preconceptions regarding

Table 6
Forest cover estimates for Santa Catarina and by forest types where SRS denotes simple random sampling estimation and MAR denotes model-assisted estimation.

Information source	Estimator	Area (km^2)	Santa Catarina	Forest type					
			Proportion forest	Mixed ombrophilous forests (MIX)		Dense rain forests (DEN)		Deciduous forests (DEC)	
				Area (km^2)	Proportion	Area (km^2)	Proportion	Area (km^2)	Proportion
IFFSC	$\hat{p}_{SRS}$	26,338	0.277	12,317	0.220	12,633	0.404	1251	0.163
LCF/SAR	$\hat{p}_{MAR}$	31,326	0.329	14,663	0.261	14,747	0.471	1921	0.250
PROBIO	$\hat{p}_{MAR}$	30,563	0.321	15,492	0.274	13,493	0.436	1600	0.209
Atlas 2008	$\hat{p}_{MAR}$	27,555	0.289	13,741	0.244	12,619	0.405	1231	0.161
PPMA	$\hat{p}_{MAR}$	35,092	0.368	19,268	0.340	14,019	0.441	1813	0.236

forest cover, nature and landscape conservation issues. The opposite approach could have produced large commission errors in the more liberal (optimistic) estimates from the LCF/SAR and PPMA maps. In fact, the smaller estimates of proportion forest cover represent a powerful argument for advocates of conservation-oriented environmental policy, whereas the larger estimates support development-oriented policy. The policy implications are important because they influence decisions on land use, new conservation units, plantation forestry, infrastructure projects such as highways and hydroelectric power plants and, in general terms, priorities for conservation and economics throughout the state. Thus, despite their differences, the maps at least partially satisfy the requirements for data at scales useful for land use planning (Fuller, 2006; Lefsky et al., 1999).

Finally, for forests consisting of formations with heights of at least 10 m and basal area of at least 10 $m^2 \cdot ha^{-1}$, the study produced a 95% confidence interval of 0.270–0.309 for a conservative estimate of Santa Catarina's proportion forest cover (Table 5). However, the study also produced good evidence in the form of a large overall map accuracy for a 95% confidence interval of 0.352–0.384 for a more optimistic estimate (Fig. 3).

5. Conclusions

Four conclusions may be drawn from the study. The first conclusion, which is specific to Santa Catarina, is that overall accuracies for the four maps were large and generally similar. Thus, the maps may serve as one source of relevant source information for land use planning. Second, the model-assisted regression estimator facilitated adjustment of estimates to compensate for estimated bias resulting from classification errors, estimation of variances, and construction of confidence intervals. Other than the map predictions, the only information necessary for estimation and inference using the model-assisted estimator is available in an error matrix. This conclusion is relevant not only for this study but also for other studies that estimate forest attributes from maps, particularly for obtaining baseline estimates in accordance with the IPCC Good Practice Guidance. Third, the large overall accuracies produced standard errors for map-based estimates of proportion forest that were smaller than for the plot-based estimate. Thus, the maps contributed to increasing the precision of estimates of proportion forest cover. Fourth, a priori clarification of the definition of forest is crucial for study areas with biological diversity and anthropogenic history similar to that of the Atlantic forest. Failure to clarify the definition could lead to the conclusion that remote sensing techniques are faulty or inadequate when, in fact, observed deviations are the result of a priori assumptions or prejudices.

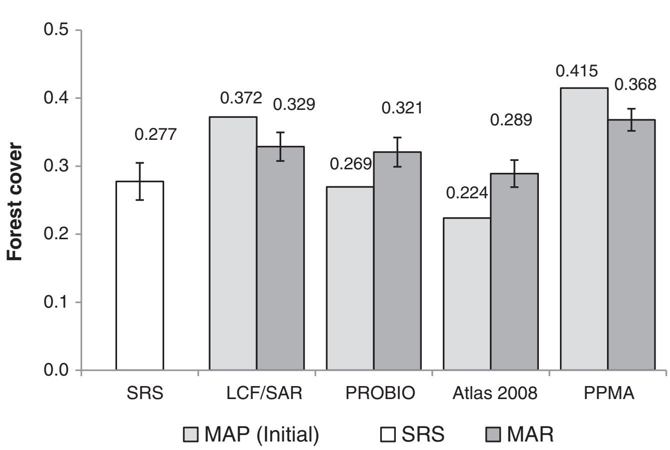

Fig. 3. Estimates of forest cover for Santa Catarina based on simple random sampling estimation (SRS, Eq. (1)) with IFFSC ground data, initial map-based estimation (MAP, Eq. (6)) and model-assisted estimation (MAR, Eq. (8)); vertical bars denote 95% confidence intervals.

Acknowledgments

The authors are grateful to the Fundação de Amparo à Pesquisa e Inovação de Santa Catarina (FAPESC) for supporting the IFFSC.

References

Achard, F., Eva, H. D., Stib, H. -J., Mayzux, P., Gallego, J., Richards, T., et al. (2002). *World atlas of biodiversity*. Berkeley: University of California Press (340 pp.).

Anderson, J., Hardy, E., Roach, J., & Witmer, R. (1976). *A land use and land cover classification system for use with remote sensing data. U.S.G.S. Professional Paper*, *964*, Government Printing OfficeWashington, DC (34 pp.).

Carvalho, L. M. T., & Scolforo, J. R. (2008). *Monitoramento da Flora Nativa 2005–2007.* Lavras: Editora UFLA (357 pp.).

Coppin, P., Jonckheere, I., Nackaerts, K., & Muys, B. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing*, *25*(9), 1565–1596.

Corlett, H. (1995). Tropical secondary forests. *Progress in Physical Geography*, *19*(2), 159–172.

Cruz, C. B. M., & Vicens, R. S. (2007). *Levantamento da Cobertura Vegetal Nativa do Bioma Mata Atlântica.* Relatório Final. Rio de Janeiro: IESB/IGEO/UFRJ/UFF (84 pp.).

Dinerstein, E., Olson, D. M., Graham, D. J., Webster, A. V., Primm, S. A., Bookbinder, M. P., et al. (1995). *Una evaluación del estado de conservación de las ecoregiones terrestres de América Latina y el Caribe. Publicado en colaboración con el Fondo Mundial para la Naturaleza.* Washington, D.C.: Banco Mundial.

Fehse, J., Hofstede, R., Aguirrea, N., Paladinesa, C., Kooijmanb, A., & Sevinka, J. (2002). High altitude tropical secondary forests: A competitive carbon sink? *Forest Ecology and Management*, *63*, 9–25.

Finnegan, B. (1996). Pattern and process in neotropical secondary rain forests: the first 100 years of sucession. *Tree*, *11*(3), 119–124.

Fuller, D. O. (2006). Tropical forest monitoring and remote sensing: A new era of transparency in forest governance? *Singapore Journal of Tropical Geography*, *27*, 15–29.

Fundação S.O.S Mata Atlântica (2009). Atlas dos remanescentes florestais da Mata Atlântica, período 2005–2008. *Relatório Final*. São Paulo: Fundação S.O.S. Mata Atlântica/Instituto Nacional de Pesquisas Espaciais (156 pp.).

Geoambiente Sensoriamento Remoto Ltda (2008). *Projeto de Proteção da Mata Atlântica em Santa Catarina (PPMA/SC).* Relatório Técnico do Mapeamento Temático Geral do Estado de Santa Catarina. São José dos Campos (90 pp.).

Giardin, C. (2010). UN-REDD/FAO to publish national forest MRV system recommendations. *UN-REDD Programme Newsletter, Issue 8, May 2010* (Available at: http://www.un-redd.org/Newsletter8_MRV_System_Recommendations/tabid/4551/language/en-US/Default.aspx. Last accessed: October 2012).

GOFC-GOLD (2012). A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals associated with deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation. GOFC-GOLD Report version COP18-1, (GOFC-GOLD Land Cover Project Office, Wageningen University, The Netherlands) (Available at: http://www.gofcgold.wur.nl/redd/index.php. Last accessed: November 2012).

Groeneveld, J., Alvesc, L. F., Bernacci, L. C., Catharino, E. L. M., Knogge, C., Metzger, J. P., et al. (2009). The impact of fragmentation and density regulation on forest succession in the Atlantic rain forest. *Ecological Modelling*, *220*, 2450–2459.

Groombridge, B., & Jenkins, M. D. (2002). *World atlas of biodiversity.* Berkeley: University of California Press (340 pp.).

Gullison, R. E., Frumhoff, P. C., Canadel, J. G., Field, C. G., Nepstad, D. C., Hayhoe, K., et al. (2007). Tropical forests and climate policy. *Science*, *316*, 985–986.

Hayes, C. J., & Cohen, W. B. (2007). Spatial, spectral, and temporal patterns of tropical forest cover change as observed with multiple scales of optical satellite imagery. *Remote Sensing of Environment*, *106*, 1–16.

Holdridge, L. R. (1947). Determination of world plant formation from simple climatic data. *Science*, *105*, 367–368.

Holdridge, L. R. (1967). *Life zone ecology*. San Jose, CA: Tropical Science Center (206 pp.).

Klein, R. M. (1978). *Mapa fitogeográfico do estado de Santa Catarina*. Itajaí: SUDESUL, FATMA, HBR. (Flora Ilustrada Catarinense, 5). 24 pp.

Lefsky, M. A., Cohen, W. B., Hudak, A., Acker, S. A., & Ohmann, J. (1999). Integration of lidar, Landsat ETM+ and forest inventory data for regional forest mapping. *International Archives of Photogrammetry and Remote Sensing*, *32*(Part 3W14), 119–126.

Liebsch, D., Marques, M. C. M., & Goldenberg, R. (2008). How long does the Atlantic Rain Forest take to recover after a disturbance? Changes in species composition and ecological features during secondary succession. *Biological Conservation*, *141*, 1717–1725.

Matthews, E., Payne, R., Rohwede, M., & Siobahn, M. (2000). *Pilot analysis of global ecosystems: Forest ecosystems.* Washington, DC: World Resources Institute (86 pp.).

McRoberts, R. E., & Walters, B. F. (2012). Statistical inference for remote sensing-based estimates of net deforestation. *Remote Sensing of Environment*, *124*, 394–401.

Oliveira, T. C. A., Carvalho, L. M. T., Oliveira, L. T., Martinhago, A. Z., Acerbi, F. W., Jr., & de Lima, M. P. (2010). Mapping deciduous forests by using time series of filtered MODIS NDVI and Neural Networks. *Cerne*, *16*(2), 123–130.

Oliveira-Filho, A. T., & Fontes, M. A. (2000). Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate. *Biotropica*, *32*(4b), 793–810.

Olschewskia, R., & Benítez, P. C. (2005). Secondary forests as temporary carbon sinks? The economic impact of accounting methods on reforestation projects in the tropics. *Ecological Economics*, *55*, 380–394.

Penman, J., Gytarsky, M., Hiraishi, T., Krug, T., Kruger, D., & Pipatti, R., et al. (Eds.). (2003). *Intergovernmental panel on climate change, good practice guidance for land use, land-use change and forestry*. Hayama, Kanagawa, Japan: Institute for Global Environmental Strategies (Available at: http://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_files/GPG_LULUCF_FULL.pdf. Last accessed: October 2012).

Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. (2009). The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, *142*, 1141–1153.

SAR. Secretaria de Agricultura e Abastecimento do Estado de Santa Catarina. (2005). *Inventário Florístico Florestal de Santa Catarina*. Relatório do Projeto Piloto. Florianópolis. (mimeo). 170 pp.

Särndal, C. -E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer (693 pp.).

Scarano, F. R. (2009). Plant communities at the periphery of the Atlantic rain forest: rare-species bias and its risks for conservation. *Biological Conservation*, *142*, 1201–1208.

Silva, J. M. C., & Casteleti, C. H. M. (2003). Status of the biodiversity of the Atlantic Forest of Brazil. In C. Galindo-Leal, & I. G. Câmara (Eds.), *The Atlantic Forest of South America: Biodiversity status, trends, and outlook* (pp. 43–59). Washington, DC: Island Press.

Steininger, M. K. (2000). Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia. *International Journal of Remote Sensing*, *21*(6–7), 1139–1157.

Teixeira, A. M. G., Soares-Filho, B. S., Freitas, S. R., & Metzger, J. P. (2009). Modeling landscape dynamics in an Atlantic Rainforest region: Implications for conservation. *Forest Ecology and Management*, *257*, 1219–1230.

Tucker, C. J., & Townshend, J. R. G. (2000). Strategies for monitoring tropical deforestation using satellite data. *International Journal of Remote Sensing*, *21*(6–7), 1461–1471.

Veloso, H. P., Rangel Filho, A. L. R., & Lima, J. C. A. (1991). *Classificação da vegetação brasileira, adaptada a um sistema universal*. Rio de Janeiro: IBGE (123 pp.).

Vibrans, A. C., Moser, P., Lingner, D. V., & Maçaneiro, J. P. (2012). Análise estatística do IFFSC e estimativas dendrométricas. In A. C. Vibrans, L. Sevegnani, A. L. Gasper, & D. V. Lingner (Eds.), *Inventário Florístico Florestal de Santa Catarina. Vol. I. Diversidade e conservação dos remanescentes florestais*. Blumenau: Edifurb. 452 pp.

Vibrans, A. C., Sevegnani, L., Lingner, D. V., Gasper, A. L., & Sabbagh, S. (2010). Inventário Florístico Florestal de Santa Catarina (IFFSC): aspectos metodológicos e operacionais. *Pesquisa Florestal Brasileira*, *30*(64), 291–302.

Vibrans, A. C., Sevegnani, L., Uhlmann, A., Schorn, L. A., Sobral, M. G., Gasper, A. L., et al. (2011). Structure of mixed ombrophyllous forests with *Araucaria angustifolia* (Araucariaceae) under external stress in southern Brazil. *Revista de Biologia Tropical*, *59*(3), 1371–1387.