



Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier



Clovis Grinand ^{a,b,c*}, Fety Rakotomalala ^{c,d}, Valéry Gond ^e, Romuald Vaudry ^c, Martial Bernoux ^g, Ghislain Vieilledent ^{e,f}

^a ETC Terra, 127 Rue d'Avron, 75020 Paris, France

^b ED SIBAGHE, Montpellier SupAgro, 2 Place Viala, 34060 Montpellier Cedex 2, France

^c ETC Terra – Madagascar, Lot VE 26 L, Ambanidilia, 101 Antananarivo, Madagascar

^d IOGA, Institut et Observatoire de Géophysique d'Antananarivo, BP 3843, 101 Antananarivo, Madagascar

^e Cirad, UPR BSEF, F-34398 Montpellier, France

^f Cirad – Madagascar, DP Forêt et Biodiversité, 101 Antananarivo, Madagascar

^g IRD, Eco&Sols, 2 Place Viala, F-34060 Montpellier, France

ARTICLE INFO

Article history:

Received 6 February 2013

Received in revised form 9 July 2013

Accepted 10 July 2013

Available online xxxx

Keywords:
Deforestation
Change detection
Classification
Land cover
Landsat TM
Machine learning
Madagascar
Random forests
REDD +

ABSTRACT

High resolution and low uncertainty deforestation maps covering large spatial areas in tropical countries are needed to plan efficient forest conservation and management programs such as REDD+ (Reducing Emissions from Deforestation and Forest Degradation). Using an open-source free software (R, GRASS and QGis) and an original statistical approach combining multi-date land cover observations based on Landsat satellite images and the random forests classifier, we obtained up-to-date deforestation maps for the periods 2000–2005 and 2005–2010 with a minimum mapping unit of 0.36 ha for 7.7 M hectares, i.e. 40.3% of the tropical humid forest and 20.6% of the tropical dry forest in Madagascar. Uncertainty in deforestation on the maps was calculated by comparing the results of the classification to more than 30,000 visual interpretation points on a regular grid. We assessed accuracy on a per-pixel basis (confusion matrix) and by measuring the relative surface difference between wall-to-wall approach and point sampling. At the pixel level, user accuracy was 84.7% for stable land cover and 60.7% for land cover change. On average for the whole study area, we obtained a relative difference of 2% for stable land cover categories and 21.1% land cover change categories respectively between the wall-to-wall and the point sampling approach. Depending on the study area, our conservative assessment of annual deforestation rates ranged from 0.93 to 2.33·yr^{-1} for the humid forest and from 0.46 to 1.17·yr^{-1} for the dry forest. Here we describe an approach to obtain deforestation maps with reliable uncertainty estimates that can be transposed to other regions in the tropical world.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Assessing changes in tropical forest cover is a challenging research and operational topic that addresses issues like climate change, biodiversity conservation and sustainable ecosystem management. Global net loss of forest area has been estimated at 5.2 million hectares per year between 2000 and 2010 (FRA, 2010) and, according to the most recent studies (Baccini et al., 2012; Harris et al., 2012), is responsible for 10% to 25% of anthropogenic greenhouse gas (GHG) emissions to the atmosphere. Tropical deforestation is a major contributor to GHG emissions because of the extent of forest being cleared each year and because of the high carbon stock per unit area (Achard et al., 2010). The REDD+ mechanism (Reducing Emissions from Deforestation and

Degradation) in the United Nations Framework Convention on Climate Change (UNFCCC) aims to encourage developing countries to slow down deforestation through a compensation mechanism. This mechanism requires accurate, transparent, and cost-effective GHG measurement and monitoring systems (Olander, Gibbs, Steininger, Swenson, & Murray, 2008). Despite well established guidelines provided by the international scientific community (GOFC-GOLD, 2010; IPCC, 2006), improvements in remote sensing techniques and data availability are still necessary to provide more accurate and cost-effective estimates of forest change.

1.1. Techniques for estimating deforestation over large forest areas

Recent reviews of methods used to estimate deforestation highlight two different methods of monitoring large areas of forest: the exhaustive mapping of forest extent (also known as the “wall-to-wall” approach) or point sampling (Achard et al., 2010; GOFC-GOLD, 2010;

* Corresponding author at: ETC Terra, 127 Rue d'Avron, 75020 Paris, France. Tel.: +33 647 32 04 37.

E-mail address: c.grinand@etcterra.org (C. Grinand).

Hansen & Loveland, 2012). The wall-to-wall approach uses satellite images (such as Landsat (Harper, Steininger, Tucker, Juhn, & Hawkins, 2007; Gutman et al., 2005), AVHRR (DeFries et al., 2002), MODIS (Freidl et al., 2002) or MERIS (Bicheron et al., 2008)) to map the full extent of forest and changes in forest area through hybrid techniques that combine automated digital segmentation and/or classification with visual interpretation (GOFC-GOLD, 2010, Rasi et al., 2012) or fully automated techniques (Hansen et al., 2008; Huang et al., 2009; Potapov et al., 2012). Point sampling approach consists in identifying partial but representative land cover observations through ground surveys or through visual interpretation of satellite images to estimate forest and deforestation extent over the entire area (Rasi et al., 2012; Steininger, Godoy, & Harper, 2009).

Point sampling reduces the operational cost of exhaustive analysis of a large number of satellite images, while improving the thematic accuracy of (for example) regrowth or degradation assessment by focusing on small areas (Achard et al., 2002; Duveiller, Defourny, Desclée, & Mayaux, 2008; FRA, 2010; Rasi et al., 2011; Stach et al., 2009). However, the accuracy of point sampling is closely linked to the quality of the sampling design (Steininger et al., 2009) and does not enable the production of deforestation maps of the entire forest area. Such maps are however required to target conservation efforts and to allow consistent spatial analysis of deforestation (Vieilledent, Grinand, & Vaudry, 2013).

Classical wall-to-wall land cover change analysis implies pairwise image comparison. This method of detecting change has been widely used (Harper et al., 2007; Stach et al., 2009) and provides useful transition matrices describing change from one land category at date 1 to another land category at date 2 (Duveiller et al., 2008; Huang et al., 2009). But when two single date maps are used in combination to derive land cover change, individual errors will be multiplied if errors on the two maps are assumed to be independent (Fuller, Smith, & Devereux, 2003). This seriously undermines the accurate detection of subtle change in forest area, and it is consequently recommended to combine satellite images acquired at many different dates in a single analysis that identifies change directly (GOFC-GOLD, chap. 2.6, 2010). In other words, supervised classification is performed on stacked images acquired at many different dates. This technique allows seasonal variations and vegetation dynamics to be taken into account (Pennec, Gond, & Sabatier, 2011). For instance, slash-and-burn plots may subsequently have a different cover (a permanent crop, fallow, secondary regrowth) and secondary forest, forested fallow and intact forest may be confused depending on the selected date. Thus, increasing the number of dates using time series of satellite images reduces the uncertainty associated with land cover change classification.

1.2. Classification algorithms

Open access to the 20-year Landsat archive (called the Landsat Global Land Survey Program) has greatly reduced the cost and facilitated the processing of large time series for the estimation of land cover change (Hansen & Loveland, 2012). Given that multi-date satellite images are needed to reduce the uncertainty associated with land cover change classification, the usual parametric classification algorithms (such as classification by maximum likelihood) may not be appropriate for the classification of combined multi-date images because of the heterogeneous spectral signature of land cover categories over large areas. To overcome this problem, data mining and machine learning techniques (including neural networks, decision trees, support vector machines and ensemble classifiers) have recently emerged in remote sensing, making it possible to deal with complex land cover status and dynamics. Such algorithms are efficient because they do not rely on the data distribution assumption (e.g. normality), are able to handle noisy observation (Breiman et al., 2001), and can be efficiently applied to large complex datasets if supported by sufficient training data (Rodríguez-Galiano, Ghimire, Rogan, Chica-Olmo, & Rigol-Sánchez, 2012). Among these algorithms, the random forests (RF) classifier has already provided

interesting results in several studies using satellite images. For instance, Schneider (2012) reported that RF outperformed maximum likelihood classifier and was better than a support vector machine classification based on accuracy assessment and visual interpretation of the resulting maps. Previous RF applications include land cover classification using Modis data (Aide et al., 2012; Clark, Aide, Grau, & Riner, 2010), Landsat data (Gislason, Benediktsson, & Sveinsson, 2006; Lawrence, Bunn, Powell, & Zambon, 2004; Rodriguez-Galiano et al., 2012; Schneider, 2012) or hyperspectral data (Ham, Chen, Crawford, & Gosh, 2005), digital soil mapping (Grimm, Behrens, Märker, & Elsenbeer, 2008) and forest biomass mapping (Baccini et al., 2012).

1.3. Accuracy assessment of land cover change

Accuracy assessment must be associated with deforestation maps, as some authors consider errors in land cover change to be the main source of error in estimates of GHG emissions (Harris et al., 2012; Pelletier, Ramankutty, & Potvin, 2011). A common approach for accuracy assessment implies computing an error matrix from independent ground survey observations or from visual interpretation of high resolution images to quantify class specific accuracies and overall map accuracy. A probability based sampling design is required for the accurate error estimation of land change categories (Stehman, 2009). Assessing a land cover map is known to be difficult, but assessing land cover change maps is even more challenging (Hansen & Loveland, 2012) mainly due to the difficulty in obtaining accurate land cover change reference datasets (Foody, 2010). Field surveys of historical change are tricky since they involve questioning people with a deep knowledge of the plot's history (e.g. the landowner). Ground truth observation via remote sensing is also constrained by the availability and resolution of the images. Another problem with change detection assessment is the extent of the change. Fuller et al. (2003) estimate that "small to medium scale changes requires levels of precision in mapping which are near impossible to achieve [...] unless it (the survey) is tailor made and rigorously applied to the recording of change". For all these reasons, accurate assessment of land cover change is not a trivial task and requires particular attention.

1.4. Objectives

Madagascar is universally recognized for its high level of biodiversity and endemism (Goodman & Benstead, 2005). The biodiversity is mainly located in its tropical forests. In the last 50 years, Madagascar has experienced a dramatic loss of forest (Harper et al., 2007) due to traditional slash and burn, pasture extension, charcoal production, illegal logging of precious wood, and mining activities, with consequences for both biodiversity conservation and GHG emissions. To curb deforestation, Madagascar is highly committed to the implementation of the REDD+ program through both the development of REDD+ pilot projects at regional scale and policy decisions at national scale. Previous work on land cover and land cover change mapping has provided valuable insight into vegetation status (Mayaux, Gond, & Bartholomé, 2000; Mayaux, Bartholomé, Fritz, & Belward, 2004; Moat & Smith, 2007) and dynamics (Harper et al., 2007). However these insights cannot be used directly since the definition of forest varies widely among studies and does not provide sufficient levels of information (a resolution of 1 km or a minimum mapping unit of 2 ha) for the detection of subtle change. Using Madagascar as a case study, we first used the RF algorithm to map land cover change using a 10-year time series of Landsat TM satellite images covering large forest areas. Second, we calculated up-to-date annual deforestation rates for the period 2000 to 2010, and third, we assessed the accuracy of forest cover change using more than 30,000 visual interpretation points on a regular grid. The specific aim of this study was to develop a transparent and cost-effective methodology to obtain reliable deforestation maps associated with

uncertainty estimates that can be easily transposed to other regions in the tropical world prone to subsistence agriculture.

2. Material and method

2.1. Study area

The study area was divided into five separate forest zones in Madagascar (Fig. 1) covering a total area of 7.7 million hectares. According to the most recent vegetation map (Moat & Smith, 2007), the study areas include 2,407,000 ha of tropical forest divided into 372,000 ha of spiny-dry forest (with precipitation < 1000 mm·yr⁻¹) and 2,035,000 ha of humid forest (with precipitation > 1000 mm·yr⁻¹) in 2005. Except for the dry region, all of these sites are located in hilly areas with medium to steep slopes, which is the only remaining typical intact humid forest landscape in Madagascar. Slash and burn and shifting cultivation are among the main agricultural practices found everywhere.

Each of the study areas – or reference areas in the REDD+ jargon – encompasses project sites of a REDD+ pilot program called Holistic Programme for Conservation of Forests in Madagascar (Fig. 1). The study areas were manually delineated using local knowledge in order to include a large forested area with homogeneous bio-physical conditions (e.g. climate or topographic range), similar to those of the project

sites. The project sites cover more than 500,000 ha of forest out of a total of 850,000 ha concerned by the creation of new protected areas, forest restoration and promotion of sustainable agriculture as an alternative to slash-and-burn practices. The extent, the project sites and the study areas, were used to produce estimates of deforestation rates.

2.2. Landsat image collection and pre-processing

The steps of the procedure are summarized in Fig. 2. Multi-temporal studies require the use of geometrically consistent images. Standard orthorectified Landsat image databases are available for the years 1975, 1990, and 2000 worldwide (Tucker, Grant, & Dykstra, 2004). Unfortunately, in our study area, the GLS 2005 Landsat product (Gutman et al., 2005) was mainly based upon Landsat 7 which has high rates of no data values (stripping) due to its failure in 2003. For this reason, Landsat TM images were collected for reference dates in 2005 and 2010. We selected images with the lowest cloud cover and which were available in a time frame of plus or minus one year compared to the reference date. These criteria were met for every site except the northern site, which had a large permanent cloud cover. For this particular site, we looked for suitable images from 1990. From the USGS EROS portal, level 1 terrain corrected (L1T) images for 2005 and 2010 were downloaded and all images were visually inspected for

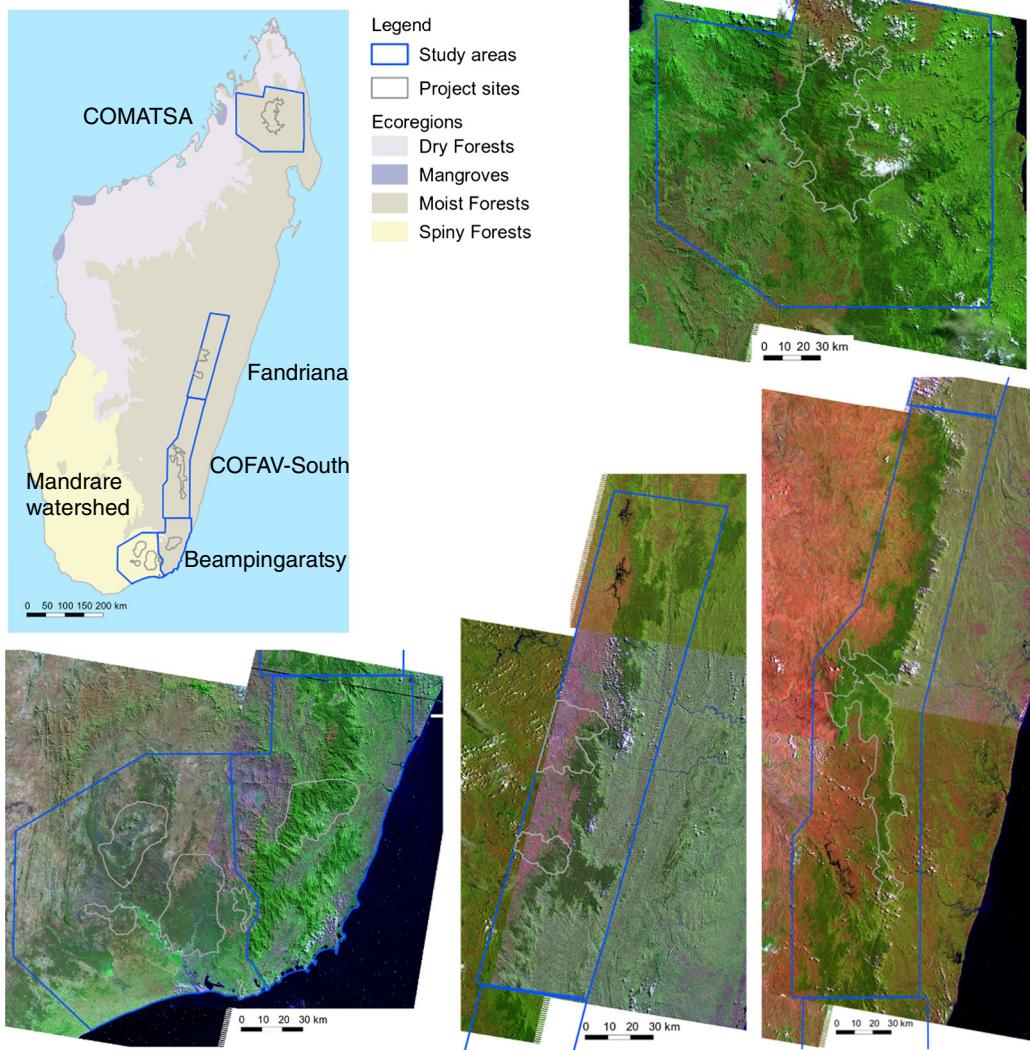


Fig. 1. Location of the study areas and project sites in Madagascar. Background satellite images are Landsat 5 from circa 2005 in a false color composite display, RGB = 542. The source of the eco-region map is the Global 200 World Wide Fund project (http://en.wikipedia.org/wiki/Global_200).

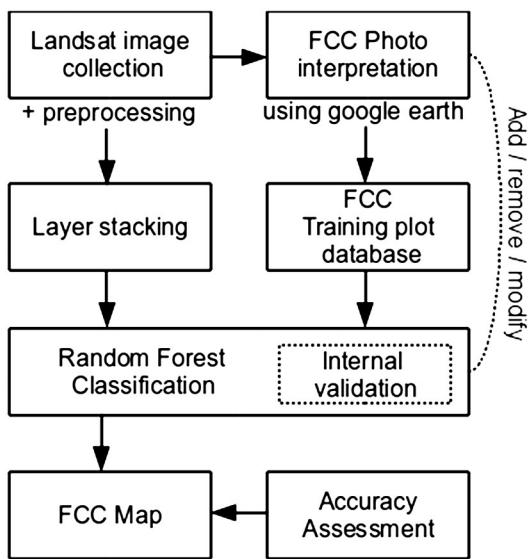


Fig. 2. Flow diagram of the processing steps. FCC stands for Forest Cover Change.

geometric precision based on the GLS 2000 image. The data showed a good overall geometric precision with less than 1 pixel offset compared with the 2000 image. Radiometric normalization was not performed, since this step is not required when images are analyzed simultaneously rather than individually (Song, Woodcock, Seto, Pax-Lenney, & Macomber, 2001 cited in Schneider, 2012) and because decision tree algorithms can handle variables on a relative scale. We produced mosaic images for each date and each site by simple overlay, with the least cloudy image at the top of the mosaic. Finally, we derived a normalized difference vegetation index (NDVI) and a normalized infrared index (NIRI) from the mosaic images to obtain an enhanced auxiliary dataset. The Landsat image database processed in this study is presented in Table 1.

2.3. Classification of land cover change

2.3.1. Collection of training plot data

Accurate training plot data are essential for supervised classification. Several studies have shown that non-parametric machine learning algorithms, such as random forests, need a larger number of training data to attain optimal results (Potapov et al., 2012; Rodriguez-Galiano et al., 2012; Schneider, 2012). We conducted a free selection of training plots to obtain the most representative coverage of land cover and land cover change either in the number of features or spatial distribution. Visual interpretation of many small subsets of images was performed where a cluster of polygons representing the diversity of land cover,

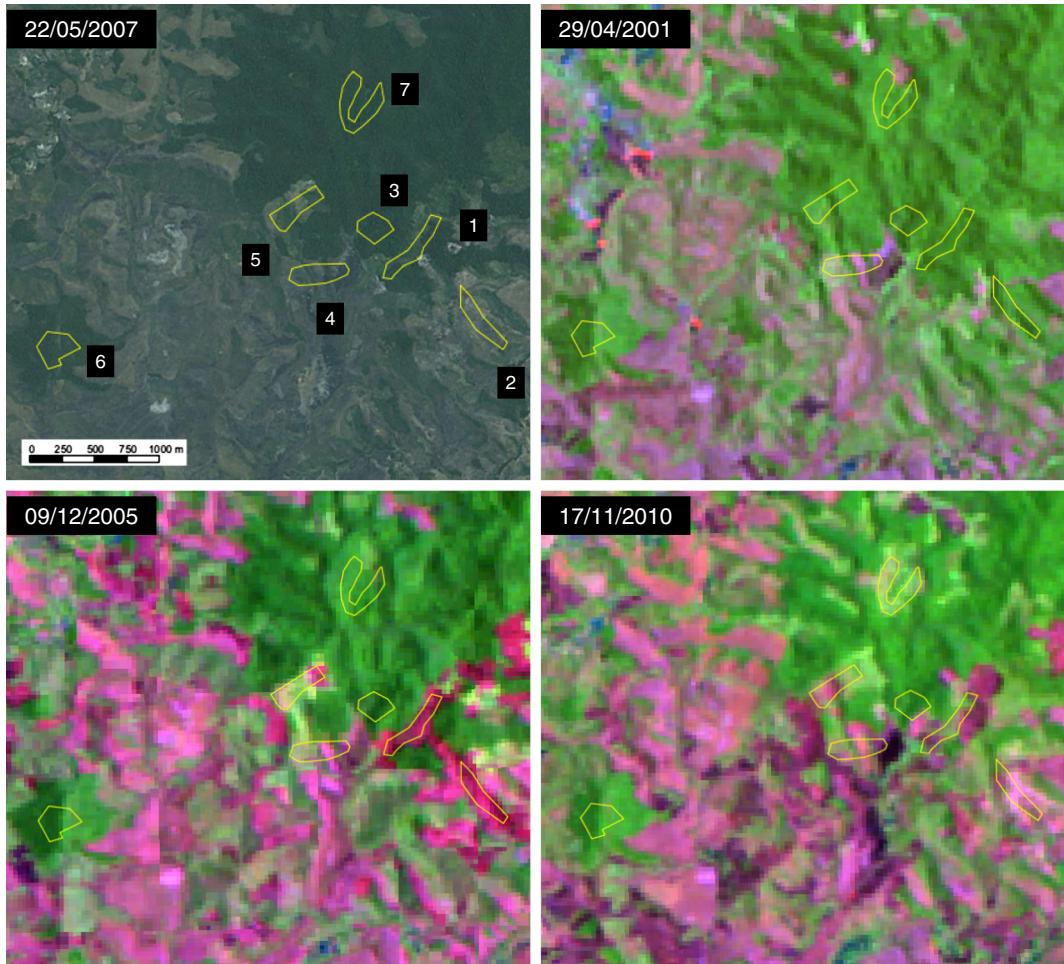
and land cover change was delineated in this particular area. By experience, we found that a scale of 1:10,000 was suitable to distinguish shifting cultivation cropland in a sufficiently large portion of the landscape. A subset of images corresponds approximately to what a photo interpreter can see in the screen display at a scale of 1:10,000. An example of visual interpretation of cluster of training plot is presented in Fig. 3. Additional sources of information, including freely available QuickBird images from Google Earth and expert knowledge were used as reference material to help interpretation. We used six land cover categories (humid forest, dry forest, cropland-savannah, wetland, rock-bare soil, cloud and shadow) which were observed for each date. For each land cover, a minimum of 20 training plots were targeted. Next, we aggregated the observation into one single forest cover change nomenclature, including six land cover categories over the 2000–2010 period and four land cover change categories (deforestation of humid and dry forest for both monitoring periods). For land cover change categories, efforts were made to capture the entire extent of the deforestation patches, contrary to land cover categories for which random or odd shapes were used. Specific rules were applied to reduce the extent of clouds and shadows. For instance, a plot observed to be forested at date 1, cloudy at date 2, and forested at date 3 was considered to be forested throughout the historical period. Although we were easily able to deduce the fate of forest or non-forest area in many cases, this was not possible in all cases. For instance, a cloud-shadow classification was used when cloud or shadow was observed at date 2, while forest was observed at date 1 and non-forest at date 3. This method enabled us to reduce the extent of clouds and shadows without entirely suppressing them.

2.3.2. Classification with random forests

Several authors have shown that land cover classifications with RF outperform classifications with other types of algorithms such as the classification by maximum likelihood and by support vector machines (Gislason et al., 2006; Schneider, 2012). RF is a decision tree algorithm considered to be an improved version of the Classification and Regression Tree (CART, Breiman, Friedman, Olshen, & Stone, 1984). Basically, it randomly selects a sample of observations and a sample of variables many times (default value of 500) to produce a large number of small classification trees. These small trees are then aggregated and a majority vote rule is applied to determine the final category (Breiman, 2001). This algorithm can be considered as a black box, as no clear class statistical signature is computed. However it provides an internal classification accuracy measurement (out-of-the-bag error) and the relative importance of the variables, which provides valuable insight into the modeled complex relationship (see Section 2.3.3). We developed models which use the selected reflective bands (Landsat TM band 1, 4 and 5) and indexes (NDVI and NIRI) for every date. These variables were selected for their well known land cover discriminating power and which, following some preliminary tests (not shown), led to no increase in accuracy when the remaining bands (2, 3 and 7) were added. Following recommendations from Rodriguez-Galiano et al. (2012), we used default parameters for the two main RF parameters (number of trees and the number of random split variables). We estimated classification accuracy using a confusion matrix obtained from an ad-hoc test procedure using 70% of the training plots as classification data and 30% of the training plots as test data. We optimized the training plot dataset until we reached maximum steady accuracy from this confusion matrix. This refinement task, also referred as internal validation (Fig. 2), was performed by adding new training plots to obviously misclassified locations or deleting some that may have been sources of confusion. This involved detailed analysis of the error matrix, training plot location, and map outputs at each classification run. Although all the categories were considered, we focused our attention on confusions related to land cover change categories. We finally used the entire training plot dataset to produce a FCC map of the whole study area.

Table 1
Landsat image database used in this study.

Region	Ref. scene	Date ~2000	Date ~2005	Date ~2010
COMATSA	158-070	22/08/1996	14/05/2006	08/09/2008
	159-070	29/06/1990	31/05/2004	30/04/2010
Fandriana	158-073	29/09/2001	24/03/2005	30/09/2009
	158-074	28/10/2000	25/04/2005	17/11/2010
COFAV-South	159-074	29/04/2001	10/01/2005	30/04/2010
	158-075	22/04/2001	03/08/2005	17/11/2010
Beampingaratsy	159-075	29/04/2001	09/12/2005	30/04/2010
	158-076	22/04/2001	03/08/2005	17/11/2010
Mandrare	159-076	29/04/2001	10/01/2005	08/11/2010
	158-077	13/09/2001	25/04/2005	17/11/2010
	159-077	28/05/2000	06/08/2006	11/01/2011
	158-077	13/09/2001	25/04/2005	17/11/2010
	159-077	28/05/2000	06/08/2006	11/01/2011



ID	2000	2005	2010	FCC000510	Legend
1	F	P	P	FPP	Deforested 2000-2005
2	F	P	P	FPP	Deforested 2000-2005
3	F	F	F	FFF	Forest 2000-2010
4	P	P	P	PPP	Non-Forest 2000-2010
5	F	P	P	FPP	Deforested 2000-2005
6	F	F	F	FFF	Forest 2000-2010
7	F	F	P	FPP	Deforested 2005-2010

Fig. 3. Stable land cover and examples of visual interpretation of land cover change using QGis, QuickBird/Google Earth images and Landsat images. From top left to bottom right: cluster of training plots overlaying QuickBird images from Google Earth (22/05/2007) and Landsat TM false color composite RGB = 4, 5, and 1 acquired at dates 29/04/2001, 9/12/2005 and 17/11/2010. Individual observations for each date and final forest cover change class for the seven training plots are presented in the table of attributes below the images.

2.3.3. Sampling intensity and importance of variables

The classification described above required a high level of expertise (visual interpretation of images and interpretation of confusion matrices), which can be fastidious and time-consuming when large areas are involved. For that reason, we investigated which sampling intensity would be appropriate to achieve satisfactory classification and whether all the variables used were relevant in the modeling process. First, we tested a sampling intensity ranging from 30% to 90% of the entire training plot dataset. Two metrics, the out-the-bag error provided by the RF algorithm and overall accuracy were calculated using the remaining test plots not used to build the model. The out-of-the-bag error is an internal accuracy metric calculated using about one third of the input dataset not used for the construction of the trees, and is considered as an unbiased estimate of the error (Breiman, 2001). Second, we evaluated the importance of the variables used for the final classification model. Two

measures exist in the RF algorithm to measure variable importance. We used only one, referred to as the mean decrease in accuracy. The latter applies variable permutation when constructing classification trees and measures the relative difference (decrease) in the out-of-the-bag error compared with a prior without-permutation test. A high decrease means the variable is important for the classification. Results are averaged on all trees and normalized on a 0–100 scale by the standard deviation of the differences.

2.3.4. Post processing

As previously mentioned in Section 2.3.1, clouds and shadows remained in the final output maps. The extent of clouds and shadows remaining in the raw classification output was less than 2.5% of the total study area for all sites except for the cloudy Comatsa site (14.1%). In order to produce a full coverage map and calculate accurate and

conservative deforestation estimates, we used an auxiliary map showing the extent of forest in 2005 (MEFT, 2009). Cloud and shadow pixels in the final FCC map were then replaced by forest or non-forest pixel values from the 2005 map (Fig. 4). In addition, we applied a minimum mapping unit filtering process using an in-house Grass script that has two folds. First, deforestation patches with an area of less than 0.36 ha (2×2 pixels) were replaced by the majority class found in a 3×3 moving window. Second, a simple majority vote using the same 3×3 moving window was applied for other categories. The 0.36 ha threshold was chosen with respect to the geometric precision of the input datasets (around 1 pixel). This approach was considered highly conservative (no overestimation of deforestation) regarding deforestation intensity as no deforestation pixels were added and small patches were removed. In another study (Steininger et al., 2009), Landsat images with a resolution of 28.5 m allowed the production of a minimum mapping unit of 0.3 ha.

2.3.5. Estimating deforestation rates

Deforestation rates can be expressed as the area deforested over a given period of time compared to the initial forest area (Puyravaud, 2003). Despite our objective of analyzing deforestation over a period spanning of five years, selected suitable images oscillated around the pivot years 2000, 2005 and 2010 (Table 1) resulting in variable time intervals between land cover observations. Time intervals ranged from three years and five months to six years and three months, with one exception for the north-east site, which showed a time interval of 13 years for the first period and three years for the second. To account for these large time differences, we used a weighted mean to estimate the average time interval between land cover observations for each study area (Vieilledent et al., 2013). The weights corresponded to the area (in km^2) of the scene with a given time interval. Following Puyravaud (2003), we used Eq. (1) to estimate the annual deforestation rate.

$$\theta = 1 - \left(1 - \frac{A_1 - A_2}{A_1} \right)^{\frac{1}{Y}}. \quad (1)$$

The annual deforestation rate θ was expressed as a function of A_1 and A_2 , the forest areas for the study areas at dates 1 and 2 and Y , the weighted average time interval. Y was expressed as a function of the number N of the different time intervals t_n (with $n = 1, \dots, N$) and the $Y = (1/N)$

$$\sum_n t_n \left(s_n / \sum_n s_n \right) \text{area of forest at date 1 with time-interval } t_n \text{ (denoted } s_n).$$

2.4. Accuracy assessment

Accuracy was assessed using a point sampling approach and following general guidelines, i.e. the use of a large sample and a low sample rate (Duveiller et al., 2008; Stehman, 2005). In this study, we used a systematic stratified sampling design with two strata (Fig. 5). The first stratum was defined to assess the extent of land cover by creating a 2-km grid, each intersection of the grid being a sample location. The second stratum was designed to capture smaller pattern such as deforestation, and a hot spot analysis was thus performed to delineate such areas. We used the fire information resource management system (FIRMS, <http://firefly.geog.umd.edu/firms>) to set up a fire location database for the last decade. We intersected the point database with the 2-km grid and selected the tile with more than five fires and within a distance of less than 500 m from the forest edge in 1990. The latter criterion was added in order to remove fire detection due to bushfire wherever possible. A 500-m grid was then created on these hot spot tiles, each intersection of the grid being a sample location. The two strata were merged and we finally obtained a total of 30,106 sampling points with 18,729 samples for stratum 1 (whole area) and 11,377 sample for stratum 2 (hot spots). At each point, the underlying pixel and its surroundings pixels (a total of 9 pixels) were visually interpreted by five photo interpreters who used the same protocol based on Qgis software. We particularly focused on the height of observation (scale) using zoom in and out functions to ensure the observation was made at pixel level. A majority vote rule was applied on the nine surrounding pixels to retain only one value for each point, thus corresponding to a sampling rate of 0.1% for the entire study area. Land cover and changes in land cover at the sampling points were classified by visual interpretation using the same method as that used for the FCC map, based on the same image mosaics. Accuracy was assessed using a confusion matrix that compared the visually interpreted sampling points and the FCC map, at the pixel level. We used the confusion matrix to estimate both the overall accuracy of the map and land category errors of commission and omission (Girard & Girard, 1999). In this study, we focused on category errors, which allowed us to derive the mean error of commission and omission for each category. This resulted in a first accuracy indicator referred to as per-pixel accuracy in this paper. We derived a second user-oriented indicator by calculating the relative difference in area derived from the point sampling approach and the FCC map. In the point sampling approach, areas were first derived using the formulas presented in Stach et al. (2009) and then using Eq. (1) to calculate deforestation rates. Relative differences in surface area were computed for the stable land cover categories and relative differences in the deforestation rate for the land cover change categories.

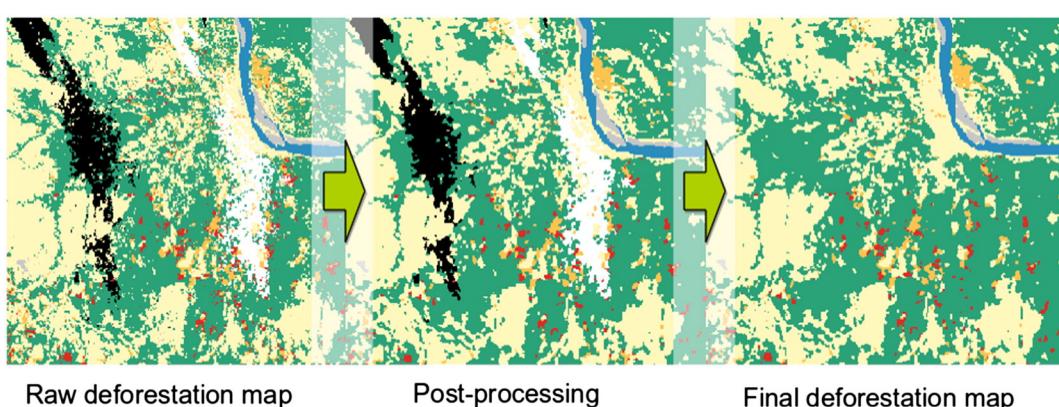


Fig. 4. Illustration of post-processing steps: 1) minimum mapping unit filtering of 0.36 ha using a 3×3 moving window that retains only 2×2 deforestation patches and applies a majority vote for other categories (first to second panel) and 2) cloud and shadow removal using the 2005 forest map (second to third panel). Green represents the extent of the forest, light yellow the non forest, orange deforestation in the 2000–2005 interval, and red deforestation in the 2005–2010 interval. White pixels represent clouds, and black pixels shadows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

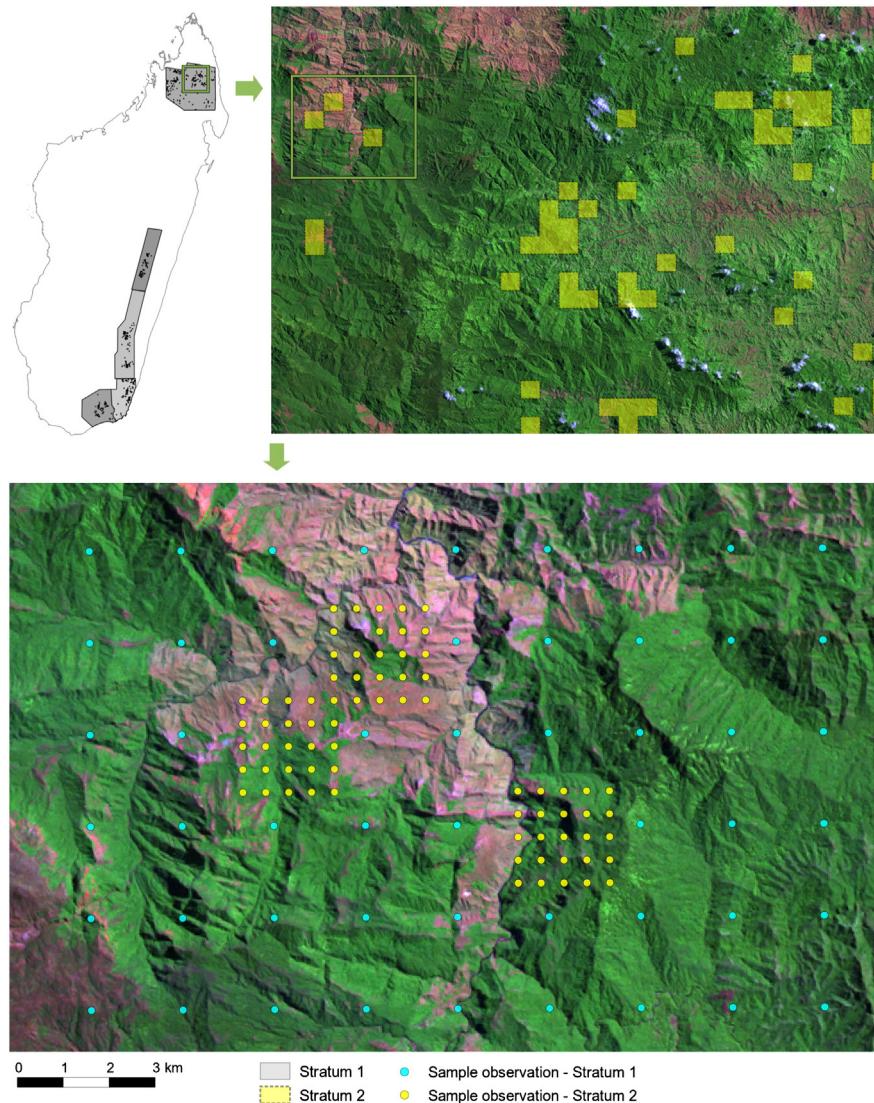


Fig. 5. Illustration of the point sampling design with the two strata used for the assessment of accuracy. Stratum 1 corresponds to one sample every 2 km over the entire study area (area in gray and blue dots) and stratum 2 corresponds to one sample every 500 m in previously located hot spot areas (area in yellow and yellow dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. Transparency and repeatability of the method

To achieve the maximum level of transparency and repeatability, efforts were made i) to minimize pre-processing steps, ii) to fully describe training plot polygons in a table of attributes, and iii) to use open source software (GRASS, QGis and R). These programs were used in a complementary fashion and, except for visual interpretation, all the steps were sequenced in a command line R-GRASS script available upon request from the authors.

3. Results

3.1. Sampling intensity and importance of the variables

Our primary goal was to test the ability of the RF algorithm to map small forest cover change over large areas in two time periods. Several interesting conclusions can be drawn from the calibration step using RF. First the results of the sampling intensity test revealed substantially different patterns for the out-of-the-bag accuracy computed by the algorithm and overall accuracy computed from the test dataset (Fig. 6 f,g,h,i,j). The former showed a relatively

homogeneous value of 0.95 with a maximum variation of 0.03 for the whole sampling intensity range, whereas the latter produced values ranging from 0.67 to 0.88. We further analyzed the trend of overall accuracy with an increasing number of plots used in the calibration. There was a clear improvement in accuracy with an increase in sampling intensity despite the fact the trend was not linear. It should be noted that at this stage, only overall accuracy was observed from the point of view of the producer and class level results will be presented from the point of view of the user in Section 3.2 below. These tests were carried out on the entire training plot dataset which was composed of 160 to 295 training plots per study area, 39–117 plots being interpreted as stable land cover and 22–51 plots as land cover change (Fig. 6 a,b,c,d,e). It should be noted that we observed only small patches of deforestation (Fig. 3), with a median size of 3.1 ha. Second, regarding the importance of the variables, results showed that certain Landsat bands played a more important role than others in the classification (Fig. 6 k,l,m,o,p). We noted the importance of band 5 and band 1 in every study area and particularly for the Comatsa, Fandriana and Mandrare watersheds. Surprisingly, the vegetation indexes (NDVI and NRI) did not have higher discriminative power than these two reflective bands. All the variables had more

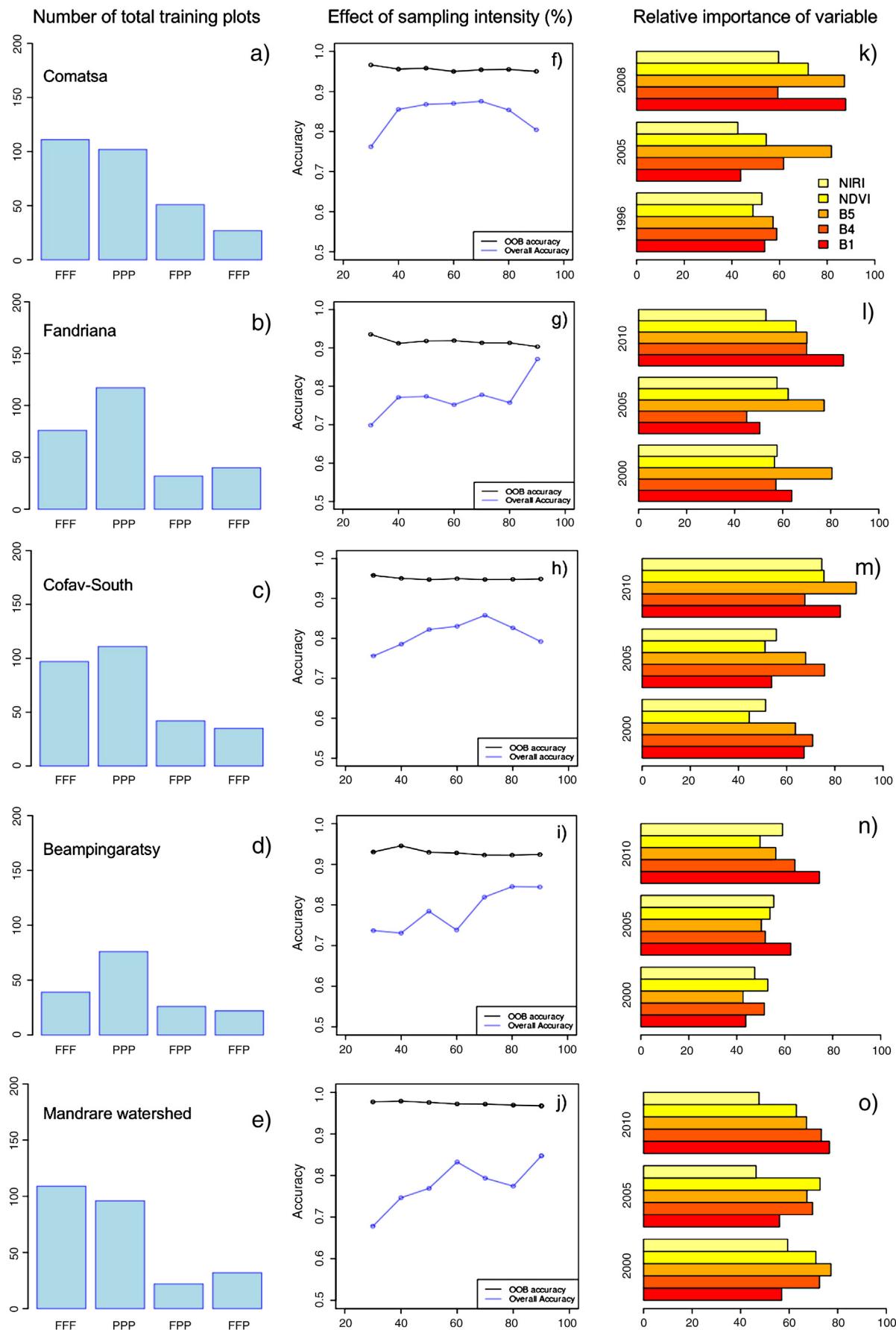


Fig. 6. Random forest calibration summary statistics. The first column corresponds to the number of overall training plots, the second column to the sampling intensity test and the third column to the relative importance of the variables.

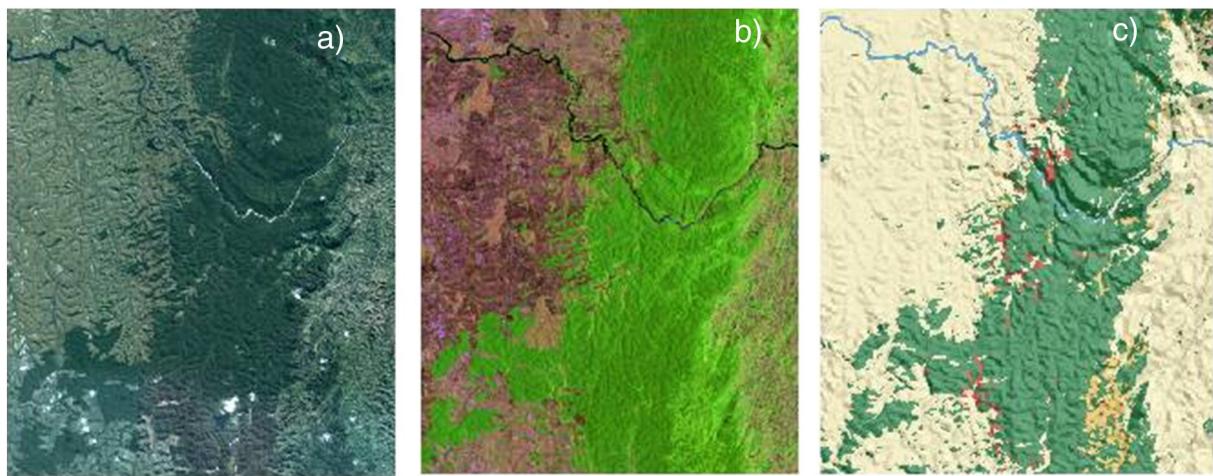


Fig. 7. Satellite images and final map of forest cover change for the periods 2000–2005–2010. a) QuickBird/Google Earth acquired on March 17, 2011. b) Landsat false color composite RGB = TM5-3-2 acquired on November 17, 2010. c) Forest Cover Change 2000–2010 map, green represents the extent of the forest, light yellow represents non-forest, orange represents deforestation between 2000 and 2005, and red represents deforestation between 2005 and 2010. The FCC map is available in kmz format in Supplementary material. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

than 40% relative importance in each study area. Finally, five cloud-free maps of forest cover change (FCC) or gross deforestation over 7.7 million hectares were produced using Landsat images. Examples of the output maps are presented in Fig. 7.

3.2. Accuracy assessment

3.2.1. Per-pixel accuracy

Results of accuracy assessment at the pixel level are presented in the confusion matrix (Table 2) and Fig. 8b. Regarding the error matrix, although overall accuracy was satisfactory (83.3%) as was the accuracy of forest and crop land categories (from 81.5 to 90.3%), significantly lower values were obtained for deforestation classes. We recorded producer accuracies of 40.5 and 41.6% for the first and second deforestation periods respectively, and user accuracies of 57.2 and 64.3% for the same periods. Confusions were identified between forest and cropland categories but to a very limited extent between forest and deforestation classes. The matrix also highlighted cloud and shadow point observations (category 999 with 581 points) which can be considered as “uncategorized” since the interpretation of the three time series did not allow these points to be attributed to other land categories. By removing these points for testing purposes, we observed a limited increase (2% to 3%) in user accuracy in the forest (FFF) and deforestation in the second period (FFP) categories. Also, only a few points ($n = 62$) were observed to be cropland or savannah (PPP) whereas they were mapped as rock or bare soil (OOO). In addition, the pattern of

deforestation on the FCC map can be considered conservative (i.e. deforestation is not overestimated) for two reasons. First, the number of deforestation observations was around 30% to 35% lower for the FCC map than the observations made at sampling points. Second, the errors of omission were significantly higher than the errors of commission, which suggests that deforestation was rarely predicted where forest or cropland was observed. We observed similar trends at each site with lower values for land cover change categories compared with steady land cover categories (Fig. 8, b).

3.2.2. Relative difference measurement

Based on surface areas computed from both the FCC map and the point sampling method, we computed the relative difference on the surface area (forest and cropland/pasture) and the annual deforestation rate (deforestation period 1 and 2) (Fig. 8c). We found very small average relative difference (2%) between the estimated surface area for forest and crop & pasture. All the study areas showed values below 10% except for the Mandrare deciduous dry forest, where a difference of 14.6% was recorded.

On the other hand, relative differences in annual deforestation rates varied (from 8.5 to 49.5%) depending on the study area and the study period. These relative differences combined multiple sources of errors in both datasets, including supervised classification errors on the FCC map and manual labeling of the land cover in the point sampling approach. However, combining all the study areas, we measured a

Table 2

Error matrix at pixel level with the FCC map and the observed sample points.

Category	Observation at sample points						P (%)	U (%)	P ^a (%)	U ^a (%)
	FFF	PPP	HHH	FFP,	FPP	999				
FCC map	9045	1113	3	338	321	279	11,099	86.7	81.5	88.5
FFF	9045	1113	3	338	321	279	11,099	86.7	81.5	84.7
PPP	1189	14,995	6	236	391	240	17,057	90.3	87.9	
HHH	7	154	13	3	1	2	180	59.1	7.2	
FFP,	80	77		439	54	33	683	40.5	64.3	41.0
FPP	108	205		69	546	27	955	41.6	57.2	60.7
OOO	1	62					63			
Total	10,430	16,606	22	1085	1313	581	30,037			

Note: Overall accuracy = 83.3%; Kappa Index of Agreement = 0.70. Note 2: FCC: Forest Cover Change; FFF: Forest 2000–2010; FFP, Deforestation between 2000 and 2005; PPP: Savannah and cropland; HHH: Wetland; OOO: Other land, including Bare soil and rock; 999: Cloud or shadow; P, producer accuracy; U, user accuracy.

^a Mean estimates of stable land cover and land cover change categories.

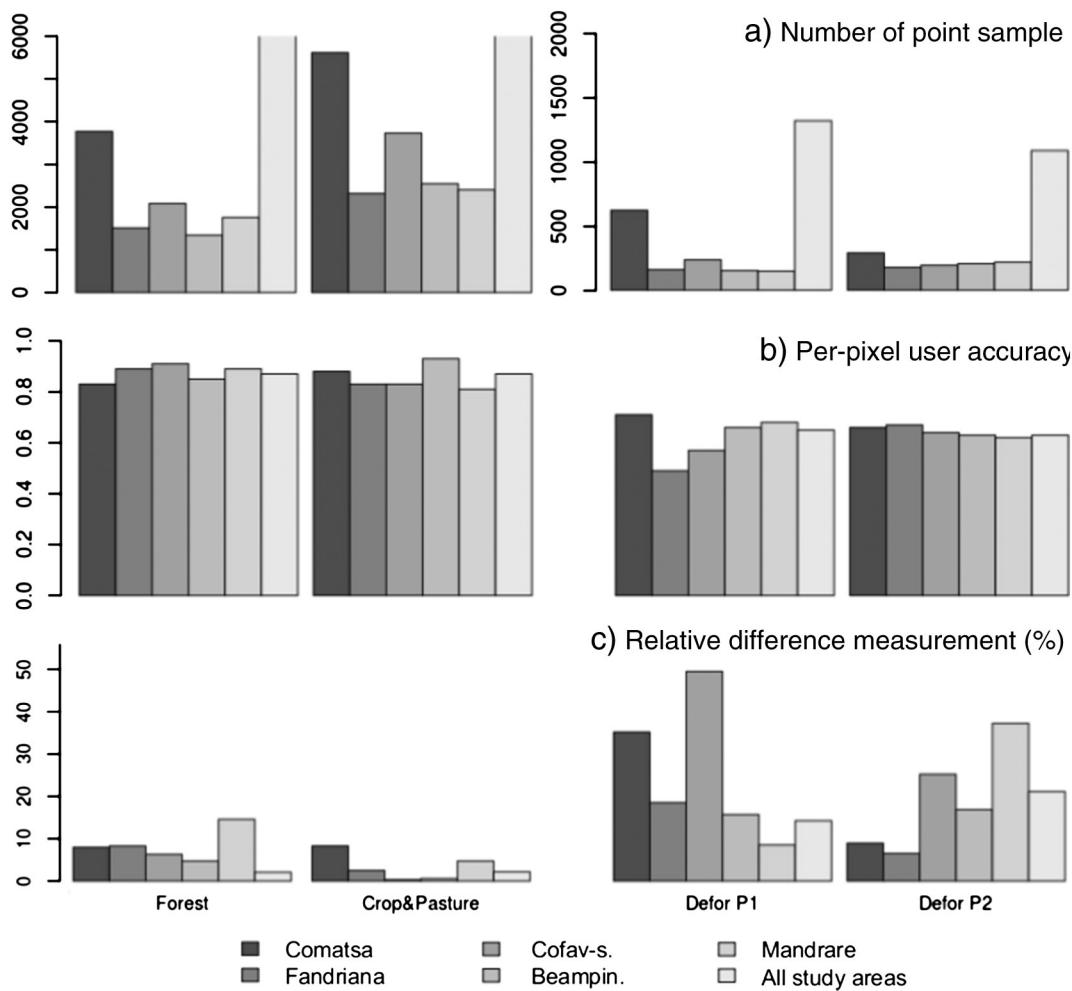


Fig. 8. Accuracy assessment. a) Number of samples observed in each category, b) per-pixel accuracy, and c) relative difference (area and deforestation rate) between point sampling and FCC approach. Defor P1: Deforestation in period 1 (2000–2005), Defor P2: Deforestation in period 2 (2005–2010).

satisfactorily limited difference between the two historical datasets, 14% for the first period and 21.1% for the second.

3.3. Analyses of deforestation estimates and trends

The FCC maps were then used to derive forest area (Table 3) and deforestation estimates (Table 4) statistics for the study areas and the project sites. We measured a total forest cover of 2407 thousand hectares (kha) in the whole study area and 528 kha in the project sites in 2010. This represents 25.5% and 5.6% of the national natural forest cover respectively (MEFT, 2009). According to the latest vegetation map (Moat & Smith, 2007), the present study covers 43.3% of the humid forest and 20.6% of dry forest. We measured 174,725 ha of deforestation in the first period and 108,983 ha in the second period corresponding to a mean deforestation rate of $1.17\%\cdot\text{yr}^{-1}$ and $0.94\%\cdot\text{yr}^{-1}$ in the study areas. In the project sites, the figures were significantly

lower, $0.64\%\cdot\text{yr}^{-1}$ and $0.55\%\cdot\text{yr}^{-1}$ (Table 4). We also observed high variability of the annual deforestation rate depending on the study area. The lowest rates were observed in the tropical humid forest northern site ($0.19\%\cdot\text{yr}^{-1}$) and in the dry spiny forest located in the southern part ($0.28\%\cdot\text{yr}^{-1}$). The highest deforestation rate ($2.48\%\cdot\text{yr}^{-1}$) was observed in the Fandriana site. We identified an overall slight decrease in the annual deforestation rate when the figures for the first period were compared with those for the second period. Only two study

Table 4

Deforestation estimates for each study area and its related project site. *The period is approximate in this table. It is valid for all sites except Comatsa. For the exact date of the processed image, see Table 1.

	Gross deforestation	Period*	ha·y ⁻¹		%·y ⁻¹	
			2000–2005	2005–2010	2000–2005	2005–2010
Study areas	Comatsa		8365	8859	0.77	0.86
	Fandriana		8255	6875	2.50	2.33
	Cofav-South		7238	4739	1.36	0.93
	Beampingaratsy		2890	2800	1.01	1.03
	Mandrare		1791	1873	0.46	0.50
	All		28,540	25,145	1.17	0.94
Project sites	Comatsa		687	402	0.31	0.19
	Fandriana		721	656	2.48	2.53
	Cofav-South		1537	745	1.27	0.64
	Beampingaratsy		336	508	0.59	0.92
	Mandrare		350	545	0.28	0.44
	All		3630	2856	0.64	0.55

Table 3

Forest surface in thousands of hectares for each study area and its related project site.

Forest areas in 2010 (ha × 1000)	Type	Study area	Project site
Comatsa	Humid	1008	216
Fandriana	Humid	275	24
Cofav-south	Humid	489	113
Beampingaratsy	Humid	263	54
Mandrare	Spiny	372	122
All		2407	528

areas showed a significant change in the deforestation rate: a decrease in Cofav-south and an increase in Comatsa. Within the project sites, the changes were more marked: four out of five project sites underwent a significant change.

4. Discussion and conclusion

4.1. The accuracy of the method

Results concerning accuracy showed averaged user accuracy of 84.7% for stable land cover and 60.7% for land cover change at the pixel level. For the whole study area, we recorded relative differences between wall-to-wall and point sampling estimates of 2% and 21.1% for stable land cover and land cover change categories respectively. Concerning per-pixel measurement accuracy, we found that overall accuracy and stable land cover accuracy were consistent with that of other studies which used Landsat imagery to map deforestation at sub-national scale (Duveiller et al., 2008; Huang et al., 2009). These authors obtained accuracy measurement of 92% and 93% or higher on average. However, our land cover change accuracies were lower. Our values cannot be compared to other studies as change at pixel level is rarely validated because of the technical challenge of collecting an accurate reference dataset (Foody, 2010). We observed significant discrepancies between producer (41%) and user accuracy (60.7%). This suggests that the map we produced tends to omit deforestation rather than predicting deforestation beyond its true extent. Relative difference measurements showed a quite different pattern. Forest and non-forested areas measured on the FCC map and the point sampling were very similar. Land cover change categories also showed very good agreement on the measured annual deforestation rate. Overall, these results suggest that stable land cover and land cover change have significantly different accuracy magnitudes. In this respect, when developing a change detection analysis, special efforts should be made to record change class accuracy (Fuller et al., 2003). The stratified systematic sampling validation procedure presented in this study was particularly ambitious (more than 30,000 points observed), since we wanted to obtain independent and statistically valid estimates. Even with this intense visual-interpretation effort, we only obtained a few hundred points that fell within the deforestation categories. It should be noted that the visual-interpretation process is also highly prone to misclassification (interpreter subjectivity errors). Therefore, the validation process applied in this study and discussed above refers more to relative accuracy assessment or consistency assessment. To better account for subjectivity errors, a preliminary trial should be conducted consisting of repeated observation of a sample subset by all the photo interpreters. Although simple, this technique is also critical and time consuming since the sample subset has to be representative of all the observations.

4.2. Comparison with other annual deforestation rate estimates

This paper shows that for all study areas combined, the mean deforestation rate during the 2000–2005 period was $1.17\% \cdot \text{yr}^{-1}$ and $0.94\% \cdot \text{yr}^{-1}$ during the 2005–2010 period. We further investigated the results of this study by comparing them with previous estimates at national and sub-regional scale. Madagascar is recognized to be the country with the highest annual deforestation rate of hot spots, up to $4.7\% \cdot \text{yr}^{-1}$ according to Achard et al. (2002). Previous studies of deforestation in Madagascar included a four-date analysis, from 1950 to 2000 (Harper et al., 2007), and a three-date analysis, from 1900 to 2000 (MEFT, 2009), both at national scale, using also Landsat TM images. The latter recorded a decrease from $0.8\% 1990–2000$ to $0.53\% \cdot \text{yr}^{-1}$ for $2000–2005$. MEFT (2009) was compared with the FCC map presented in this paper using the same study areas and for all the periods of time covered (Table 5). We found limited agreement between the two FCC maps on the overlapping period (2000–2005). We identified four possible explanations for these marked differences. First, the date of the satellite images used in MEFT (2009) was not known exactly, and so a fixed time step of 5 years was used.

Table 5
Comparison of annual deforestation rates for different periods.

Study area	Period	MEFT et al., 2009	FCC, present study
Comatsa	1990–2000	0.33	
	2000–2005	0.11	0.77
	2005–2010		0.86
Fandriana	1990–2000	1.57	
	2000–2005	1.45	2.50
	2005–2010		2.33
Cofav-s	1990–2000	1.21	
	2000–2005	0.19	1.36
	2005–2010		0.93
Beampingaratsy	1990–2000	0.88	
	2000–2005	1.76	1.01
	2005–2010		1.03
Mandare watershed	1990–2000	0.19	
	2000–2005	0.92	0.46
	2005–2010		0.50

From our experience, we are aware that the time interval is rarely exactly five years (Table 1). This can lead to marked variations in annual deforestation rates (Puyravaud, 2003; Vieilledent et al., 2013). Second, in the map by MEFT (2009), there is still a remaining cloud and shadow class which may conceal large local forested/deforested areas. Third, the minimum mapping unit post-processing step may be responsible for a fair amount of difference. Indeed, the FCC map produced in our study used a 0.36 ha MMU whereas the MEFT (2009) FCC map used a 2 ha MMU, which may overlook many small shifting cultivation plots. Lower values in the first three study areas may be partly explained by this. Finally, the definition of forest was not exactly the same in the two studies, especially the definition of dry forest (2 m vs. 5 m minimum height). However, this may not be a real reason in tropical humid forest study areas because forest plots above the defined height are easily recognized. Moreover, the trend measured is highly informative due to the fact that an increase or decrease in the rate of deforestation that occurs several times in a five-year period is unlikely at sub-regional scale. In this respect, the FCC map produced in this study was considered coherent, as the maximum change recorded was a 30% decrease in the rate of deforestation between the two periods studied in this paper.

4.3. On the methodology

The method presented in this study was tested in one of the most difficult tropical regions for forest cover change monitoring since it covers a wide area with small deforestation plots (slash and burn practices) on steep slopes with frequent to permanent cloud cover. The study produced two main results.

First, using stacked multi-date images is a powerful way of reducing classification errors, and enables better characterization of small complex land cover change status. While it is clearly useful to perform a bi-temporal land cover change analysis to produce a transition matrix with no prior information, this method is limited when the change monitored is small compared to the image resolution. In this context, small geometric and thematic errors such as confusion between subtle land cover classes (e.g. two different forest strata, thicket/dry forest, and regeneration/fallow) may greatly compromise true detection of change. As previously observed by Fuller et al. (2003), classification errors may be of the same magnitude as the change itself or even greater. On the other hand, this study confirmed recent observations (Schneider, 2012) that multiplying image datasets acquired at different dates allows successful characterization of land cover and land cover change by incorporating seasonal and vegetation dynamics even in complex and heterogeneous situations. The drawback of the method is that the change class needs to be identified prior to image processing. Consequently the approach developed in this study should be used with clear objectives and known change categories.

Second, with complex datasets and land cover, a machine learning algorithm should be used. Random forest was found to significantly improve accuracy and the related visual output when a supplementary training plot was added. This study confirmed that RF (like any decision tree algorithm) requires additional high quality training plot data (Gislason et al., 2006). The availability of VHR images such as the QuickBird/Google Earth image partially fulfills this need for information without extensive field surveys. Refinement of the output was mainly based on the confusion matrix set using a test dataset. Unlike Rodriguez-Galiano et al. (2012), who suggested using the out-of-the-bag estimator as an unbiased classification accuracy estimator, we found that this indicator was not appropriate for reliable assessment of overall accuracy. Furthermore, optimal sampling intensity was difficult to assess since, contrary to expectations, the trends of accuracy measurement with increasing sampling intensity did not show the same pattern depending on the study area. The reason may be the variability of the results, which, as shown by Schneider (2012) can be high, depending on the experimental design. In this study, the test set was sufficiently large to provide unbiased estimates but not to assess the variability of the error using k-fold cross validation for some categories. Therefore, further calibration tests using k-fold cross validation or repeated bootstrap samples as well as additional training samples would be required to draw valid conclusions on minimum sampling requirements.

4.4. On the forest monitoring perspectives

Madagascar is one of the tropical countries involved in the UNFCCC REDD+ mechanism, which requires setting up a national forest monitoring system that enables accurate and timely monitoring of deforestation, degradation and increase in carbon stocks. DeFries et al. (2007) drew up guidelines to fulfill the need for monitoring tropical deforestation and identified both the sampling strategy and exhaustive mapping (wall-to-wall) as an operational solution. Broich, Stehman, Hansen, Potapov, and Shimabukuro (2009) suggested considering the point sampling approach as a timely and cost effective component of a monitoring system. This paper shows that, contrary to general guidelines on GHG reporting techniques, which suggest using one or the other (GOFC-GOLD, 2010), the two methods are both essential and complementary.

On one hand, exhaustive mapping provides invaluable spatially explicit information which can be used not only for greenhouse gas monitoring but also for forest conservation planning (Harper et al., 2007) and spatial modeling of deforestation (Vieilledent et al., 2013). On the other hand, point sampling has several non-negligible advantages: i) it is less demanding technically, ii) it can address more subtle changes in land use (e.g. degradation, regeneration) and iii) it reduces systematic errors thanks to full visual inspection. The main drawback is the time required for the visual interpretation and for the training session on the image interpretation protocol. In our study, we reached 400–600 point observations per photo interpreter per day, which is a large number of points but with a small sampling rate (see Section 2.4). In this paper, we argue that both approaches should be implemented, point sampling being an integral part of the wall-to-wall work flow. To date, there are no widely accepted validation standards for studies of forest cover change, partly because retrieving a reference dataset is challenging (Foody, 2010). Even if the point sample includes errors, the analysis of consistency errors between the two datasets is repeatable, transparent and produces intelligible results (e.g. in this study, less than 10% errors for stable land cover categories, around 20% for land change cover categories), provided that a consistent sample design is used. More generally, if a point sample database on land cover status (current and past) is maintained, it could be used as an independent validation dataset.

In addition, this study focused on gross deforestation, which was previously identified as the main source of anthropogenic GHG

emission in Madagascar (RPP, 2010). More subtle changes, such as regeneration or degradation, also take place in Madagascar and should be monitored as part of GHG reporting tasks but more importantly with respect to biodiversity and forest resources.

Although the Landsat resolution was acceptable for a five-year deforestation study, it did not allow visual discrimination between a regenerated forest and a fallow and between a degraded forest and an intact forest. Regarding clear cut of forest, in this study we measured a median plot size of 3.1 ha in a five year period, which roughly corresponds to a plot size 0.6 ha on an annual basis (~7 Landsat pixels). The detection of clearings may then be jeopardized if one wants to implement a forest monitoring system with measurements made at intervals of less than five years. These difficulties were summarized by DeFries (2007) “Smaller clearings [compared with clearing for settlement] and more heterogeneous landscapes require data with higher spatial resolution (5–15 m) and greater involvement of an interpreter for visual analysis and more complex computer algorithms that detect less pronounced differences in spectral reflectances”.

Therefore 10-m resolution images would greatly improve deforestation monitoring and would also enable detection of regeneration and degradation. Timely monitoring could be achieved in the future thanks to the recent activation of the SEAS-OI (*Surveillance Environnemental Assistée par Satellite dans l'Océan Indien*), a satellite image reception antenna based in Reunion island whose reception area covers the whole of Madagascar, and the ESA Sentinel-2 platform due to be launched in 2013. Based on open source software that is able to handle large datasets, this method will facilitate further implementation of accurate, cost-effective and reproducible forest monitoring systems.

Acknowledgment

We thank three anonymous reviewers for their helpful comments leading to major improvements in the manuscript. The authors would like also to thank the five photo interpreters, who carried out the visual interpretation for the accuracy assessment, and the WWF local coordinators for their useful feedback on the resulting maps. This study is part of the Holistic Conservation Programme for Forests (PHCF) in Madagascar. The PHCF is a REDD+ pilot-project funded by Air France and jointly implemented by WWF and GoodPlanet foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi: <http://dx.doi.org/10.1016/j.rse.2013.07.008>. These data include Google map of the most important areas described in this article.

References

- Achard, F., Eva, H. D., Stibig, H. -J., Mayaux, P., Gallego, J., Richards, T., et al. (2002). Determination of deforestation rates of the World's humid tropical forests. *Science*, 297, 999–1002.
- Achard, F., Stibig, H. -J., Eva, H. D., Lindquist, E. J., Bouvet, A., Arino, O., et al. (2010). Estimating tropical deforestation from Earth observation data. *Carbon Management*, 1, 271–287.
- Aide, T. M., Clark, M. L., Grau, H. R., L-C. D., Levy, M. A., Redo, D., et al. (2012). Deforestation and reforestation of Latin America and the Caribbean (2001–2010). *Biotropica*, <http://dx.doi.org/10.1111/j.1744-7429.2012.00908.x>.
- Baccini, A., Goetz, S. J., Walker, S., Laporte, N. T., Sun, M., Sulla-Menashe, D., et al. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon density maps. *Nature Climate Change*, 2, 182–185.
- Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., et al. (2008). GlobCover 2005 – Products description and validation report, Version 2.1, 2008. Available on the ESA IONIA website. <http://ionia1.esrin.esa.int/>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees – Wadsworth & Brooks Wadsworth statistics probability series*.
- Broich, M., Stehman, S. V., Hansen, M. C., Potapov, P., & Shimabukuro, Y. E. (2009). A comparison of sampling designs for estimating deforestation from Landsat imagery: A case study of the Brazilian Legal Amazon. *Remote Sensing of Environment*, 113, 2448–2454.

- Clark, M. L., Aide, T. M., Grau, H. R., & Riner, G. (2010). A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America. *Remote Sensing of Environment*, 114, 2816–2832.
- DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B., et al. (2007). Earth Observation for estimating greenhouse gas emissions from deforestation in developing countries. *Environmental Science & Policy*, 10, 385–394.
- DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D., & Townshend, J. (2002). Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *PNAS*, 99(22), 14256–14261.
- Duveiller, G., Defourny, P., Desclée, B., & Mayaux, P. (2008). Deforestation in Central Africa: Estimates at regional, national and landscape levels by advanced processing of systematically-distributed Landsat extracts. *Remote Sensing of Environment*, 112, 1969–1981.
- Foddy, G. M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114, 2271–2285.
- Freidl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, 83, 287–302.
- Fuller, R. M., Smith, G. M., & Devereux, B. J. (2003). The characterisation and measurement of land cover change through remote sensing: Problems in operational applications? *International Journal of Applied Earth Observation and Geoinformation*, 4, 243–253.
- Girard, M. C., & Girard, C. M. (1999). *Traitements des données de télédétection*. Paris: Dunod Ed (529 pp.).
- Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, 27, 294–300.
- Global Forest Resource Assessment (2010). *Main report*. Rome, Italy: Food and Agriculture Organization of the UN.
- GOFC-GOLD (2010). *A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation, Report version COP16-1*. Alberta, Canada: GOFC-GOLD Project Office, Natural Resources Canada (210 pp.).
- Goodman, S. M., & Benstead, J. P. (2005). Updated estimates of biotic diversity and endemism for Madagascar. *Oryx*, 39, 73–77.
- Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island – Digital soil mapping using random forests analysis. *Geoderma*, 146, 102–113.
- Gutman, G., Byrnes, R., Covington, M. S., Justice, C., Franks, S., & Headley, R. (2005). Towards monitoring land-cover and land-use changes at global scale: The global land use survey. *Photogrammetric Engineering and Remote Sensing*, 64, 6–10.
- Ham, J., Chen, Y., Crawford, M. M., & Gosh, J. (2005). Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 492–501.
- Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, <http://dx.doi.org/10.1016/j.rse.2011.08.024>.
- Hansen, M. C., Stehman, S. V., Potapov, P. V., Loveland, T. R., Townshend, J. R. G., DeFries, R. S., et al. (2008). Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings of the National Academy of Sciences of the United States of America*, 105(27), 9439–9444.
- Harper, G., Steininger, M. K., Tucker, C. J., Juhn, D., & Hawkins, F. (2007). Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation*, 34, 1–9.
- Harris, N. L., Brown, S., Stephen, C., Hagen, S. C., Saatchi, S. S., Silvia Petrova, S., et al. (2012). Baseline map of carbon emissions from deforestation in tropical regions. *Science*, 336, 1573–1575 (22).
- Huang, C., Kim, S., Song, K., Townshend, J. R. G., Davis, P., Alstatt, A., et al. (2009). Assessment of Paraguay's forest cover change using Landsat observations. *Global and Planetary Change*, 67, 1–12.
- IPCC (2006). *IPCC guidelines for national greenhouse gas inventories*. Japan: IGES (www.ipcc-nccc.iges.or.jp/public/2006gl/index.html).
- Lawrence, R. L., Bunn, A., Powell, S., & Zambon, M. (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, 90(3), 331–336 (15).
- Mayaux, P., Bartholomé, B., Fritz, S., & Belward, A. (2004). A new land-cover map of Africa for the year 2000. *Journal of Biogeography*, 31, 861–877.
- Mayaux, P., Gond, V., & Bartholomé, E. (2000). A near real-time forest cover map of Madagascar derived from SPOT VEGETATION data. *International Journal of Remote Sensing*, 21(16), 3139–3144.
- MEFT, Ministry of the Environment, Forest and Tourism of Madagascar (2009). *Evolution de la couverture de forêts naturelles à Madagascar, 1990–2000–2005*. Antananarivo, Madagascar: USAID-Cl, pp 132.
- Moat, J., & Smith, P. (2007). *Madagascar vegetation atlas*. Kew, UK: Royal Botanic Gardens.
- Olander, L. P., Gibbs, H. K., Steininger, M., Swenson, J. J., & Murray, B. C. (2008). Reference scenarios for deforestation and forest degradation in support of REDD: A review of data and methods. *Environmental Research Letters*, 3, 025011.
- Pelletier, J., Ramankutty, N., & Potvin, C. (2011). Diagnosing the uncertainty and detectability of emission reductions for REDD+ under current capabilities: An example for Panama. *Environmental Research Letters*, 6, 1–12.
- Pennec, A., Gond, V., & Sabatier, D. (2011). Tropical forest phenology in French Guiana from MODIS time series. *Remote Sensing Letter*, 2(4), 337–345.
- Potapov, P. V., Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Alstatt, A., et al. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM + data. *Remote Sensing of Environment*, 122, 106–116.
- Puyravaud, J.-P. (2003). Standardising the calculation of the annual rate of deforestation. *Forest Ecology and Management*, 177, 593–596.
- Rasi, R., Beuchle, R., Bodart, C., Vollmar, M., Seliger, R., & Achard, F. (2012). Automatic updating of an object-based tropical forest cover classification and change assessment. *IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing* 1939–1404, 1–8, <http://dx.doi.org/10.1109/JSTARS.2012.2217733>.
- Rasi, R., Bodart, C., Stibig, H.-J., Eva, H., Beuchle, R., Achard, F., et al. (2011). An automated approach for segmenting and classifying a large sample of multi-date Landsat imagery for pan-tropical forest monitoring. *Remote Sensing of Environment*, 115(12), 3659–3669.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Ochoa, M., & Rigol-Sánchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93–104.
- RPP-Madagascar (2010). *Redd Readiness Preparatory Proposal - Madagascar*. Antananarivo, Madagascar: Forest Carbon Partnership Facility, pp 107.
- Schneider, A. (2012). Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat data and data mining approach. *Remote Sensing of Environment*, 124, 689–704.
- Song, C., Woodcock, C. E., Seto, K. C., Pax-Lenney, M., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, 75, 230–244.
- Stach, N., Salvador, A., Petit, M., Faure, J. F., Durieux, L., Corbane, C., et al. (2009). Land use monitoring by remote sensing in tropical forest areas in support of the Kyoto Protocol: The case of French Guiana. *International Journal of Remote Sensing*, 30(19), 5133–5149.
- Stehman, S. V. (2005). Comparing estimators of gross change derived from complete coverage mapping versus statistical sampling of remotely sensed data. *Remote Sensing of Environment*, 96, 466–474.
- Stehman, S. V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20), 5243–5272.
- Steininger, M. K., Godoy, F., & Harper, G. (2009). Effects of systematic sampling on satellite estimates of deforestation rates. *Environmental Research Letters*, 4, 034015.
- Tucker, C. J., Grant, D. M., & Dykstra, J. D. (2004). NASA's global orthorectified Landsat data set. *Photogrammetric Engineering and Remote Sensing*, 70, 313–322.
- Vieilledent, G., Grinand, C., & Vaudry, R. (2013). Forecasting deforestation and carbon emissions in tropical developing countries facing demographic expansion: A case study in Madagascar. *Ecology and Evolution*, <http://dx.doi.org/10.1002/ee.3550> (Article first published online: 3 MAY 2013).