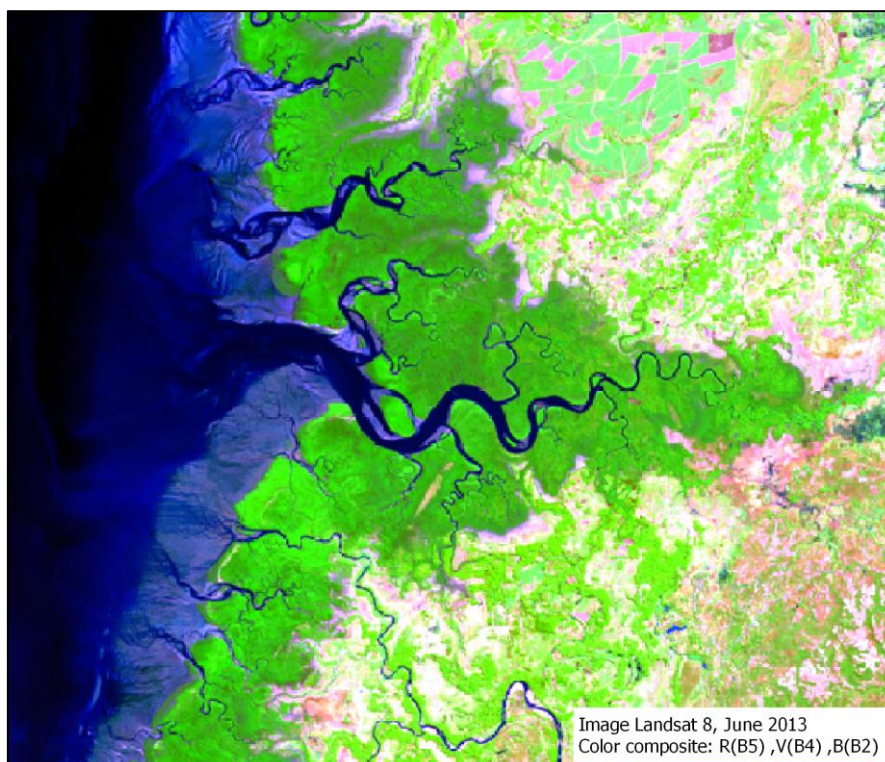


Mangrove deforestation analysis in Northwestern Madagascar

Stage 1 - Analysis of historical deforestation

Frédérique Montfort, Clovis Grinand, Marie Nourtier



March 2018

1. Context and study area :

The study area is the adjacent bays of Ampasindava, Tsimipaika and Ambaro in northwestern Madagascar (*Figure 1*). This area of 24 402 ha, contain the second country's largest mangrove ecosystem on the island, which is also currently experiencing the most rapid rates of mangrove deforestation nationwide. Almost 500 hectares of mangroves are being lost from this ecosystem every year, primarily due to charcoal production. From both a biodiversity and sociological perspective, there is an urgent need for long-term mangrove conservation and restoration action in this region.

Mangrove forests are essential ecosystems, not only for exceptional biodiversity that they shelter, but also for the fundamental contribution in services and goods thta they offer to the coastal populations. There is now international recognition of the exceptional capacity of mangrove forests to sequester carbon. This carbon has a value on the international carbon market. If this value can be realised and transferred to the people whose livelihoods depend on the exploitation of mangroves, this benefit has the potential to both incentivise and fund sustainable, locally-led mangrove management. Thus, preventing the continued wholesale loss of these invaluable ecosystems and ensuring the long-term sustainability of coastal livelihoods.

Blue Ventures is working to realise this potential by partnering with local stakeholders to develop a mangrove carbon project in Tsimipaika Bay, validated under the VCS standard. A key part of project development is establishing a 'without-project' deforestation scenario. Etc Terra – Rongead is in charge of realized this part. This report outlining the method and data used for the stage 1, consisting in the analysis of historical deforestation from 2000-2014 in the area of interest.

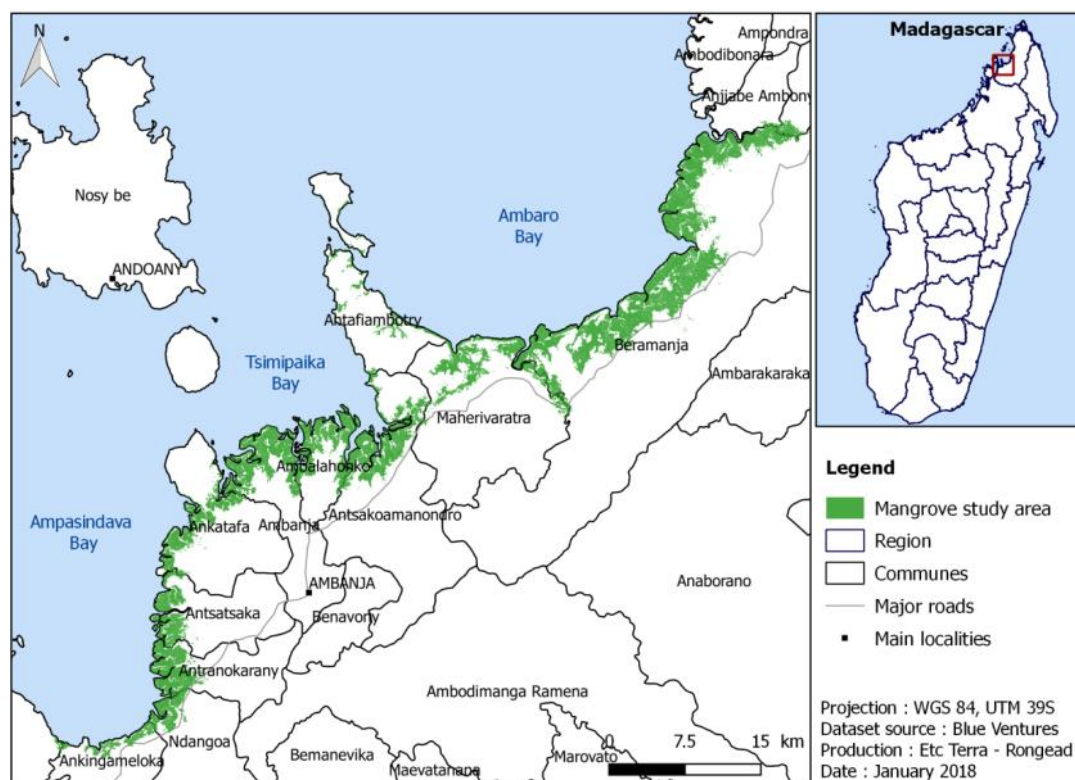


Figure 1 : Area of interest in northwestern Madagascar

2. Methodology :

This step aims to analyze deforestation during the historical reference period (2000-2014) within the RRD.

2.1. Collection of appropriate data source

The historical analysis must respect the following criteria :

- Be produced for 2 or more points in time in a period no more than 20 years prior to project start ;
- Use remotely sensed spatial data that have medium resolution (30X30 m or less) ;
- Produce a map with 90% accuracy in the classification of forest versus non-forest (the accuracy is assessed via high resolution data or ground truthing points on the last date analysed).

The analysis follows the method presented by Grinand et al. (Grinand et al., 2013) based on a multi-dates analysis for a direct classification of land uses and changes using the algorithm *RandomForest*. Main steps are detailed in the following section.

Accuracy assessment was specifically done on the last Mangrove/Non-Mangrove map of the reference period, in 2014. A sample of validation points, were classified on Landsat images and very high resolution images available in Google Earth. The overall accuracy is 98%. For mangrove and non-mangrove categories, accuracy is respectively 93% and 99% and are in accordance with the methodology requirements. Results are presented in Table 7.

2.2. Satellite image database

LANDSAT image from 2000 to 2014 were used with a spatial resolution of 30 m. Those images are available on the USGS data servers (Earth Explorer, www.earthexplorer.usgs.gov) for free. These images come from three different LANDSAT missions: 5,7 and 8/OLI, which have slightly different sensors in terms of width and number of spectral bands. Images were uploaded by bands; therefore it was primarily necessary to combine these single bands into multispectral images (stacking) to be comparable from one date to another (*Table 1*).

Table 1 : Spectral band used

LANDSAT 5/7		LANDSAT 8	
Spectral bands	Wavelength (nm)	Spectral bands	Wavelength (nm)
Bande 1 - Blue	0,45 - 0,52	Bande 2 - Blue	0,45 - 0,52
Bande 2 - Green	0,52 - 0,60	Bande 3 - Green	0,53 - 0,60
Bande 3 - Red	0,63 - 0,69	Bande 4 - Red	0,63 - 0,68
Bande 4 - Near- Infrared (NIR)	0,76 - 0,90	Bande 5 - Near- Infrared (NIR)	0,85 - 0,88
Bande 5 - Short-wave Infrared (SWIR 1)	1,55 - 1,75	Bande 6 - Short-wave Infrared (SWIR 1)	1,56 - 1,66
Bande 7 - Short-wave Infrared (SWIR 2)	2,08 - 2,35	Bande 7 - Short-wave Infrared (SWIR 2)	2,10 - 2,30

The study area is covered by one LANDSAT scene, presenting the following identifiers : 159/69 (path/row). The selected LANDSAT scenes and time interval between dates are presented in the following tables.

Table 2 : Date of selected LANDSAT image

Satellite	Sensor	Processing Correction Level	Date of acquisition	Scene cloud cover (%)
Landsat 7	ETM+	L1TP	13/06/2000	1
Landsat 7	ETM+	L1TP	02/03/2003	4
Landsat 7 - GLS2005	ETM+	L1T	17/08/2006	-
Landsat 5	TM	L1TP	30/08/2008	10
Landsat 7 - GLS2010	ETM+	L1T	09/06/2010	-
Landsat 8	OLI/TIRS	L1TP	09/06/2013	2.2
Landsat 8	OLI/TIRS	L1TP	12/06/2014	3.49

To ensure good geometrical quality images, LANDSAT Global Land Survey products (GLS) and Level-1T (L1T) were used. According to Gutman et al. (2008), these data have sufficient radiometric and geometric qualities to perform land use change analysis. Additionally, we performed a visual inspection of each scene to check their geometric consistencies. No additional geo-rectification was performed.

2.3. Data pre-processing

In order to improve the classification, several spectral indexes were derived from the primary bands as presented in the following table.

Table 3 : Spectral indexes calculated

Index	Formula
NDVI (<i>Normalized Difference Vegetation Index</i>) –Vegetation spectral enhancement	$NDVI = \frac{PIR1 - R}{PIR1 + R}$
NIRI (<i>Near Infrared Reflectance Index</i>) – Soil spectral enhancement	$NIRI = \frac{PIR2 - PIR1}{PIR2 + PIR1}$
NDWI (<i>Normalized Difference Water Index</i>) – Water spectral enhancement	$NDWI = \frac{PIR1 - V}{PIR1 + V}$

2.4. Supervised classification

After data pre-processing, the method to establish a deforestation map follows three main steps:

- Definition of land use and land cover changes classes;

- Delimitation of training plots;
- Classification with a specific algorithm.

2.4.1. Definition of land use and land cover changes classes:

Land use and land cover change (LULCC) classes that exist in the areas and are detectable with Landsat imagery are presented in the following table.

Table 4 : Typologie of land use and land cover changes classes for the study

Identification code	Description of the class
2	Deforestation between 2000-2003
3	Deforestation between 2003-2006
4	Deforestation between 2006-2008
5	Deforestation between 2008-2010
6	Deforestation between 2010-2013
7	Deforestation between 2013-2014
8	Mangrove over the over the 2000 - 2014 period
9	Water
10	Bare soil, sand, rocks, settlements
12	Mosaic of cropland, fallows and savannah land
14	Permanent forest - Other vegetation
15	Bare soil temporarily flooded
16	Paddy field, Flooded area

2.4.2. Delimitation of training plots:

Delimitation of trainings plots is a necessary step to calibrate the classification algorithm when applying a supervised classification. The accuracy of the classification mainly depends on the quality of the delimitation of these training plots. Therefore, a standardized and rigorous photo-interpretation work was conducted. Photo-interpretation was carried on the basis of field knowledge, LANDSAT images patterns and high-resolution images from *Google Earth*. Number of polygons and area delimited are presented in the table below.

Table 5 : Number of polygons and associated delimited area used as training plots

LULCC Class ID	Number of training polygons	Cumulated area (ha)
2	20	27
3	12	18
4	17	28
5	22	31
6	36	68
7	19	26

8	58	516
9	12	33
10	29	194
12	56	477
14	66	699
15	46	67
16	10	90
Total	403	2276

First, in order to improve the localization and determination of changes, those area where highlighted by performing a multi-dates color composite. Then, training plots were located in cluster *i.e.* by grouping several plots of different categories on a same landscape unit or small area (Figure 2). A landscape unit is defined according to the scale of study and represents here roughly an area of analysis below 3 km² and/or at 1:10 000 scale. In order to reduce noise in training data and to guarantee the good consideration of the forest definition, plots contours were verified by superposition on very high-resolution images available on *Google Earth*.

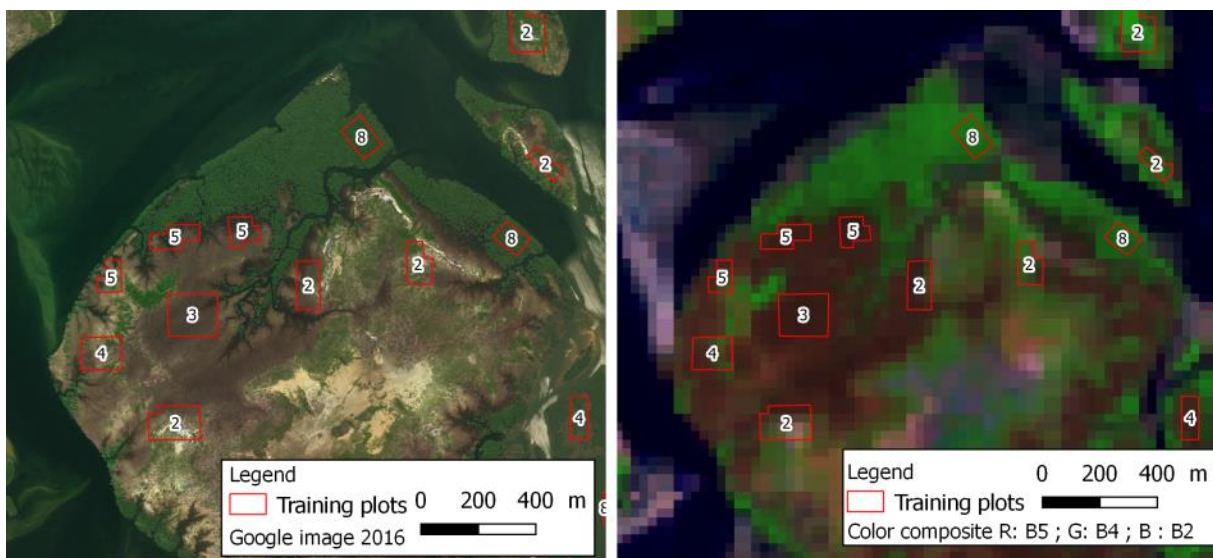


Figure 2 : Example of landsat image (2014) on the right and comparison to Google Earth on the left.

2.4.3. Classification:

Afterward, the training plot spatial database was correlated with the multi-date stacked image database using a statistical algorithm. The RandomForest algorithm, developed by Breiman (2002) and available in R software was used. It is a data-mining algorithm that combines bagging techniques and decision tree. It was successfully applied in land cover change studies in humid forests of Madagascar (Grinand et al., 2013).

RandomForest calibration was performed using 2/3 of randomly selected training plots. The remaining plots (1/3) were used to perform an “internal validation” by the algorithm. Based on a confusion matrix, this validation enabled the operator to identify the remaining confusions in order

to add, remove or change the training plots on the GIS and redo the classification until satisfactory results were obtained.

2.4.4. Post-classification treatments:

After classification, some isolated pixels of forest were found, giving a noisy appearance to the map. To respect the requirements on Minimum Mapping Unit (MMU - linked to the forest definition), those pixels were removed during post-classification processing. In the present study, MMU is 1 ha for forest. A majority filter with a 3x3 window was first used to remove isolated pixels. The classified image was filtered with a Grass/R script for forests patches.

2.4.5. External validation of results

This step entails a statistical analysis of the classification results accuracy, with a points sampling approach. This validation was carried out on Mangrove / Non-mangrove map of the last date of the reference period, 2014, as provided by the BL-UP module. Validation points were selected independently of training plots that were used for the classification.

Following the good practices recommendations proposed by Olofsson et al (2014), a stratified random sampling was used with 3 map classes (Mangrove, Non-mangrove and Water). Stratification is of interest for reporting accuracy results per classes and improves the precision of the accuracy and area estimates (Olofsson et al., 2014). The overall sample size was calculated with the formula proposed by Cochran in 1977, and can then be distributed among the different classes. N is number of pixels in the area of interest, $S(\hat{O})$ is the standard error of the estimated overall accuracy that we would like to achieve, W_i is the mapped proportion of area of class i , and S_i is the standard deviation of the stratum i :

$$n = \frac{(\sum W_i S_i)^2}{[S(\hat{O})]^2 + (1/N) \sum W_i S_i^2} \approx \left(\frac{\sum W_i S_i}{S(\hat{O})} \right)^2$$

With a standard error for overall accuracy of 0.01 and a user's accuracies of 0.95, 0.90 and 0.95 for Non-Mangrove, Mangrove and Water, respectively, the resulting sample size $n = 498$ (Table 6).

Table 6 : Sampling size results

	Non-Mangrove	Mangrove	Water	Total
Area in pixels	1939068	263240	1883970	4086278
W_i (Mapped proportion)	0,47	0,06	0,46	
U_i (Expected user's accuracy)	0,95	0,90	0,95	
S_i (Standard deviation)	0,22	0,30	0,22	
$W_i * S_i$	0,10	0,02	0,10	0,22
SE overall accuracy				0,01
Total number of samples				498

The overall sample size resulting from this calculation can be allocated among the stratum in different way: proportional, equal, optimal and power allocation as described in Olofsson et al, 2014.

An equally allocation between the classes was chosen: 166 validation points per classes were spread out on the study area (Figure 3).

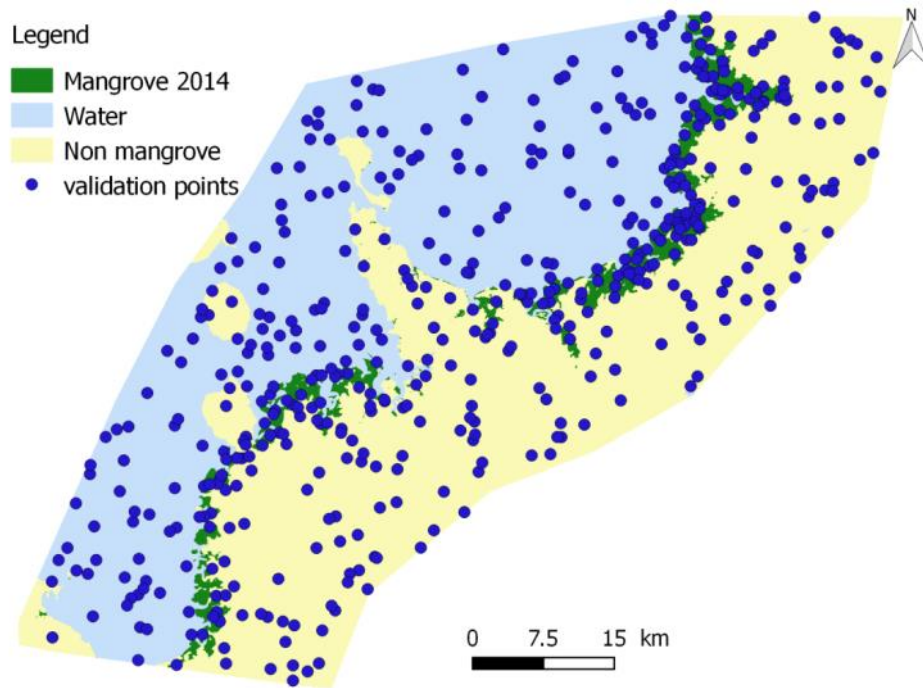


Figure 3 : Distribution of the validation points randomly selected on the reference map of 2014 (Mangrove/Non-mangrove map)

The state of the forest was visually inspected on every point and gathered in a spatial database. The inspections were based on very high-resolution Google Earth images and on the LANDSAT images that had been used for the classification. The result of the photo-interpretation (reference dataset) was finally compared with the map to produce a confusion matrix. This confusion matrix is used to calculate the accuracy of the map which is presented in **Erreur ! Source du renvoi introuvable.7**.

According to Olofsson et al. 2014, in the case of a stratified sample and with a number of sample units per stratum disproportionate relative to the area of the stratum, is necessary to estimate the surface proportions (\hat{p}_{ij}) for each cell in the error matrix rather than the number of points before proceeding with the analysis. The area proportions are estimates as $\hat{p}_{ij} = W_i \times n_{ij} \div n_i$ where W_i is the mapped proportion of area of class i , n_{ij} is the sample count in cell ij and n_i is the total number of sample counts in map category i . These proportions are then calculated in hectares, multiplying the totals of the columns by the size of the stratum and the size of the pixels in hectares ($30^2/100^2$). The standard errors of the area estimates were estimates, by the following equation for a stratified random sample:

$$S(\hat{p}_{\cdot j}) = \sqrt{\sum_i \frac{W_i \hat{p}_{ij} - \hat{p}_{ij}^2}{n_i - 1}}$$

Finally the accuracy of the map was calculated. Three different accuracy measures are of interest: i) overall accuracy which is simply the sum of the diagonals in the error matrix of estimated area

proportions; ii) user's accuracy which for a map category i is given by $U_i = \hat{p}_{ii} \div \hat{p}_{i\cdot}$ and iii) producer's accuracy for map category j given by $P_j = \hat{p}_{jj} \div \hat{p}_{\cdot j}$ where $\hat{p}_{i\cdot}$ and $\hat{p}_{\cdot j}$ are the row and columns totals respectively.

The overall accuracy is 98%. For Mangrove and non-mangrove categories, accuracy is respectively 93% and 99% and are in accordance with the methodology requirements. All the results are presented in the *Table 7*.

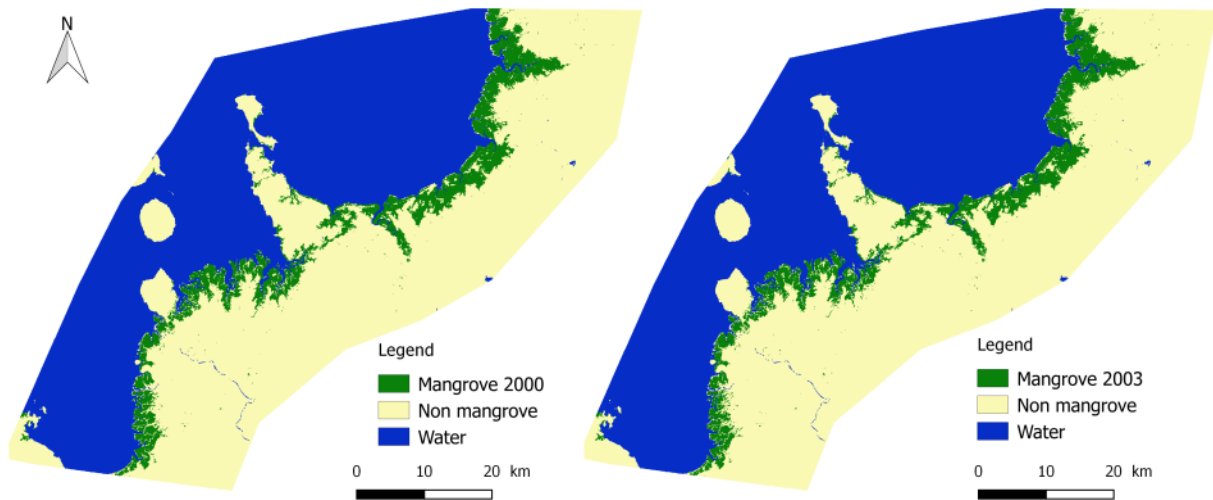
Table 7 : Confusion matrices (number of points above and estimates area proportions below) on the external validation of the reference Mangrove Non-mangrove map of 2014

		Observed (Landsat/Google earth)			Total	Pixels	Wi
		Non-mangrove	Mangrove	Water			
Predicted (Forest / Non-forest map)	Non-mangrove	164	2	0	166	1 939 068	0,47
	Mangrove	11	155	0	166	263 240	0,06
	Water	0	0	166	166	1 883 970	0,46
Total		175	157	166	498	4 086 278	1

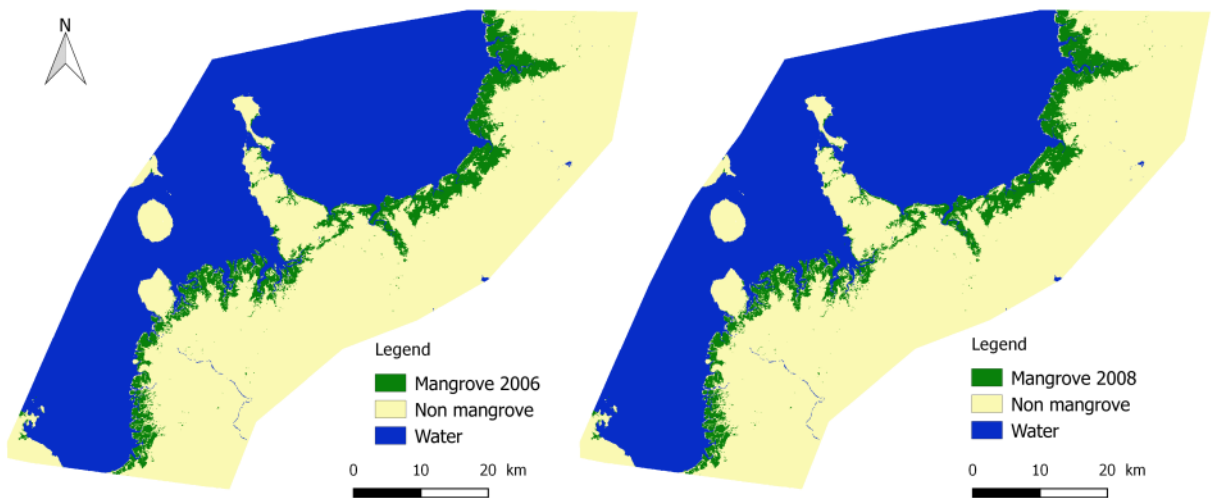
		Observed (Landsat/Google earth)			Total	Pixels	Wi
		Non-mangrove	Mangrove	Water			
Predicted (Forest / Non-forest map)	Non-mangrove	0,4643	0,0057	0,0000	0,4700	1 939 068	0,47
	Mangrove	0,0040	0,0560	0,0000	0,0600	263 240	0,06
	Water	0,0000	0,0000	0,4600	0,4600	1 883 970	0,46
	Total	0,4683	0,0617	0,4600	1	4 086 278	1
Area [pix]		1 913 658	252 069	1 879 688	4 086 278		
Area [ha]		172 229	22 686	169 172			
S(Area)		0,0042	0,0042	0,0000			
S(Area) [ha]		1 529	1 529	0			
95% CI [ha]		2 997	2 997	0			
User's		0,99	0,93	1,00			
Producer's		0,99	0,91	1,00			
Overall		0,9804					

2.5. Mapping of historical deforestation

Mangrove cover maps for the 7 dates of analysis on the RRD are presented in *Figure 4* and



deforestation maps are presented in *Figure 5*.



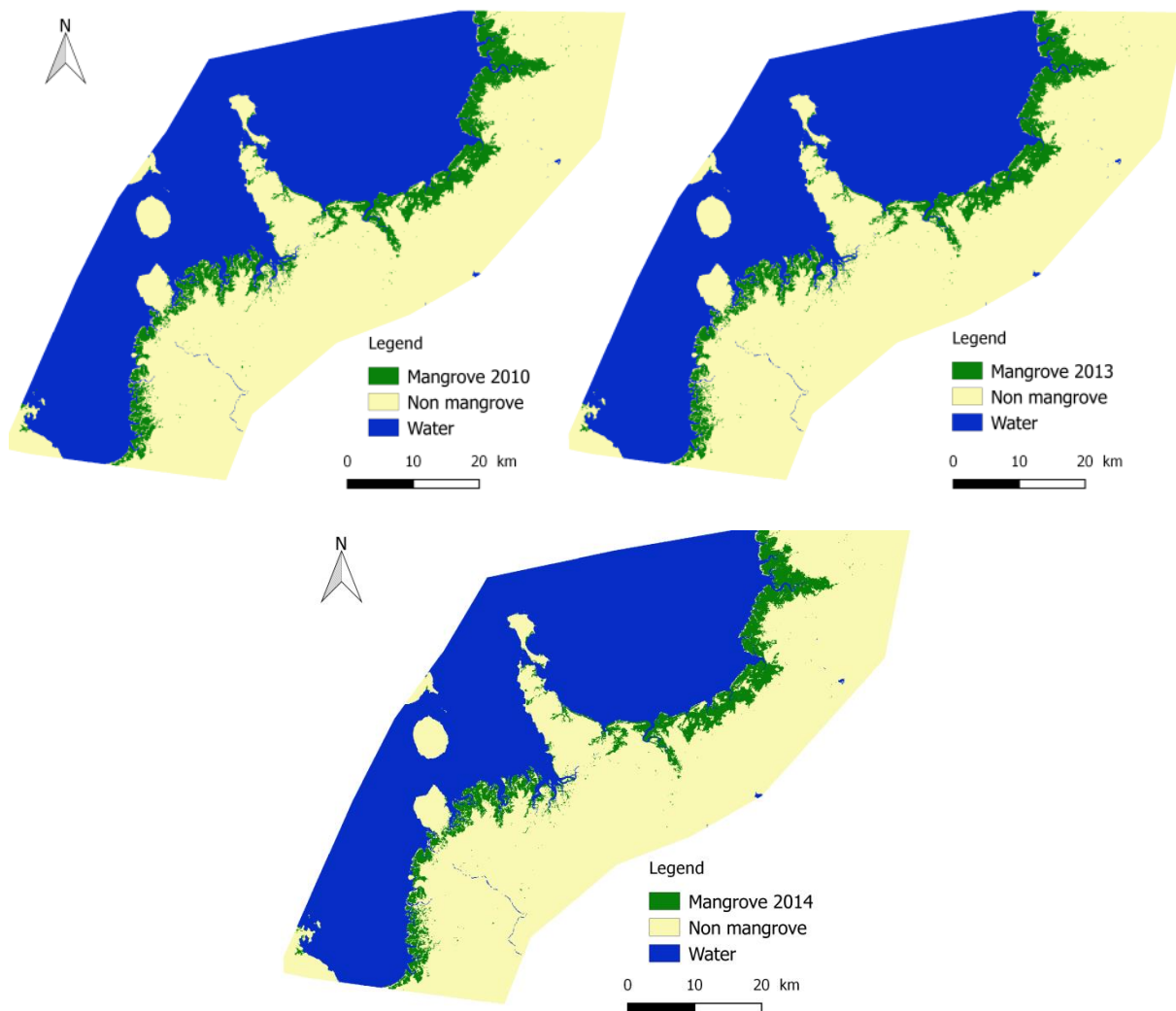


Figure 4 : Mangrove cover maps for the 7 dates of historical analysis (2000, 2003, 2006, 2008, 2010, 2013, 2014) on the Project zone

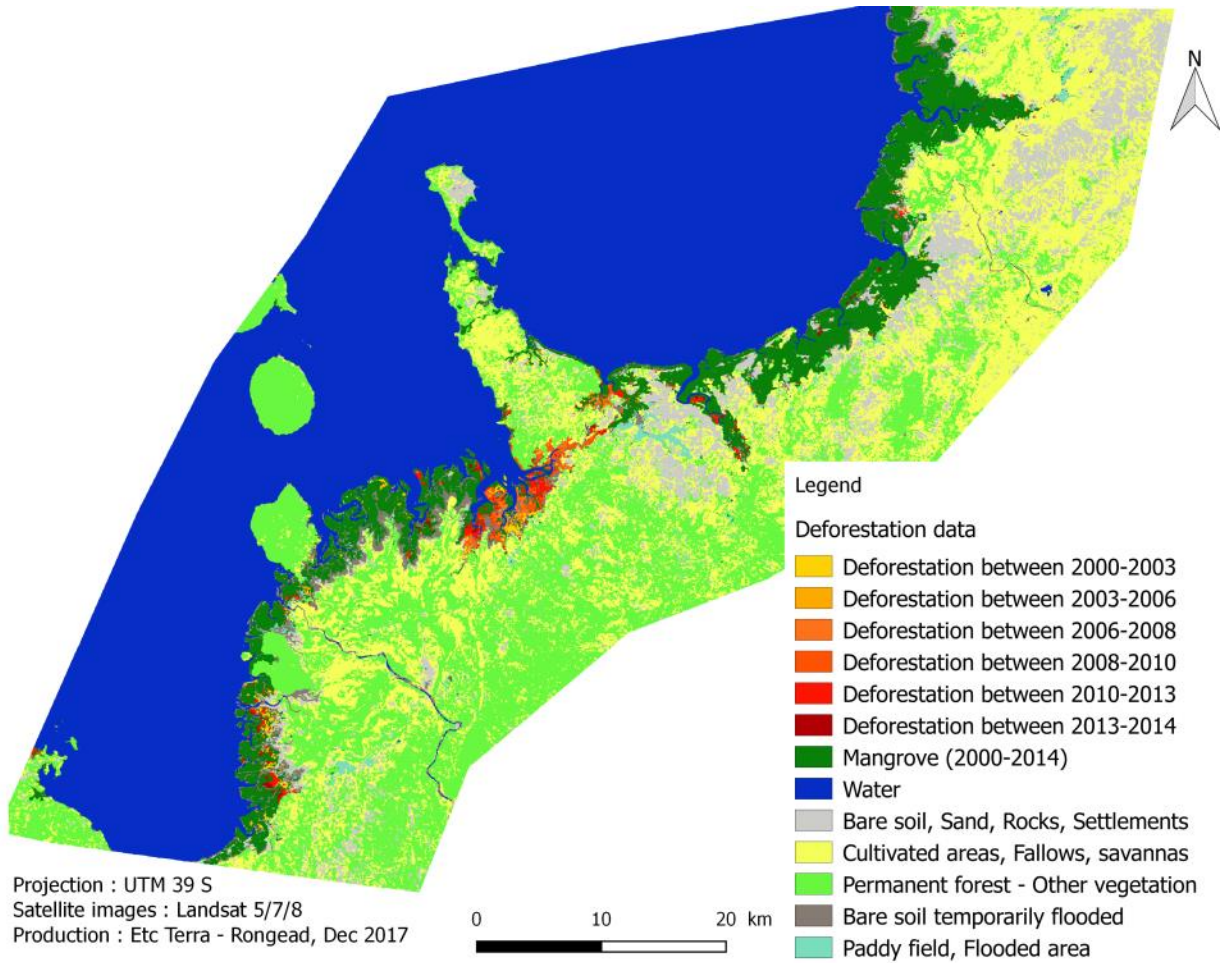


Figure 5 : Deforestation map between 2000 – 2014 on the Project zone

2.6. Calculation of the historical deforestation rate

In a first approach, an annual deforestation rate is a ratio between the deforestation area over a period and the number of years between the two dates of the same period (Menon and Bawa 1997). However, several publications explained that this simple ratio could not be used as deforestation rate dynamics followed a compound interests rule because the ratio changed with forest area during the period of interest as deforestation continued (Puyravaud 2003). Hence, an adaptation of this law was done to calculate annual deforestation rate. The following standardized equation proposed by Puyravaud (2003) was used in the present study:

$$\theta = -\frac{1}{t_2 - t_1} \ln \frac{A_2}{A_1}$$

Where A_i is the forest area during the year t_i .

This calculation approach requires knowing exactly the interval between the two dates (t1 and t2) of the considered period. Therefore, a table summarizing the exact interval between images was established (Table 8).

Table 8 : Time interval between reference year

Period		Time interval (decimal year)	
		Day	Year
13/06/2000	02/03/2003	992	2,71780822
02/03/2003	17/08/2006	1264	3,4630137
17/08/2006	30/08/2008	744	2,03835616
30/08/2008	09/06/2010	648	1,77534247
09/06/2010	09/06/2013	1096	3,00273973
09/06/2013	12/06/2014	368	1,00821918

Gross deforestation expressed in hectares and percentage of the RRD extracted from the deforestation map is presented in the following table.

Table 9 : Results of historic deforestation on RRD during the reference period (statistic with filter)

	Mangrove area (ha)						
	2000	2003	2006	2008	2010	2013	2014
RRD	27403	26759	26510	26058	25155	24006	23692

	Annual mangrove loss (ha/an)					
	2000 - 2003	2003 - 2006	2006 - 2008	2008 - 2010	2010 - 2013	2013 - 2014
RRD	236,87	72,07	221,52	508,87	382,54	311,45

	Annual deforestation rate (%)					
	2000 - 2003	2003 - 2006	2006 - 2008	2008 - 2010	2010 - 2013	2013 - 2014
RRD	0,87471352	0,27058125	0,84280576	1,9874845	1,5565681	1,3059423

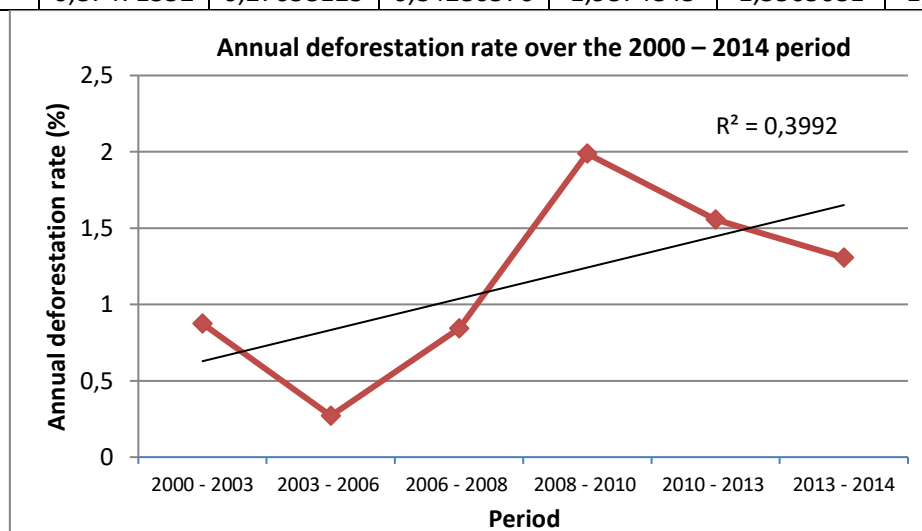


Figure 6 : Annual deforestation rate over the 2000 – 2014 period

Table 10 : Results of historic deforestation on RRD during the reference period (statistic without filter)

	Mangrove area (ha)						
	2000	2003	2006	2008	2010	2013	2014
RRD	26350	25605	25306	24836	23888	22607	22091

	Annual mangrove loss (ha/an)					
	2000 - 2003	2003 - 2006	2006 - 2008	2008 - 2010	2010 - 2013	2013 - 2014
RRD	274,12	86,34	230,58	533,98	426,61	511,79

	Annual deforestation rate (%)					
	2000 - 2003	2003 - 2006	2006 - 2008	2008 - 2010	2010 - 2013	2013 - 2014
RRD	1,05528528	0,33918789	0,91972682	2,19213889	1,8355453	2,2901079

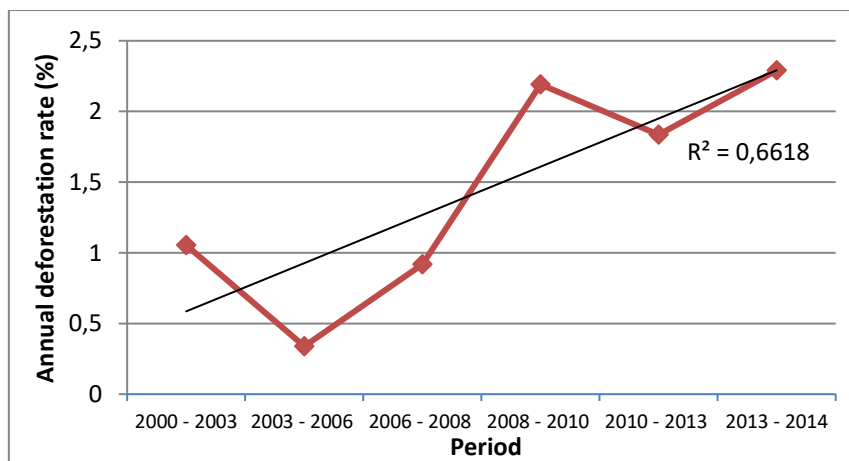


Figure 7 : Annual deforestation rate over the 2000 – 2014 period

The period between the 2013 and 2014 images is too short to observe any visible changes in deforested area. Deforestation is often only a few isolated pixels, those pixels were removed during post-classification processing, which explains the differences between the results with and without filter.