Centre for Geo-Information

Thesis Report GIRS-2015-23

Exploring Relations Between Human Migration and Land Cover Changes in the Tropics

A master thesis investigating relationships between human migration and deforestation, agricultural expansion and urbanization in the decade 2000-2010 on a large scale

Tei Overgoor

10-07-2015

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Exploring Relations Between human Migration and Land Cover Changes in the Tropics

A master thesis investigating relationships between human migration and deforestation, agricultural expansion and urbanization in the decade 2000-2010 on a large scale

Tei Overgoor

Registration number 881216-637080

Supervisors:

Kathleen Neumann

Sytze de Bruin

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre,

The Netherlands**.**

29-06-2015

Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2015 -23
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Abstract

Rapid land cover changes in the tropics are a major threat to the world's valuable natural assets such as biodiversity, carbon stock and ecosystems. Demographic dynamics, of which migration is an important component, are often considered to be related to the alterations in land cover. Although the relations between migration and land cover changes have been researched in multiple studies, it is has not yet been conducted using a global scale and a spatial resolution higher than country level simultaneously, due to a lack of spatial explicit data with sufficient extent and resolution. However, in recent years datasets with improved quality have been becoming more available. In this study it was searched for the most suitable global spatial explicit datasets that could be used for testing relations between land cover change and migration. This resulted in the MODIS land cover and the Centre for International Earth Science Information Network (CIESIN) migration dataset. With the use of linear correlations and regression trees relations between the datasets were assessed. This analysis was repeated using a national extent and a higher resolution to compare the results for different spatial scales. It was done for the countries Colombia and Indonesia because spatially distinct patterns with a variety of migration and land cover processes occurred here. The results showed that no significant relation could be found between the land cover change and migration variables both on a global and a national scale. Rather, the forest cover change depended on the size of existing forest stock in the area. Improvements of the study could be higher quality global data with less unknown errors and higher resolution. Furthermore, interactions with other factors, such as biophysical characteristics and local policies, should probably be taken into account.

Keywords: Land cover change, Migration, Geographic information systems, GIS, Tropics

Acknowledgements

I would like to thank my supervisors Kathleen Neumann and Sytze de Bruin (Laboratory of Geo-Information Science and Remote Sensing) for all their help and support. In regular (skype) meetings they provided me with useful feedback and advice. Also, they were always available for questions and meetings.

Table of Contents

1 Introduction

1.1 Context

Land cover change with both human and natural causes has been an always ongoing process (Lambin, Geist, & Lepers, 2003) and its recent rates and scales are alarming (Hansen et al., 2013; Rockström et al., 2009). Especially agricultural and urban lands have been expanding rivalling natural area's (Foley et al., 2005). These changes have an impact on both local and global environments (Foley et al., 2005), including biodiversity (Hahs et al., 2009; Reidsma, Tekelenburg, van den Berg, & Alkemade, 2006), ecosystem services (Lambin & Geist, 2006; Millennium Ecosystem Assessment, 2005), carbon stocks (van der Werf et al., 2009), climate (Lawrence & Chase, 2010) and soils (Hartemink, Veldkamp, & Bai, 2008). Especially the tropics are threatened by the increasing rates of deforestation and urban expansion (Hansen et al., 2013; Seto, Fragkias, Güneralp, & Reilly, 2011). This is particularly worrisome since the tropics hold the most valuable assets of the world (Costanza et al., 1997; Ricketts, Daily, Ehrlich, & Michener, 2004), such as high biodiversity and large carbon stock (Pan et al., 2011).

Among multiple factors such as climate change and land use intensification, demographic dynamics is often considered to be contributing to land cover change (Aide & Grau, 2004; D. L. Carr, 2004; Rudel et al., 2005). The growing world population with an increasing demand for natural resources puts more and more pressure on natural lands, which are converted to agricultural land (Ramankutty, Foley, & Olejniczak, 2002). Furthermore, since population growth typically is highest in cities – largely driven by immigration – expansion of urban areas is ongoing (Lambin et al., 2003). Human migration from rural areas to urban locations however, does not reduce the pressure on land for food production as urban inhabitants have a relatively high consumption (Grimm et al., 2008). At the same time also rural-rural migration is taking place where people abandon areas, for example because of land degradation for more favourable areas that are often at the front of the forest (Carr, 2009).

1.2 Problem Definition

Multiple studies have investigated statistical relationships between land cover change and its potential drivers on a local scale. Huijun et al. (2002) report a decrease and fragmentation of forest areas in the tropical mountains of Xishuanbanna in Southwest China, which were caused by both demographic growth and local policies. In contrast, deforestation has been found to be uncorrelated to demographic pressure in recent years on Bellona Island, Solomon Islands, as the increasing food demand was covered by import (Birch-Thomsen, Reenberg, Mertz, & Fog, 2010). An increasing population is more often a poor predictor of deforestation on a local scale because intensive agriculture does not inhere many people (Gasparri & Grau, 2009).

Aide & Grau (2004) describe the trend of rural-urban migration in Latin America, resulting in on the one hand expanding cities and on the other hand abandoned land with the potential for forest recovery. Seto et al. (2011) concluded that population increase was the most important driving force for the largest urban expansions in the world between 1970 and 2000, in China, Africa and India. Urban populations have larger ecological footprints than just the area they live in (Grimm et al., 2008). Their relatively high food consumption puts pressure on the surrounding lands of the cities.

This relation can be found on a regional level rather than a local level. On the other hand, the large scale soy production in Brazil for export to China (Nepstad, Stickler, & Almeida, 2006) is not related to population change on a regional level. DeFries et al. (2010) concluded for most countries in the tropics that a relatively strong relation can be found between agricultural export and forest change.

As illustrated by this overview, trends found on the local scale in one region are not necessarily applicable to another region. Additionally, multiple scale studies about relations between land cover change and demographic dynamics and environmental change (Aide et al., 2013; Kok & Veldkamp, 2001; Lele, Nagendra, & Southworth, 2010; Redo, Aide, & Clark, 2012; Sánchez-Cuervo & Aide, 2013), show that for the same study areas often different relations can be found depending on the spatial scales. For example, Aide et al. (2013) state that net deforestation in Latin America and the Caribbean often does not coincide with trends on small scales where much forest recovery is found. According to them, causes on multiple scales play a role in land cover change, as for example local migration can stimulate forest recovery when land is abandoned while global demand for food drives agriculture expansion. Furthermore, Lele et al. (2010) emphasized that in order to understand drivers of land cover change, studying large area's with variations in local context is important. In their study in the Cauvery basis in India, they found that protected areas and accessibility played a major role at different spatial scales for deforestation between 2001 and 2006.

In conclusion, relations found between demographic trends and land cover changes on a local scale are not necessarily the same for other scales. Previous research about these relationships was mostly executed on a local scale (Birch-Thomsen et al., 2010; Huijun et al., 2002; Sánchez-Cuervo & Aide, 2013; Tsegaye, Moe, Vedeld, & Aynekulu, 2010), thus leaving relations for the full extent of the tropics unknown. The few global studies mostly used aggregated data at country level (DeFries et al., 2010; Kauppi et al., 2006). Furthermore, most of the studies used population increase as a demographic dynamic indicator and not migration. Also, the demographic data used was never spatially explicit but semi-quantitative. Nevertheless, a promising development is an increasing amount of spatial explicit global data with not only improving quality but also higher resolutions (Klein Goldewijk & Ramankutty, 2004; Lele et al., 2010), which opens doors to enable filling the knowledge gap of relations between migration and land cover change on a global scale using a subnational resolution. Earth observing satellites have improved global land cover mapping (Giri, 2005; Loveland et al., 1999). Recently a global dataset on human migration was released by the Center for International Earth Science Information Network (CIESIN) (2011) with a resolution of 1 km^2, which makes it possible to do a spatial explicit analysis on the relation between migration and land cover change on a global scale at sub-national resolution.

1.3 Research Objectives & Research Questions

The research objective is to explore global spatial relations between migration and land cover change with currently existing data.

The related research questions are:

RQ1 Which spatially explicit datasets of migration and land cover changes exist and what are their main characteristics and metadata?

RQ2 What are the relationships between migration and land cover changes if analysed over the tropics?

To answer this research question the following hypotheses shall be tested:

H1: Deforestation is positively related to immigration.
H2: Agricultural land cover increase is positively related to immigration.
H3: Urban land cover expansion is positively related to immigration.

RQ3 Are relationships stronger or different when considering selected countries and finer resolution data?

First, the availability of land cover and migration datasets is investigated. Selected datasets are explored for their characteristics. The next step concerns statistical analysis of relationships between demographic dynamics and land cover change. Three land cover types are tested: forest, agricultural and urban. Although research has found some contrasting relations, the main global processes seemed to be that a positive migration led to deforestation, urban expansion and agricultural increase. Also, since the scale has such an important impact on the results, a comparison will be made with the same statistical analysis on a national scale with the same data.

1.4 Reading Outline

In chapter 2, the methodology and data used for this study is explained. In chapter 3, results of the data exploration, statistical analysis and a scale comparison are presented. In chapter 4 the results are interpreted, discussed and compared with other studies. The approach of the study and its limitations will be discussed and recommendations for future research are presented. Chapter 5 lists the main conclusions of the study.

2 Methods

This research can be structured into three parts (Figure 1). First, available Land Cover and Migration datasets were explored (research question 1). Second, relationships between land cover change and the migration were examined at sub-national scale (research question 2). Third, relationships were explored for the same variables at a smaller extent and higher resolution for two selected countries to test their scale dependency (research question 3).

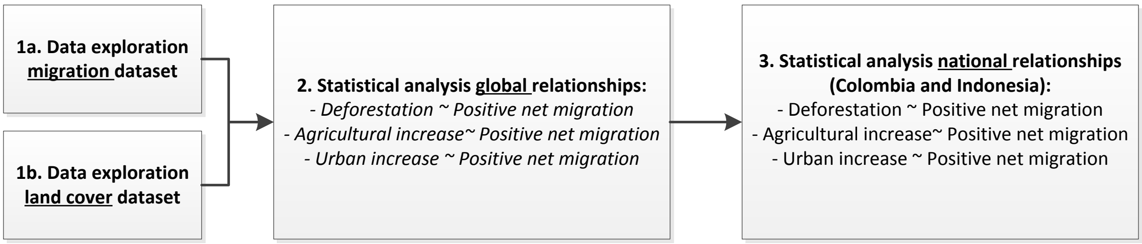


Figure 1: General flowchart of the steps in this study.

In this chapter the area studied and the data used for this end will be explained in paragraph 2.1. Paragraph 2.2 will describe the methods for data pre-processing and finally paragraph 2.3 will explain the methods used for analysis of the data.

2.1 Study Area and Data

2.1.1 Overview

The spatial extent for this research was the tropics (23.5◦N–23.5◦S), because most dynamics in land cover change have been taking place in this region. To remediate known inaccuracies in the migration data at its original resolution of 1 km^2 (see 2.1.2 Migration Data), the data was aggregated to sub-national administrative units. The research was repeated at national level for two selected countries: Colombia and Indonesia, in order to compare the results using different scales. These countries were chosen after visual exploration of the data sets, which revealed that they both exhibit spatially distinct patterns of land cover and migration that include regions with positive and negative land cover change and migration. The temporal extent of this study was the decade 2000-2010, as sufficient land cover data was only available for this period (see section 2.1.2 Land Cover Data).

2.1.2 Land Cover Data

Table 1 shows available land cover datasets, which could solely be found for the period 2000 till 2012. Since not only forest but also agricultural and urban land cover is assessed in this research, the Land Cover dataset derived from the Moderate Resolution Imaging Spectro radiometer (MODIS) was chosen. The Hansen et al. (2013) dataset was deemed unsuitable because it includes only forest land cover information. Another option with sufficient thematic resolution would have been using Global Land Cover 2000 (GLC 2000) and GLOBCOVER 2009. These products are based on the same classification scheme, i.e., The United Nation Land Cover Classification System (LCCS) (Congalton, Gu, Yadav, Thenkabail, & Ozdogan, 2014). However, the data are partly inconsistent since they are based on data from different sensors, (GLC 2000: SPOT 4 (Bartholomé & Belward, 2005); GLOBCOVER: ENVISAT MERIS (Bontemps et al., 2011)) with different spatial resolutions, and employed slightly different classification techniques.

Table 1: Overview of available global raster land cover datasets and their charasteristics.

Name	Spatial Resolution	Classification scheme	Temporal extent	Sensor	Source	Link Source
Tree canopy cover for year 2000	1 arc-second (30 m)	Percent Tree Cover in 2000	2000	Landsat	Hansen et al (2013)	http://earthenginepartners.appspot.com
Global forest cover loss 2000–2012	1 arc-second (30 m)	Loss/No loss/Water or no data	2000-2012	Landsat	Hansen et al (2013)	
Global forest cover gain 2000–2012	1 arc-second (30 m)	Gain/No gain/Water or no data	2000-2012	Landsat	Hansen et al (2013)	
Landsat Forest Cover Change 2000-2005	1 arc-second (30 m)	Persistent Forest/Forest Loss/Forest Gain/Persistent Non-forest	(1975;1990);2000-2005	Landsat	www.landcover.org	http://www.landcover.org/data/
MODIS Vegetation Continuous Fields product (VCF)	250 m	Percent of pixel area covered by land cover type	2000- 2009	MODIS		http://glcf.umd.edu/data/vcf/
MODIS Land Cover	500 m/5' x 5'(10 km)/0.5° x 0.5°(50 km)	IGBP: 17 land cover classes	2001-2012	MODIS	www.landcover.org	http://glcf.umd.edu/data/lc/
Global Land Cover (GLC) 2000	1km at Equator	LCCS	2000	SPOT 4	European Commission Joint Research Center	http://www.eea.europa.eu/data-and-maps/data/global-land-cover-2000-europe
ESA GLOBCOVER 2009 Land Cover map	1km at Equator	LCCS	2009	ENVISAT MERIS	European comission	http://due.esrin.esa.int/globcover/

The Classification System International Geosphere-Biosphere Programme (IGBP), used for the MODIS Land Cover dataset, distinguishes 17 land cover classes, which include the forest, cropland and urban classes required for this study. The method of producing the map was supervised classification using decision tree trained on sample sites from high resolution imagery (mostly Landsat) together with additional information (Friedl et al., 2002). A cross-validation analysis identified an overall accuracy of all land cover classes in the global product of 75% correctly classified. However, the range in class-specific accuracies is large (Friedl et al., 2010). Especially the forest classes and agricultural classes show good accuracies (Friedl et al., 2010). The urban land cover class overall's accuracy was also high with 93% (Schneider, Friedl, & Potere, 2009). Except for North America, the overall accuracy per continent turned out to be higher than the average 75%.

The original spatial resolution of the Modis Land Cover dataset is 500 m but the product is available at multiple resolutions (see table 1). For this study, the 0.5° (10 km) resolution was chosen in which the majority of a land cover classes were assigned to each cell global analysis. For the presented study the maps of the years 2001 and 2010 were used to derive land cover changes. These came closest to the period of the migration data, which is available per decade between 1970-2010 (see next section).

2.1.3 Migration Data

The migration dataset from CIESIN was only available that covers global extent and is spatially explicit at the moment. This dataset contains estimates on net migration per 30 arc-seconds (+- 1 km) cells, based on time-series population distribution grids combined with United Nations (UN) and other data on birth on death rates (CIESIN, 2011). Net migration is defined as population change minus the births-deaths rate (natural increase).

The migration data is available per decade from 1970 until 2010. As a result of the land cover data availability the decade 2000 – 2010 was chosen. However, since the population distribution grid for the year 2010 is not based on real census data, which were not available at the time of creation of migration dataset, the migration estimates of the decade 2000 – 2010 are potentially less accurate than the other decades (CIESIN, 2011). To reduce data uncertainty the net migration data was aggregated to sub national level. Since the census data, on which the migration estimations are largely based, are also per administrative level it made sense to use this type of units instead of square cells. Next to the net migration values, density grids are also available for the years 2000 and 2010 which could be used to calculate migration relative to population density.

2.1.4 Sub National Administrative Boundaries

The Global Administrative Areas (2012) dataset contains administrative boundary features of countries and sub national regions at multiple levels. Level 1 is the coarsest level after national, which is level 0. The administrative units at level 1 differ between countries. For example, level 1 units correspond to provinces in Indonesia but to states in Brazil. Level 2 is a subdivision of the level 1 units and can correspond to for example regencies or municipalities. For all countries included in this study level 1 units exist. While for most countries level 2 administrative boundaries are provided, only a few countries have data at the third or fourth level, which include village and commune respectively as units.

2.2 Data Pre-processing

For all pre-processing steps GIS and statistical programs ArcGIS and R were used.

2.2.1 Choosing Sub National Administrative Units

Since the sizes of the administrative units vary widely over the 168 countries situated in the tropics when using a single level of the Global Administrative Areas dataset (GADM), a mix of the first and second level of the data was used to equalize the areas sizes more. Level 1 sub national regions were the first choice because it is available for all countries and resulted in 1718 regions. However, as can be seen in figure 2, some level 1 regions are very large with sizes exceeding 30 times the mean size across all countries. The standard deviation of the area of level 1 units is high with 97,884 km^2, which is three times the mean value of 32,148 km^2 (see table 2). Visually, the distribution of the level 1 regions can be divided into a majority of regions with an area between 0 and 20,000 km^2 and a few very large outliers (see figure 2). Also, table 2 shows that 75% of the areas are smaller than 25,716. The other 25% includes regions sizes up to almost 2 million km^2, which is 8 times as large. To reduce the great dispersion of region areas, for countries containing regions with a size larger than 20,000 km^2 level 2 units were chosen. This was done for 18 countries: Angola, Australia, Bolivia, Brazil, Chad, Democratic Republic of the Congo, Ethiopia, India, Indonesia, Kenya, Mali, Mauritania, Niger, Peru, South Sudan, Sudan, Venezuela and Yemen.

The distribution of the regions sizes after harmonization can be seen in figure 3. Unfortunately, the distribution still has a high peak for smaller values. This was caused mainly by Brazil as this country was responsible for the very large regions in level 1, but after harmonization also for a many very small regions (see figure 3). However, the outliers are less extreme high (table 2). Combining level 1 and 2 results in 8155 regions; the mean size of the selected regions is 6,768 km^2.

In the national level analysis for the countries Colombia and Indonesia it was decided to use level 2 subnational administrative boundaries to maximize the resolution. In the global analysis Indonesia is also used with this level, while Colombia is used with level 1. Level 2 was the finest unit available for these countries.

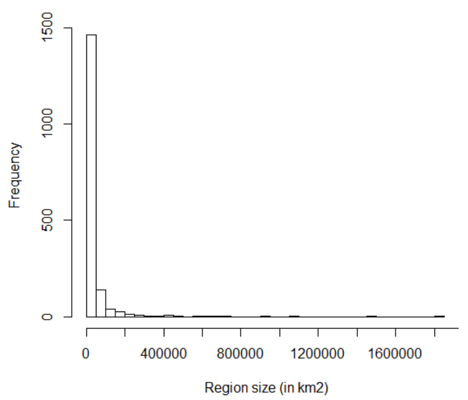


Figure 2: Histogram distribution area sizes of regions level 1.

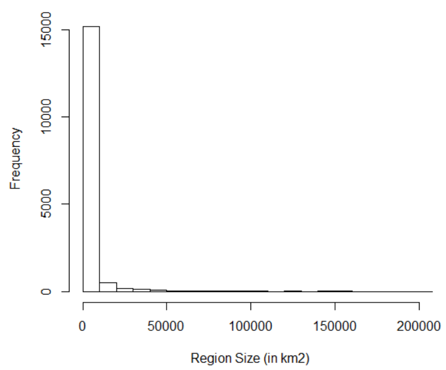


Figure 3: Histogram distribution area sizes of regions level 1 and 2 combined.

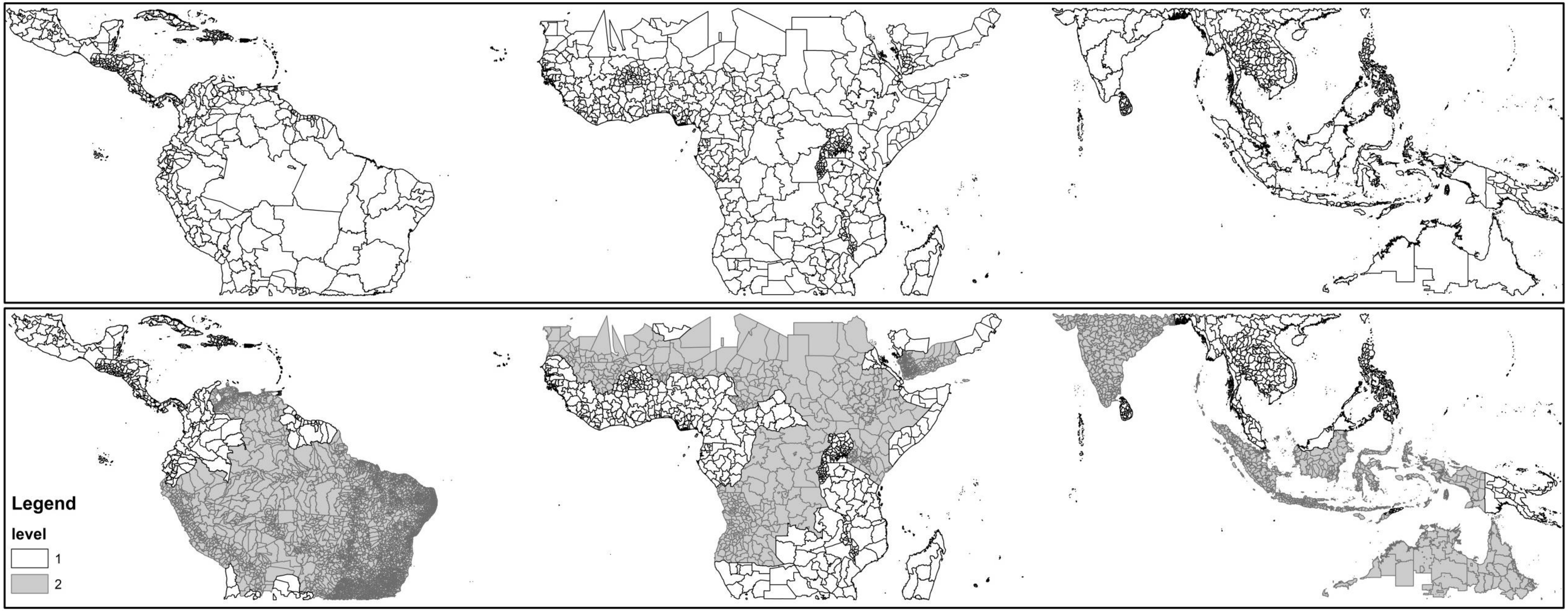


Figure 4: Upper map: administrative regions in the tropics according to level 1. Bottom map: the selected combination of level 1 and level 2 regions as used in the analysis.

Table 2: Summary statistics of subnational region sizes in km^2 for level 1 and for the mix of level 1 & 2.

Level	Number of Regions	Median	Mean	Std.dev	Min	Quantile Areas 25%	50%	75%	100%
1	1,718	5,502	32,148	97,884	1	757	5,502	25,716	1,839,446
1&2 combined	8,155	1,010	6,770	19,668	1	311	6,770	4,390	454,000

2.2.2 Land Cover and Migration Data

Figure 5 presents a flowchart including all data pre-processing steps that were done in order to create the required datasets for both exploratory and statistical analysis. From the global raster input datasets, change maps for forest, cropland, urban and net migration in the tropics in the period 2000 - 2010 had to be derived with an aggregated spatial resolution of sub national regions.

First, all the datasets were cropped to the extent of the tropics (x = -180, 180, y= 23.5, 23.5) to reduce the datasets to the part of interest. Second, the IGBP land cover classes of the MODIS data were generalized to reduce the land cover classes considered in this study. Table 3 shows that the original MODIS land cover data contain 17 land cover classes that were aggregated for this research to 10 classes. For answering the research questions and testing the hypotheses it was not important to distinguish between individual forest types and hence all individual forest classes were generalised to one forest class. The choice for reclassifying the agricultural land was more difficult as classes such as "Cropland/Natural Vegetation mosaic" and "Grassland" could also represent managed land. However, since it was uncertain how much they represented agriculture land, as they also include natural grassland and vegetation, the choice was made to use solely the cropland class for testing the second hypothesis of the relation between migration and agriculture.

Table 3: Reclassification table from original IGBP land cover classes to 10 generalized classes.

MODIS value	IGBP LC Classes MODIS data	Reclass Name
0	Water	Water
1	Evergreen Needleleaf forest	Forest
2	Evergreen Broadleaf forest	Forest
3	Deciduous Needleleaf forest	Forest
4	Deciduous Broadleaf forest	Forest
5	Mixed forest	Forest
6	Closed shrublands	Other natural vegetation
7	Open shrublands	Other natural vegetation
8	Woody savannas	Other natural vegetation
9	Savannas	Other natural vegetation
10	Grasslands	Grasslands
11	Permanent wetlands	Wetlands
12	Croplands	Croplands
13	Urban and built-up	Urban
14	Cropland/Natural vegetation mosaic	Cropland/Natural vegetation mosaic
15	Snow and ice	Snow and ice
16	Barren or sparsely vegetated	Barren or sparsely vegetated

Third, change maps were created for each land cover type; forest, cropland and urban, by comparing the simplified land cover maps from 2001 and 2010. Per pixel it was identified if the respective land cover type has disappeared or newly emerged. For a pixel that had changed into the particular land cover type, for example forest, a value of +1 was signed to. If a forest pixel in 2001 was also forest in 2010, the value 0 was assigned. When a forest pixel turned into another land cover type it was given the value -1. The raster change maps for cropland and urban were created with the same method.

Fourth, with the zonal statics function in ArcMap the sum of the values of these change raster's were calculated for each administrative unit at sub national level. For the land cover datasets this resulted in a net change value in km^2 per region. The zonal statistics sum function was also applied to the net migration raster dataset, resulting in a net migration value of persons that migrated from or into a region.

Additionally, relative changes were calculated by dividing the net land cover change values per region by the sum of total km^2 of the respective land cover type in 2001. Calculating a relative increase was a problem for regions where a land cover type newly appeared in 2010 since it would lead to a division by zero. For forest this was the case in 170 regions and for cropland in 345 regions. As a solution the change values were taken relative to an area of 10 km^2 in 2001 as this is the area size of one land cover MODIS pixel, which led to representative increase percentages for these regions. For the migration data, changes were relative to the sum of the gridded population density in 2000 over each region. Since the population density had a unit of persons per km^2, it equalled population counts per grid cell at the datasets resolution of 1 km^2. All produced variables that were used in this study can be found in table 4 with a description.

Table 4: Variables for land cover change and migration used in further analysis.

	Variable Name	Variable Description
Land cover Change	**Forest Change**	Signed total forest change per region 2001- 2010 in km^2
	Relative Forest Change	Forest change relative to the forest area in the region in 2001 as a percentage
	Deforestation	Binomial values where 0 = no net deforestation (net change >= 0) and 1 = net deforestation (net change<0)
	Forest 2001	Area of forest per region in the year 2001 in km^2
	Forest 2001 %	Area of forest in the year 2001 in km^2 relative to the total area of the region
	Cropland change	Signed total cropland change per region 2001- 2010 in km^2
	Relative Cropland	Cropland change relative to the cropland are in the region in 2001 as percentage
	Cropland 2001	Area of cropland per region in the year 2001 in km^2
	Cropland 2001 %	Area of cropland relative to the total area of the region in the year 2001 in km^2
	Urban 2001	Area of urban per region in the year 2001 in km^2
	Urban 2001 %	Area of urban relative to the total area of the region in the year 2001 in km^2
Migration	**Migration**	Net migration in persons per region 2000 -2010
	Relative Migration	Net migration 2000-2010 relative to population size in the region in 2000

All the land cover change variables are net in the sense that the total increase of the land cover type minus the loss of this land cover type within a subnational region is taken. Within a region the land cover type can have newly appeared in one location while it can have disappeared in another at the same time resulting in a net value.

The variables did not include land cover change value for all the 8155 regions. Solely for the regions that had either the particular land cover type in 2001 or in 2010 a change value was calculated. The regions where the land cover type was never located are labelled not available (NA) and left out of the analyses.

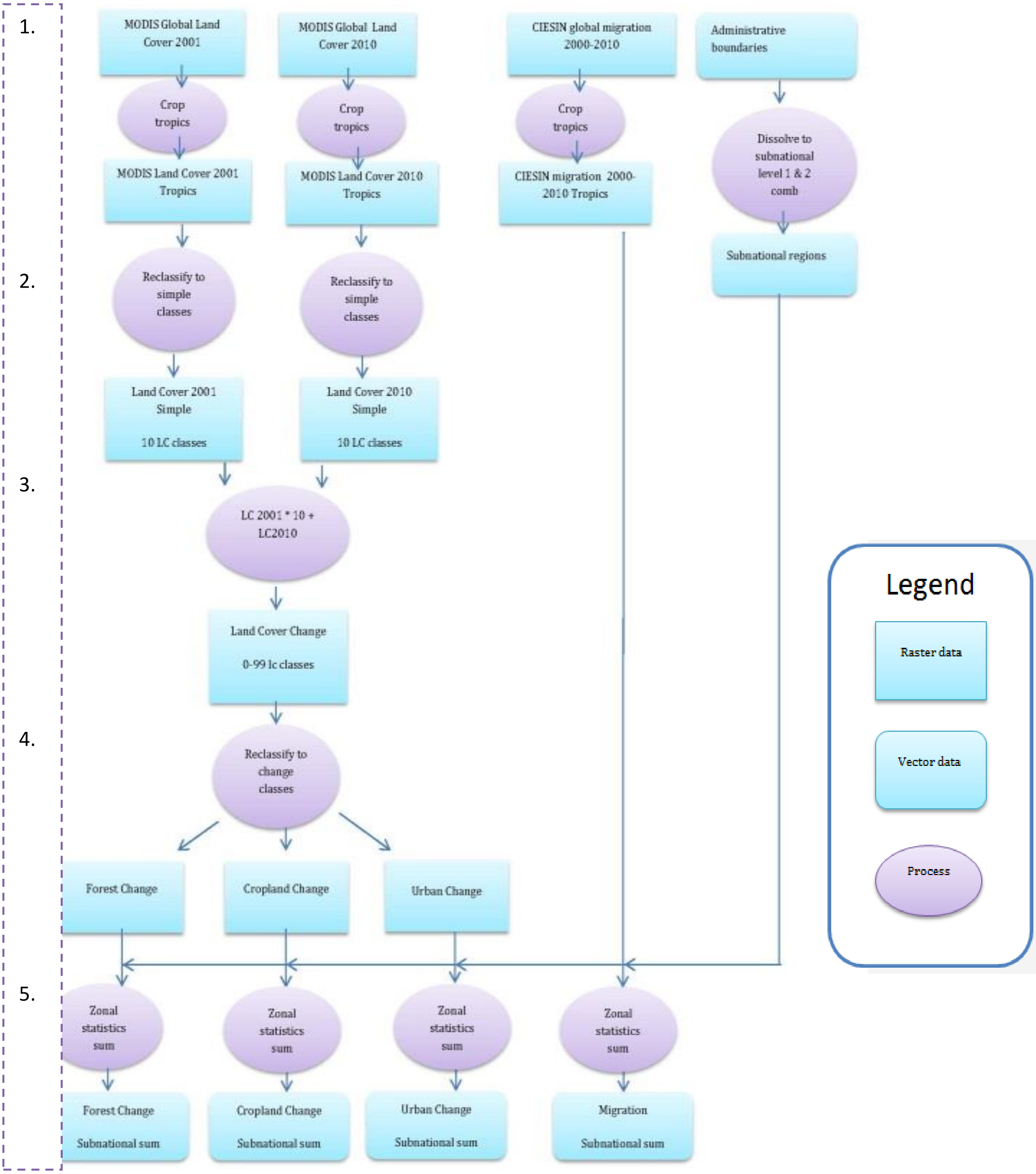


Figure 5: Flowchart of the pre-processing steps for preparing the input data for exploration and statistical analysis.

2.3 Analysis Methods

2.3.1 Exploratory Data Analyses

The characteristics of the datasets were explored by both visual assessment and descriptive statistics. Histograms, ranges, means, stand deviations and quartiles were calculated for all variables. The histograms show the data distributions, which can help selecting the method for analysing statistical relationships and which gave information about extreme values. Furthermore, the datasets were evaluated for possible artefacts by zooming in on the extreme values and finding explanations for them.

2.3.2 Statistical Relationships Analysis

To answer research question 2 the following hypothesis were tested:

H1: Deforestation is positively related to immigration.
H2: Agricultural land cover increase is positively related to immigration.
H3: Urban land cover expansion is positively related to immigration.

Choosing the statistical approach for exploring the relationships was an explorative process of trial and error to identify the most suitable approach for the datasets. First, the land cover change values per sub national region were plotted against the migration data to see if a relationship was visible. This was done for both absolute and relative values of each land cover type: forest, cropland and urban. Additionally, Pearson's correlation coefficients(*r) w*ere calculated between all variables (for an overview of the variables see Table 4), which gives information about the strength and direction of a possible linear relationships. Furthermore, since the forest cover change data included many zero values, a logistic regression was tested, with the forest change data transformed to binary values 0 (no deforestation) and 1 (deforestation). The model fit was evaluated based on pseudo R square. The McFadden pseudo R-square (McFadden, 1973) was chosen, for which a value close between 0.2 and 0.4 to one is considered to be a good fit (Clark & Hosking, 1986; Domencich & McFadden, 1975).

Next, the data were analysed using regression trees, which can deal with non-linear relationships as well as interaction between explanatory factors. The classification and regression trees (CART) method classifies data into more homogenous subsets by recursively splitting the data in two based on a threshold value of a single explanatory variable (Breiman, 1984; Waheed, Bonnell, Prasher, & Paulet, 2006), with the goal to create subsets that are more homogenous. With this method for each land cover change type (expressed as both absolute and relative change values) it was tested if and to what extent it could be explained by the absolute and relative migration variables. Also, the absolute and relative amount of each land cover type per region in 2001 were included as explanatory variables, in order to test for their possible interaction with land changes. Lastly, for forest change, the cropland change variables were also included, to detect potential relations. Table 5 presents the variables used in each regression trees.

Table 5: Response and explanatory Variables used to test relationships between land cover change and migration with the CART regression trees method. All the absolute variables are in km^2.

Response Variable	Explanatory Variables
Absolute Forest Change	Forest 2001 + Cropland 2001 + Urban 2001 + Forest 2001 % + Cropland 2001 % + Urban 2001 % + Migration + Relative Migration + Cropland Change + Relative Cropland Change

Relative Forest Change	Forest 2001 + Cropland 2001 + Urban 2001 + Forest 2001 % + Cropland 2001 % + Urban 2001 % + Migration + Relative Migration + Cropland Change + Relative Cropland Change
Absolute Cropland Change	Forest 2001 + Cropland 2001 + Urban 2001 + Forest 2001 % + Cropland 2001 % + Urban 2001 % + Migration + Relative Migration
Relative Cropland Change	Forest 2001 + Cropland 2001 + Urban 2001 + Forest 2001 % + Cropland 2001 % + Urban 2001 % + Migration + Relative Migration

The R package rpart (Therneau & Atkinson, 2012) was used to produce the regression trees. The rpart package includes multiple parameters which can be adapted (Therneau, Atkinson, & Ripley, 2006). The minsplit, a minimum number of observations required in a node to even calculate a new split or the minbucket, and the minbucket, a minimum number of observations in an end node, can be adjusted. The default settings for minsplit is 20 and for minbucket minsplit/3. Furthermore, the complexity parameter (cp) defines the minimum factor of R square improvement that each step should have. The function of this parameter is to reduce computation time because trees are immediately pruned off for splits that are not significant. The default setting is 0.01.The xval parameter determines how many cross-validation that will be done (default = 10). Since trying different setting for all these parameter did not improve the quality of the trees, for all the trees the default setting were maintained.

The predictive power of a tree was assessed with the use of r-square and 10-fold cross validation. In order for an additional split to be significant the r-square should be increased. The cross validation, which takes a sample from the dataset and compares the model performance of that sample to the model performance of the complete dataset, was also shown in a plot of relative error against number of splits. A split is significant if the relative error ($1- r^2$) decreases. The statistical analysis was done at the global scale covering the tropics and repeated at the national scale for Colombia and Indonesia.

3. Results

In this chapter the results of the global data exploration will be presented in section 3.1. In section 3.2 the statistical analysis of the relation will be shown. Section 3.2 is the results of the national analysis.

3.1 Global Data Exploration

3.1.1 Land Cover Data

Table 6 presents land cover changes in the tropics between 2001 and 2010 based on the MODIS land cover data. The largest change was an increase in cropland area of 378,000 km^2, which is an increase of 21% relative to the cropland area in 2001. The second largest change was a negative change in forest area of 326,000 km^2, which is 2% relative to the total forest area in 2001. According to the data, urban land decreased with 5,000 km^2, which is contrary to the general perception that urban land increased over this period. Although the accuracy of the land cover class urban was high (Schneider et al., 2009), it could be that the chosen resolution of the dataset (10 km) was too low. Consequently, it was decided that this dataset was not suitable for urban change detection, resulting in the exclusion of urban land cover change in this research. Another remarkable result is that wetlands increased. In Appendix 1 plots are presented that show the yearly number of pixels for forest, cropland, urban and wetland derived from the yearly MODIS land cover datasets. From this can be derived that although the number of pixels for each land cover type can vary within the period, especially forest and urban, there is a main trend, indicating that the calculated net changes can be considered to be reasonable. However, it could still have had an effect.

Table 6: Areas per land cover type in 2001, 2010 and the absolute and relative change per land cover type in the tropics. The relative change is the area change relative to the total area of the landcover type in 2001.

LC Class	Area 2001 (x 1000 km^2)	Area 2010 (x 1000 km^2)	Absolute Net Change (x 1000 km^2)	Relative Net Change
Water	159,534	159,481	-53	0%
Forest	13,838	13,512	-326	-2%
Other natural vegetation	19,645	19,570	-75	0%
Grasslands	3,345	3,261	-84	-3%
Wetlands	130	319	188	145%
Croplands	1,832	2,210	378	21%
Urban	79	74	-5	-7%
Cropland/Natural vegetation mosaic	4,339	4,474	135	3%
Barren or sparsely veg	7,988	7,830	-158	-2%

Forest Cover Change

In figure 6, maps are shown for forest change and relative forest change in the tropics between 2001 and 2010. Within each continent and almost in every country both deforestation and reforestation occur. The relative change map shows several extreme values, which are typically caused by a low value of the forest area in 2001.

From the 8155 regions 2872 have a value for forest change. Regions with no forest in 2001 and in 2010 are not included. Regions with forest area in 2001, which did not change in 2010, have the value zero. Overall, there are slightly more region with net deforestation (1222) than with net reforestation (1068 regions).

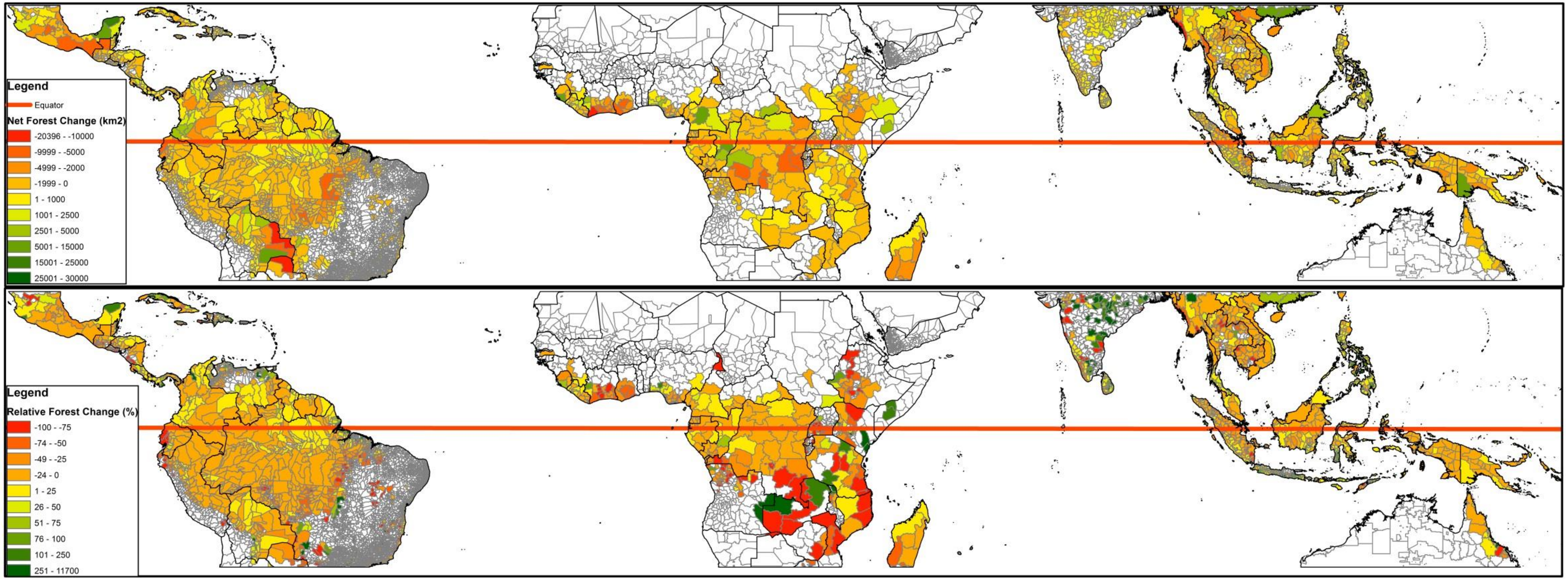
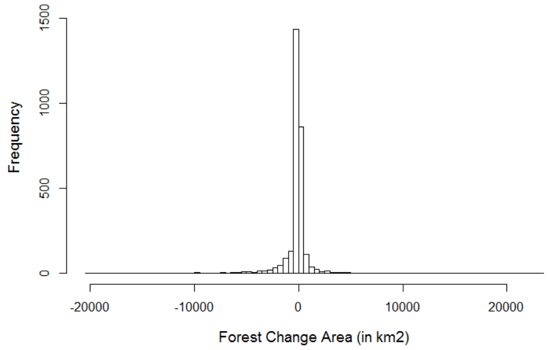


Figure 6: Forest cover change per subnational region 2001 – 2010 in the tropics. upper map: the absolute net change in km^2. Map below: forest change relative to the forest area in the region in 2001.

Table 7: Summary statistics of regional forest change in km^2 and in precentage.

	Median	Mean	Std.dev	Quantile Areas				
				Min	25%	50%	75%	Max
Forest Change (km^2)	0	-133	1,298	-20,400	-115	0	81	23,000
Relative Forest Change (%)	0	817	6,673	-100	-18	0	12	217,000

Table 7 shows that half of the values of forest change of the tropics were between – 115 and 82, while the range of the values was much wider: -20,400 to 23,000. This can also be seen in the left histogram in figure 7. There is high peak around zero and only a few with much higher values.



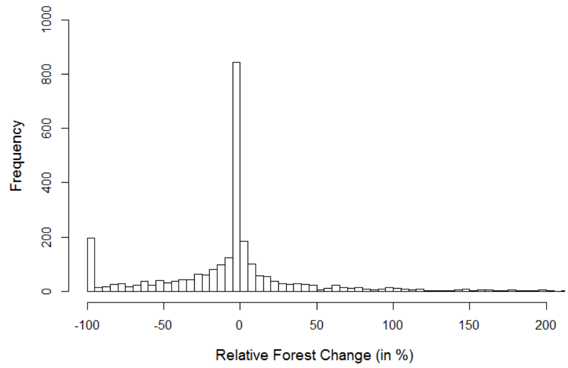


Figure 7: Global forest change per region histograms. Left: absolute change and Right: net change relative to forest area in 2001. 269 of the 2872 regions with forest change values above 200% forest increase (not shown).

The relative values include some extreme high percentages because in regions with almost no forest in 2001, a small increase in forest area is recorded as a large relative increase. However, the majority of regions are characterised by a forest cover change ranging between -100% and 200%. The peak in the histogram with relative values shows a peak for -100%, which were also often cases of a low value of forest area in 2001 in the region. Most of these regions, such as a few in Brazil, have an initial forest area in 2001 of less than 200 km^2. Remarkable were two regions in Tanzania where in the entire area of forest in 2001 was deforested in 2010, which was in both regions an area of around 1600 km^2. One of the regions named Singida is characterised by a savannah ecosystem and the deforestation is mostly driven by charcoal production. When looking at the yearly data of MODIS the deforestation took place between 2001 and 2003. The cells all turned into other natural vegetation.

In appendix 1 an overview of all basic statistics for forest change can be found not only for the harmonized mixed level 1 & regions used here, but also for only level 1 to show what was the influence of choosing this mixed level. In general the shapes of the distributions looked the same. However, after harmonization, the ranges were smaller and the standard deviations decreased.

Cropland Cover Change

Cropland dynamics were calculated for 1,822 regions from which 175 had no change, but did have cropland in 2001. Furthermore, more regions (894) had a cropland increase than the 753 regions with a cropland decline.

Figure 8 shows that tropical regions with a positive net change are mostly located in the northern hemisphere, especially in North Africa. In contrast, except for the countries in South America, most regions with a negative change are located on the southern hemisphere.

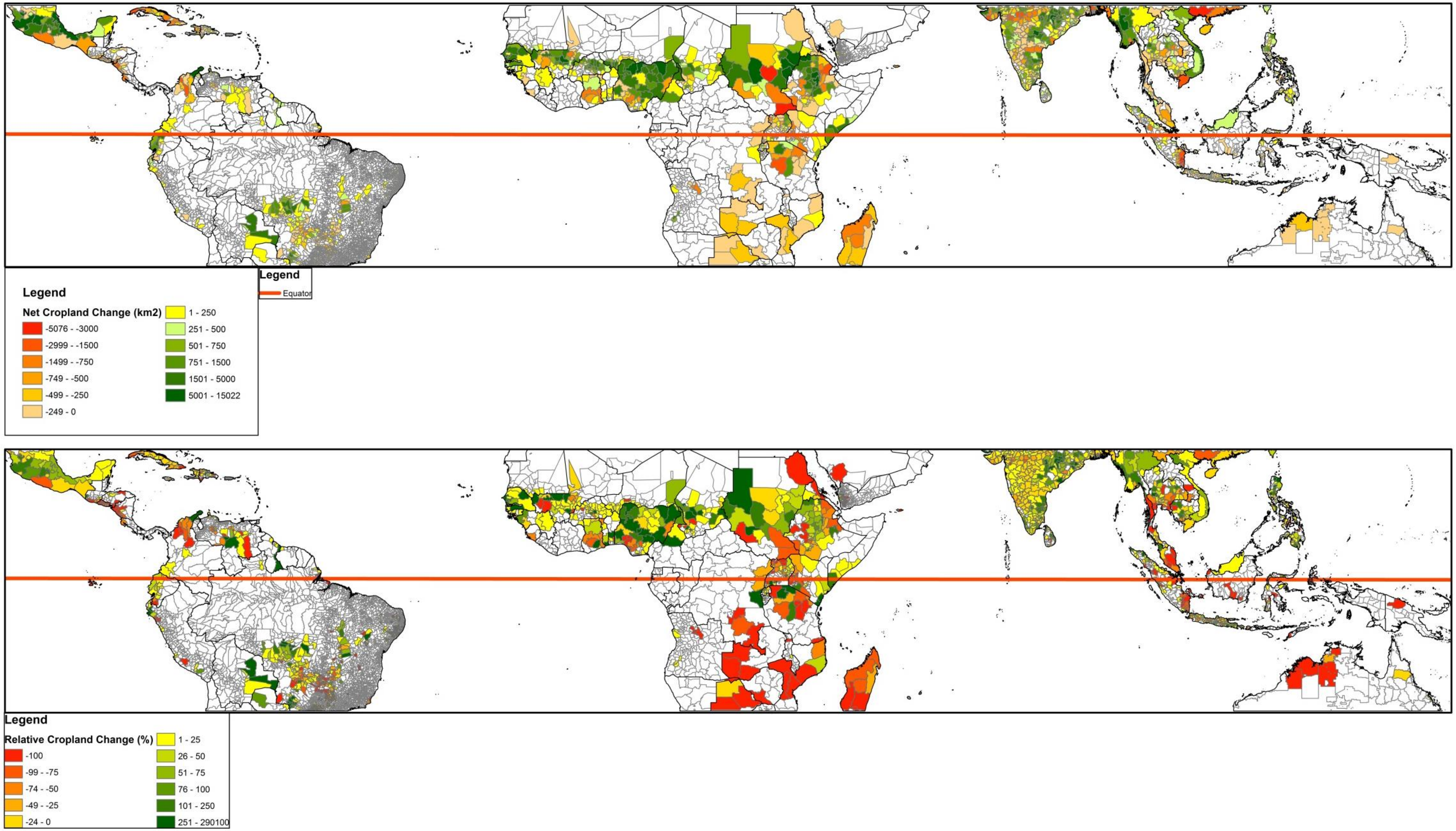
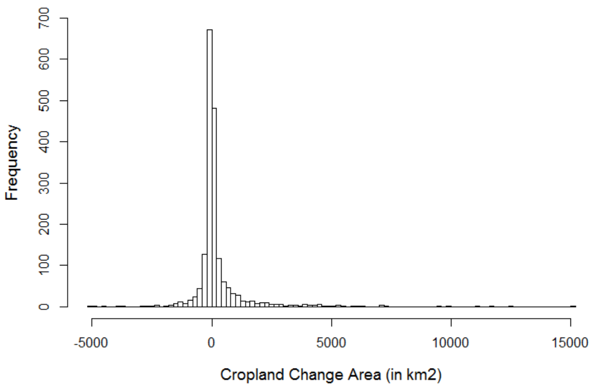


Figure 8: Cropland cover change per subnational region 2001 – 2010 in the tropics. Upper map: the absolute net change in km^2. Map below: cropland change relative to the cropland area in the region in 2001.

Table 8: Summary statistics of regional cropland change in km^2 and percentage.

	Median	Mean	Std.dev	Quantile Areas				
				Min	25%	50%	75%	Max
Cropland Change (km^2)	0	229	1,150	-5,076	-86	0	163	15,022
Relative Cropland Change (%)	0	6,149	31,141	-100	-64	0	230	529,300

In table 8 it can be seen that half of the cropland change per region had a value between -86 and 163 with a median of 0. However the high standard deviation shows that cropland change across regions is highly variable. In figure 9 it can be seen that the frequency of regions with 100% loss of cropland was very high.



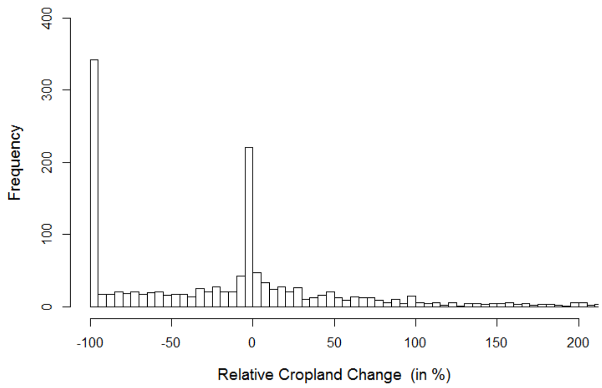


Figure 9: Global cropland change per region histograms. Left: absolute change and right: net change relative to cropland area. 417 of the 1822 regions with cropland change had values above than 200% (not shown).

3.1.2 Migration Data

In figure 10 it can be seen that in the majority of the regions is characterized by outmigration stronger than in migration. The most apparent is the lower map showing typical relative migration values between -24% and 0%. Positive net migration is mostly located along the west coast of Africa, the north of Africa and in Malaysia in Asia. Hence, in Mexico and central Africa are many regions with a negative net migration.

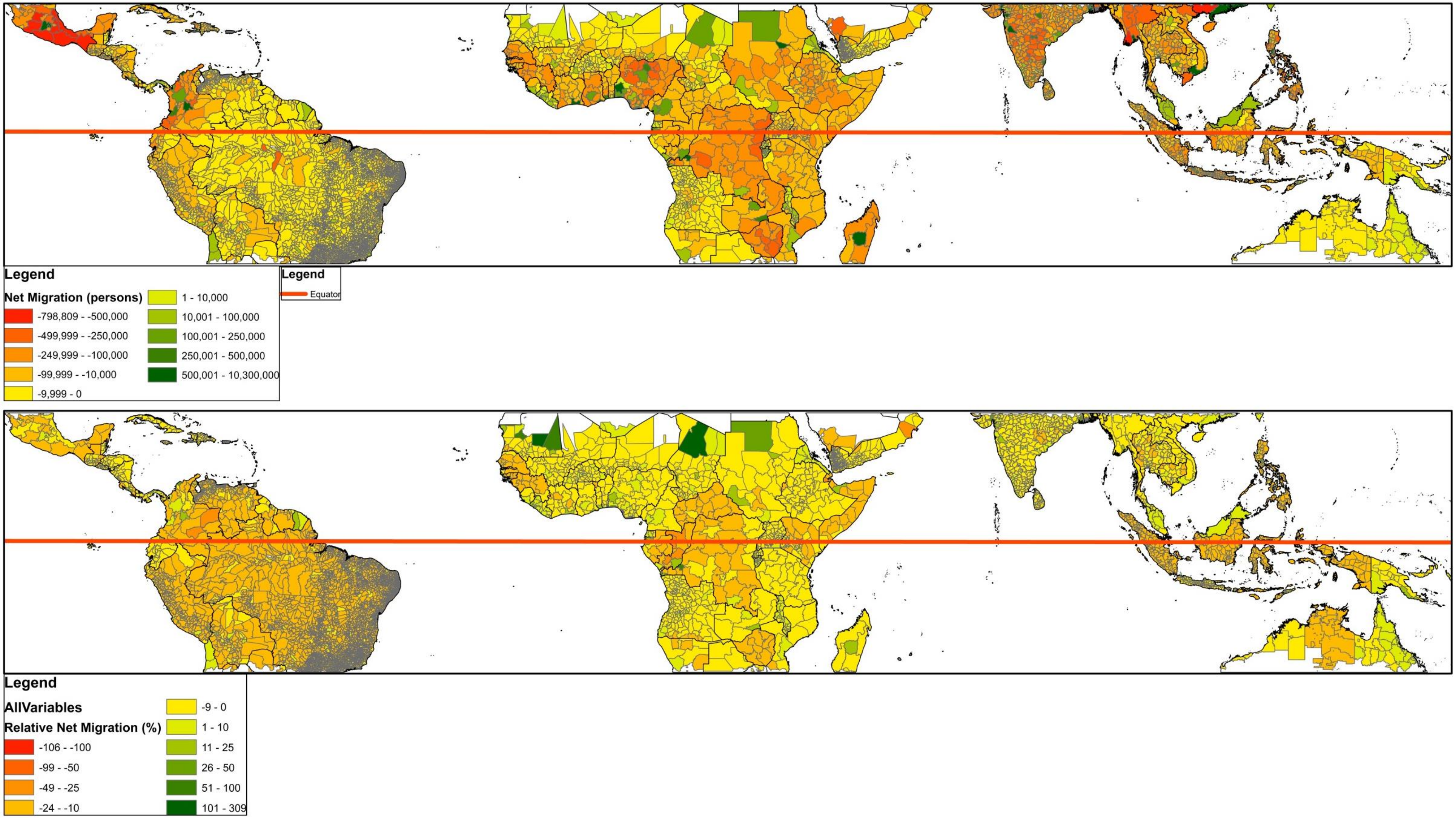


Figure 10: Migration per subnational region 2001 – 2010 in the tropics. upper map: the absolute migration in persons. Map below: migration relative to the population in the region in 2001.

Table 9: Summary statistics of regional migration in persons and in percentage.

	Median	Mean	Std.dev	Quantile Areas				
				Min	25%	50%	75%	Max
Migration (persons)	-2,700	-824	222,890	-788,000	-7,440	-2,700	-990	10,300,000
Relative Migration (%)	-15	-11	11	-106	-18	-15	-6	309

Table 9 provides summary statistics of migration at regional level. More than 75% of the regions had a negative net migration value.

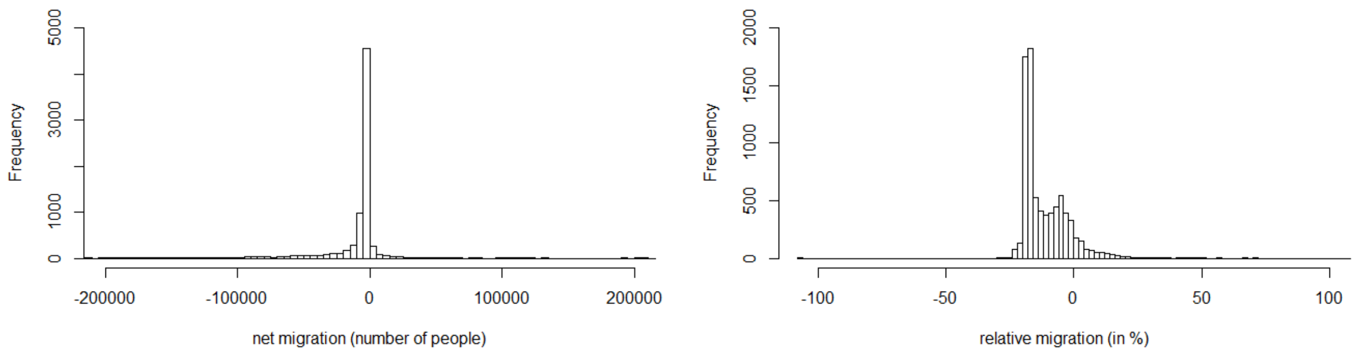


Figure 11: Global regional migration histograms. Left: absolute change. 102 regions had a values above 200,000 (not shown) and Right: change relative to population in 2000.

The histogram of the absolute values in figure 11 show a peak a little bit left to zero. Remarkable are some extreme high values which are not shown the first histogram. The histogram for the regional relative migration values shows a bimodal distribution with one peak at – 15 and the other one around -5 similar to the absolute data.

Table 10 presents the five highest regional migration values, also expressed in terms of migration per km^2. The net migration in the region Maharashtra in India was the highest with 9537 people per km^2.

Table 10: Top 5 regions with highest positive migration values.

Country	Region Name	Net – migration 2000-2010	Relative Migration	Size Region (km2)	Migration/Km2
India	Maharashtra	+10,290,391	5469%	1,079	9537
India	West Bengal	+5,126,048	4798%	7,375	695
India	West Bengal	+6,801,093	4072%	6,972	975
China	Guangdong	+7,434,353	896%	97,721	76
Thailand	Bangkok Metropolis	+5,732,148	7632%	1,883	3044

3.2 Global Data Relations Exploration

3.2.1 Linear Correlations to Explain Land Cover Changes

In figure 12 scatter plots are presented of each forest variable against each migration variable. Unfortunately no clear relation is visible for any pair of variables. In addition, as can be seen in figure 13, the correlation coefficients between any of the land cover change and migration variables were low. Even the highest correlation found [r = 0.16, n = 1822], between the relative cropland and relative migration variables, was lower than 0.19, an indicator boundary from which a relation has a (weak) relationship. Next to the correlation between the change in land cover variables and migration variables the correlation between forest change and cropland change were also calculated (see correlation plot figure 13). Since the correlations with migration were low, it was interested to see if expected relation of cropland expansion resulting in deforestation could be found. However, these correlations were also low [r < 0.19].

There are especially many zero values of regions that had a particular land cover type in 2001 which did not undergo any change till 2010. A binary variable was created that codes the data as ones for

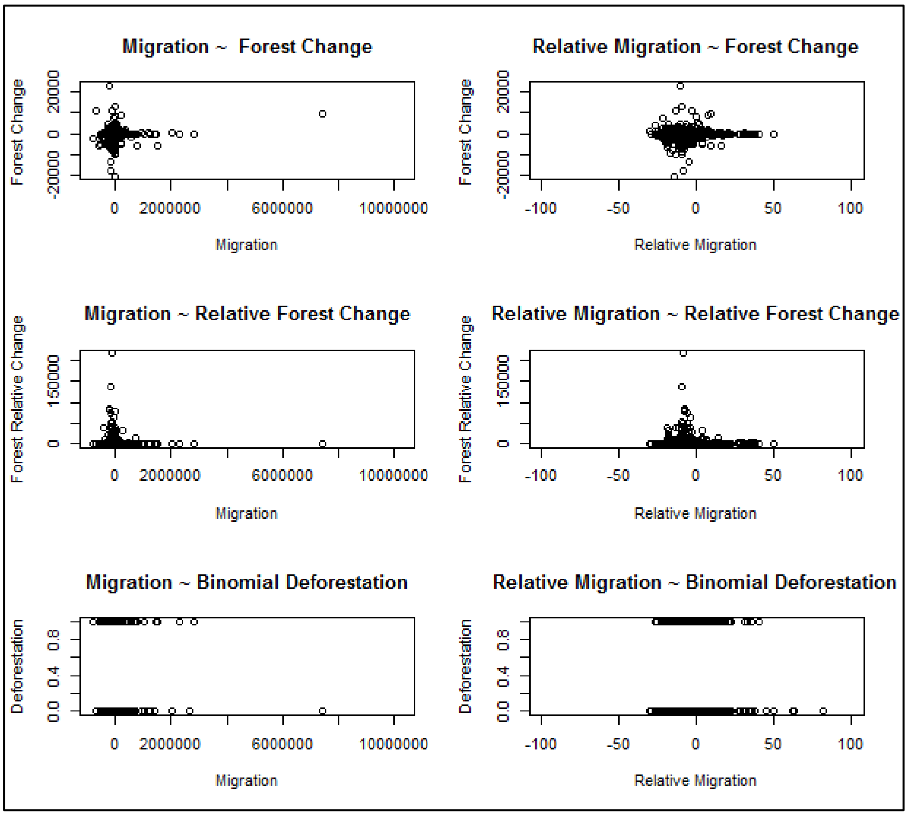


Figure 12: Scatter plots of land cover change variables versus migration variables Top: forest change to migration variables, middle: relative forest to migration variables and below: the binomial forest cover change variable to the migration variables.

regions with deforestation and zero for regions with no change or reforestation. However, as can be seen in the plots, also this variable did not reveal any relationship between forest change and migration. In addition, McFadden's pseudo-R squared of logistic regression of binary forest change on migration was very low: 0.0096.

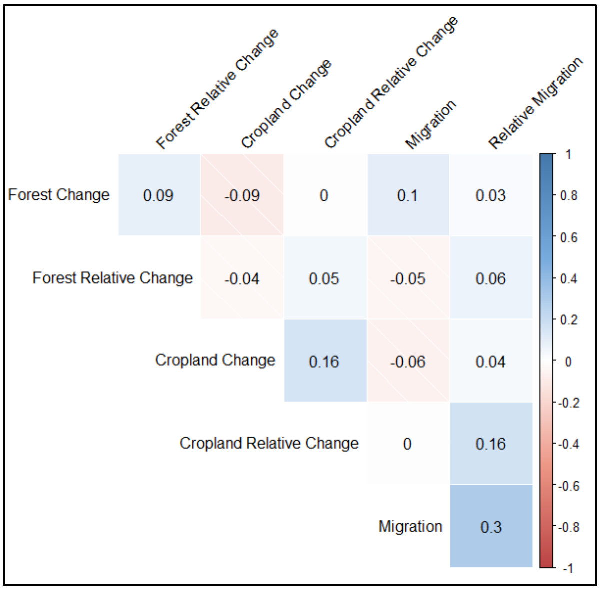


Figure 13: Correlation coefficients between all possible combinations of the forest change and cropland change variables to the two migration.

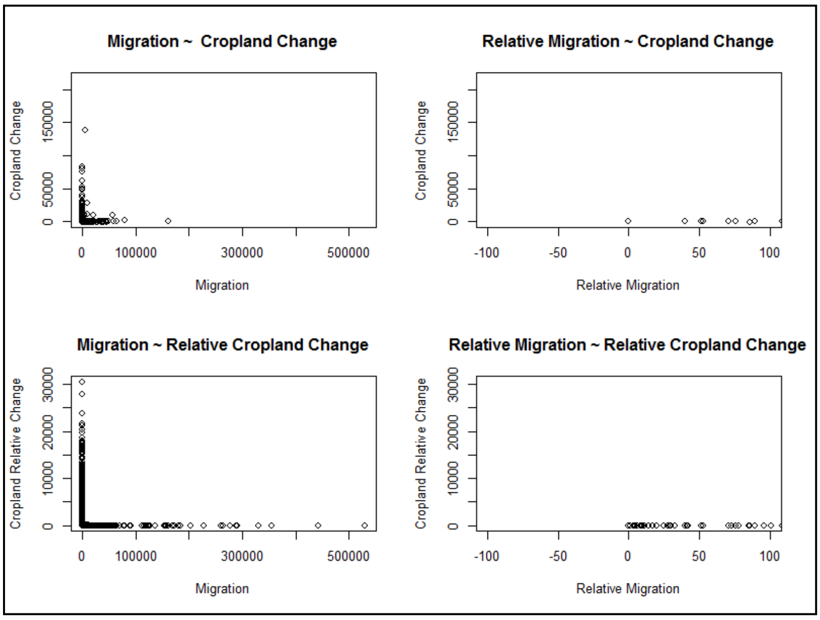


Figure 14: Scatter plots of cropland variables to migration variables. Up: cropland change to migration values, Below: relative cropland change to migration variables.

The scatter plots presented in figure 14 did not provide evidence for clear linear relations between cropland cover change and migration. Accompanied by the weak correlation values for cropland versus migration it can be concluded that no linear relation was found between dynamics of cropland area and migration at sub-national level.

3.2.2 Regression Trees

Regression trees were produced to explain forest change, relative forest change, cropland change and relative cropland change. In an additional analysis, cropland change was included as an explanatory variable for forest change as it would be interesting to assess whether deforestation could be explained by cropland expansions.

The regression tree for absolute forest change first splits data based on the explanatory variable forest area in the region in 2001. For the regions were this was below 5239 km^2 the mean forest change was -20 km^2. The second split uses the variable Cropland cover area in 2001. At the third and fourth split the migration variables are used. However, these splits are not significant as the relative error did not decrease (figure 17). As a result, pruning this tree led to total deletion of the tree, except for the root. There was no explanation from the independent variables that can make a significant split of the data.

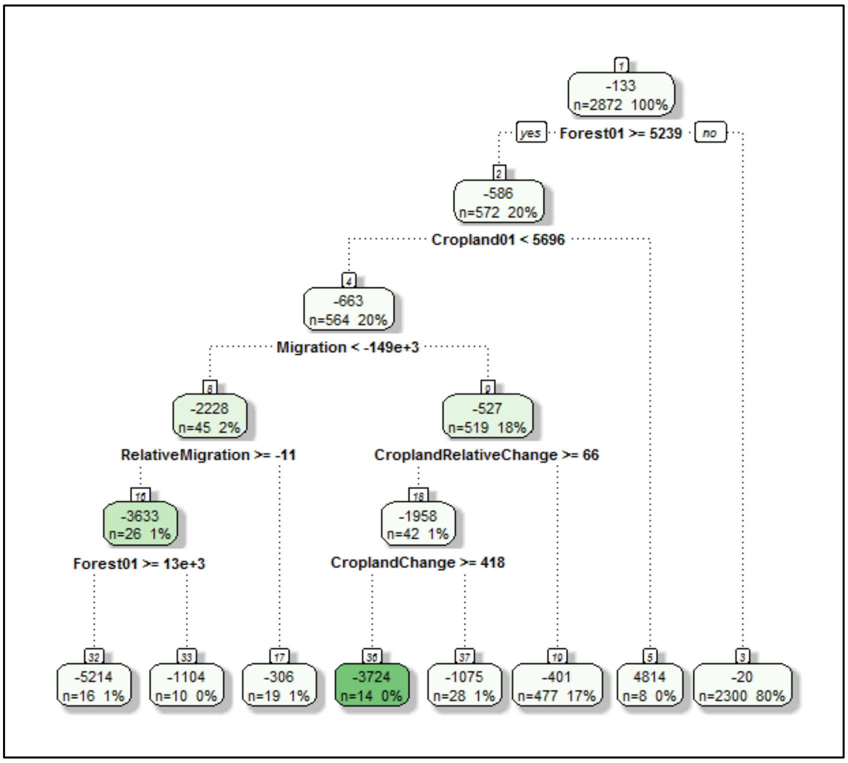


figure 15: rpart regression tree for predicting forest change using land cover and migration as explanatory variables. The binary splits based on predicting variables divide the data into more homogeneous groups. The first number in each leaf (bottom row) is the mean of the group, n = the number of regions in the group, and the percentages are the number of regions in the group relative to total number of regions used for the tree.

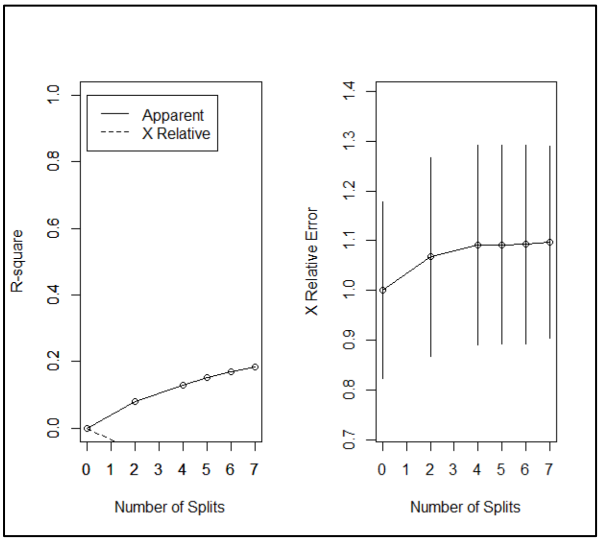


Figure 16: Validation of the regression tree in figure 15 for dependent variable forest change. Left: the of r-square versus the number of splits. Right: cross-validation relative error +/- 1-SE from cross-validation versus the number of splits.

The produced tree for regional relative forest change includes two significant splits (figures 17 to 19) based on explanatory variables forest area in the region in 2001 and on the regional migration. If the forest area in 2001 in the region is above 2.5 km^2, migration is an explanatory variable. Then if the regional migration was below -81,000 persons the relative forest change was + 900 km^2, or when above -81,000 persons 3645 (figure 19). However, these are both high relative forest increase values which cannot support the hypothesis that a positive migration is related to forest cover decline.

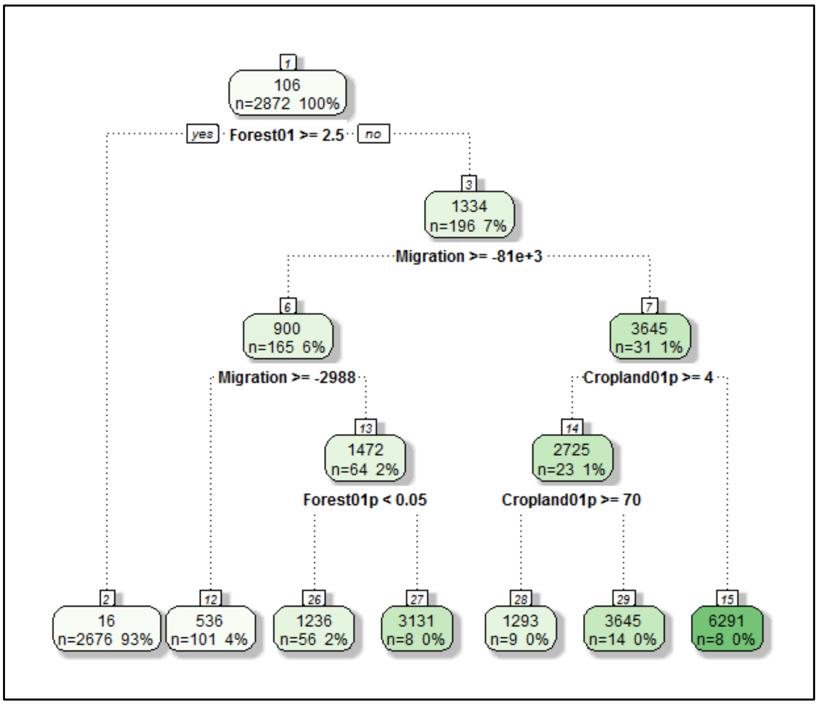


Figure 17: regression tree for predicting relative forest change using migration and land cover as explanatory variables.

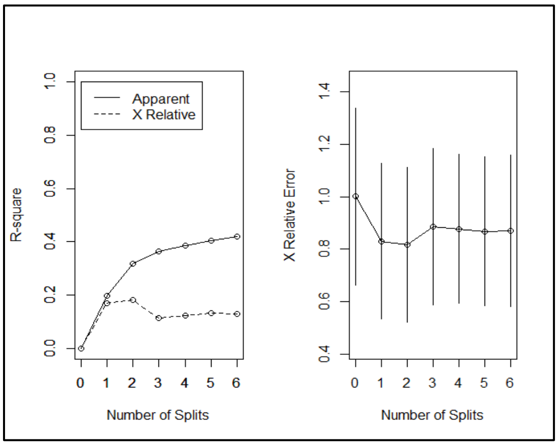


Figure 18: Validation of the regression tree for relative forest change (figure 17).

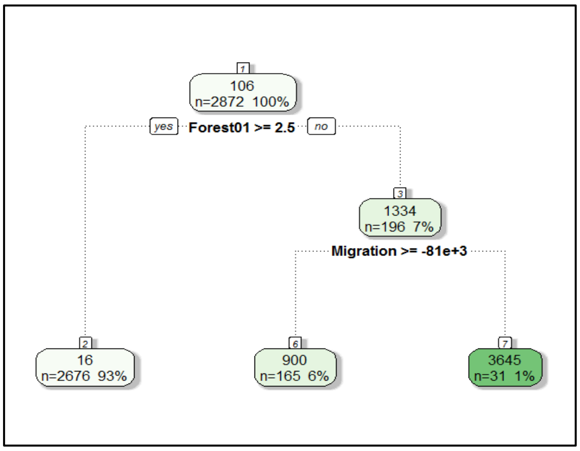


Figure 19: Pruned regression tree for relative forest change.

The regression tree for cropland change (figure 20) does decrease the relative error (figure 21). However this was only caused by the first two splits in cropland 2001 and cropland 2001 percentage. As a result these are only splits left in the pruned tree (figure 22). Only the first two splits turned out to be significant by cross validation. Hence, a cropland area in 2001 smaller than 7,387 km^2 ends up in leaf where mean cropland change is 157 km^2. If cropland in 2001 in the region is higher, depending on the cropland cover percentage in a region in 2001 being smaller or larger than 59%, the fitted tree predicts a cropland change of respectively 100 and 4201 km^2. In summary, the more cropland in the region in 2001, the more cropland change increase between 2001 and 2010.

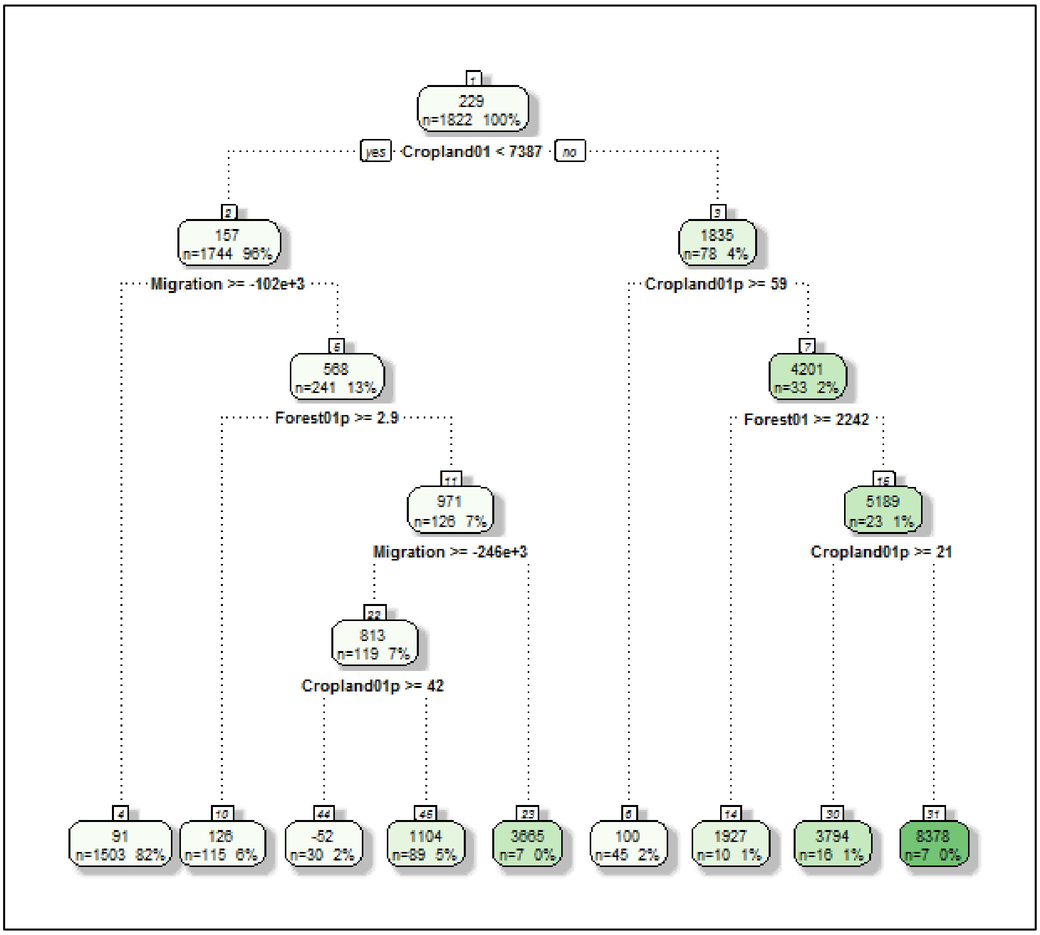


Figure 20: rpart regression tree for predicting cropland change using migration and landcover as explanatory variables.

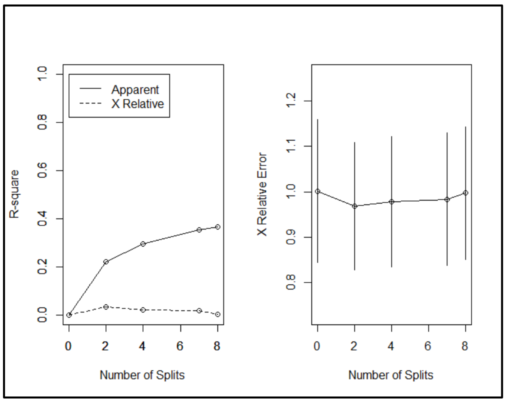


Figure 21: Validation of the cropland change regression tree

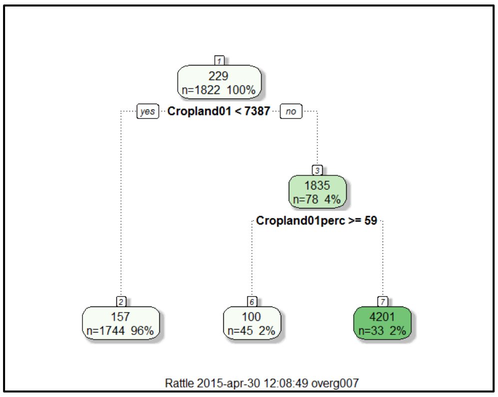


Figure 22: The pruned tree cropland change. based on the validation, the insignificant splits were pruned from the tree resulting in a tree with only the significant splits.

The produced tree for relative cropland change did not result in any significant splits (figure 23). After pruning the tree, only the root is left whereby no explanatory variable is involved.

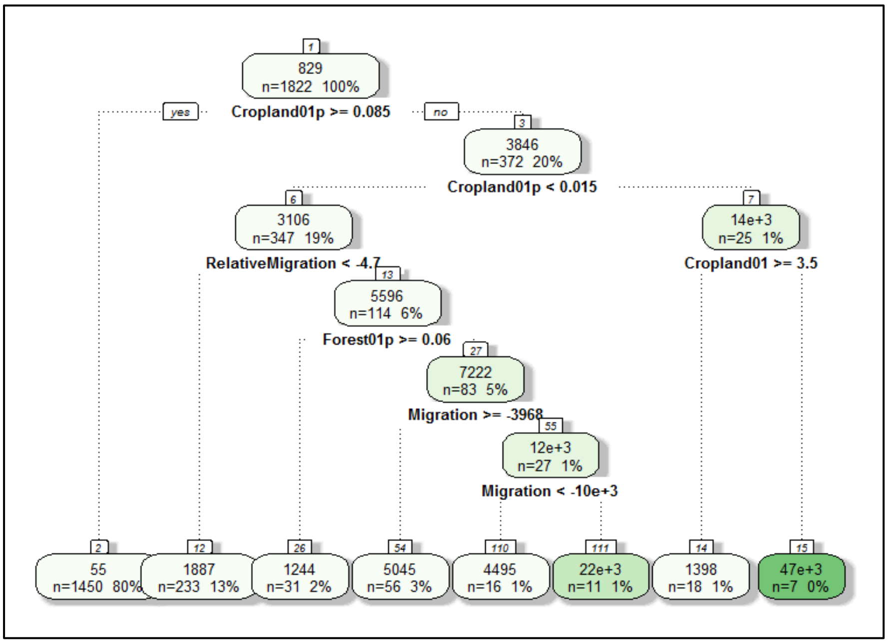


Figure 23: Regression tree for relative cropland change.

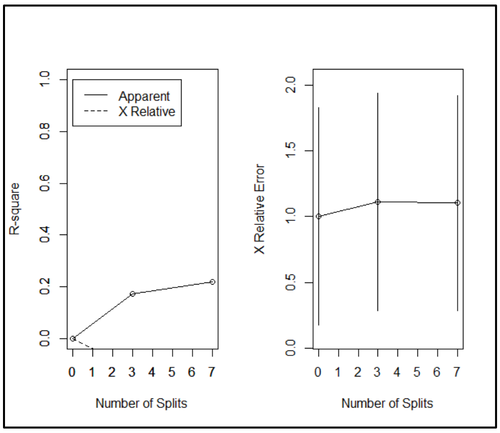


Figure 24: Validation of the regression tree for relative cropland change (figure 23).

3.3 National Studies: Colombia & Indonesia

In this section the data exploration and statistical analysis were repeated on a country level for the two countries Colombia and Indonesia. For Colombia also a finer resolution was used.

3.3.1 Colombia

Data Exploration

Forest Cover

Figure 25 shows that both deforestation and reforestation has occurred in parallel in Colombia between 2001 and 2010. From the 1065 regions, 846 had a forest change value. Along the Amazon the deforestation values are low. Higher deforestation values are located near the central part of the country and a few along the coast.

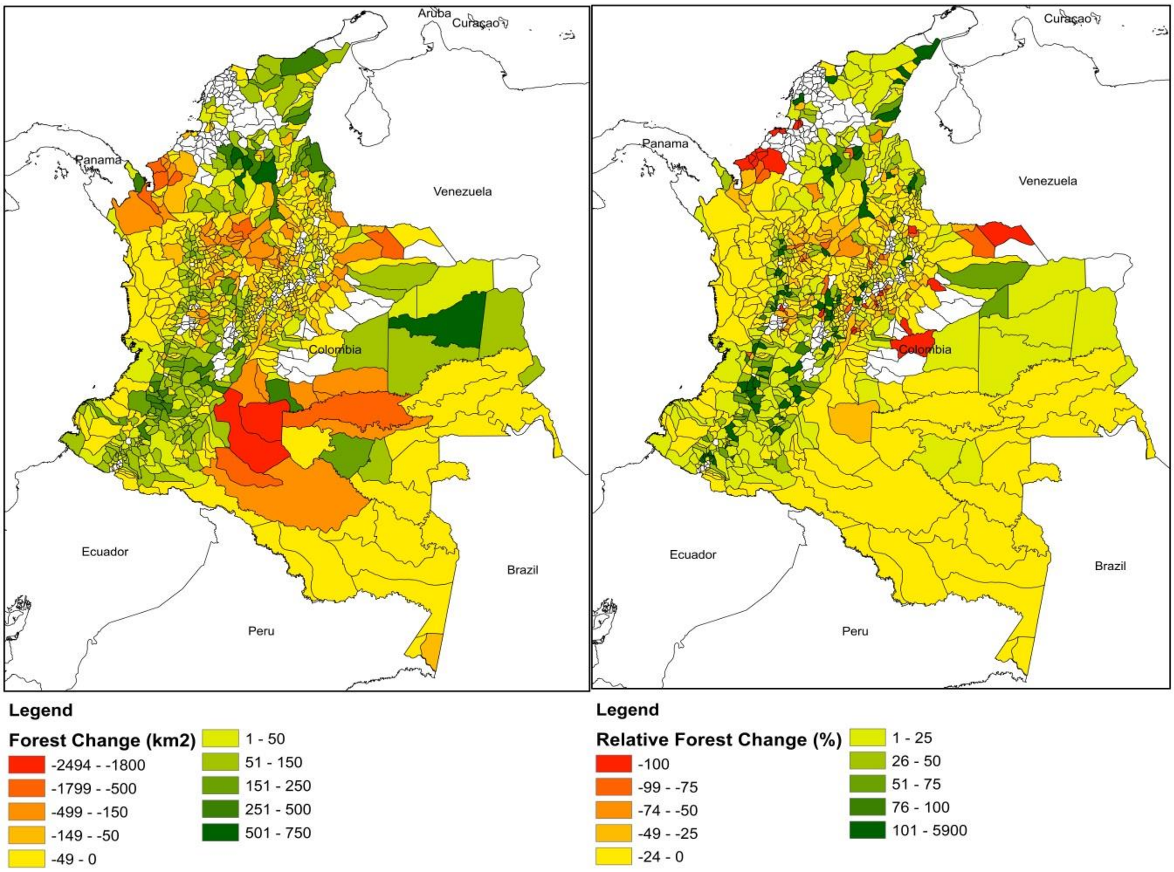


Figure 25: Regional forest cover change maps Colombia. Left: absolute forest change in km^2. And right: relative forest change in km^2.

The statistics in table 11 and the histograms in figure 26 show a distribution around zero for the regional forest change values in Colombia. A few high outliers in relative forest change exist.

Table 11: Summary statistics of regional forest change in Colombia in km^2 and in percentage.

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Forest Change (km^2)	9	164	-2,490	0	0	52	731
Relative Forest Change (%)	65	403	-100	0	0	10	5,900

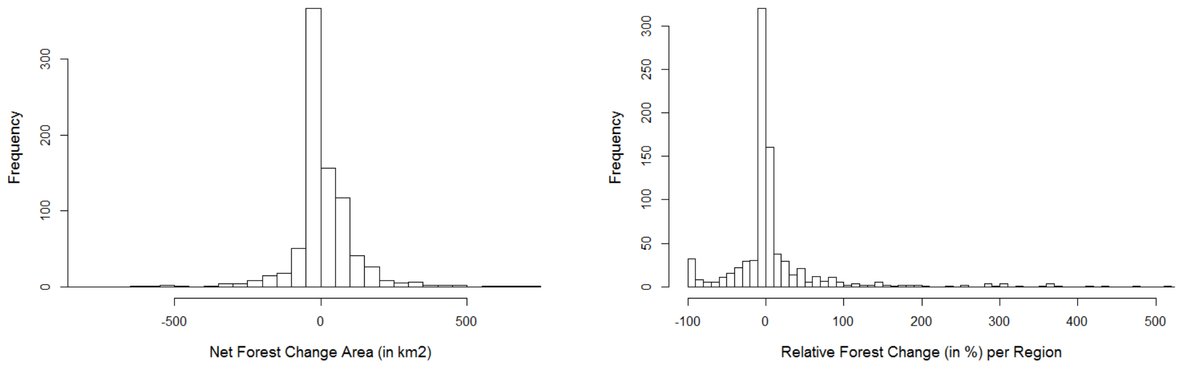


Figure 26: Histogram Forest change Colombia. Left: forest change. 5 from the 846 regions with forest change have a value lower than -500 km^2 (not shown). Right: relative forest change. 25 regions have a value higher than 500% (not shown).

Cropland Cover

Only 35 out of the 1065 regions in Colombia experienced a change in cropland cover. Both an increase and a decrease of cropland per region took place (figure 27). Table 12 shows that the values were mainly zeros, no change.

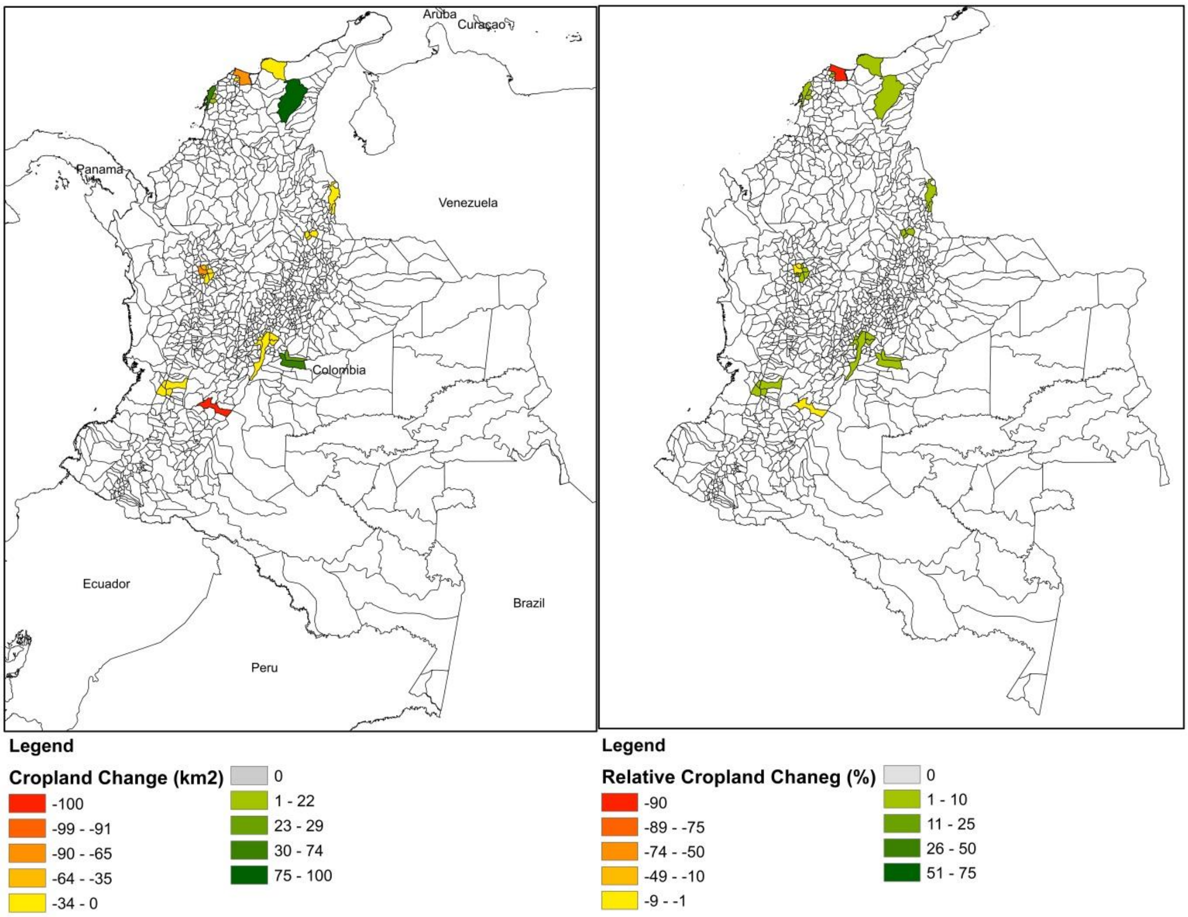


Figure 27: Regional cropland cover change maps Colombia. Left: absolute cropland change in km^2. And right: relative cropland change in km^2.

Table 12: Summary statistics of regional cropland change in Colombia in km^2 and in percentage.

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Cropland Change (km^2)	0	37	-100	0	0	0	100
Relative Cropland Change (%)	83	567	-100	0	0	0	3,180

Migration

In figure 28, the absolute and relative migration maps, it can be seen that in the west of the country there was mainly a small negative regional net migration, while in the east there was more variation in both directions: positive and negative values. The relative migration values had a different spatial distribution with less extreme values.

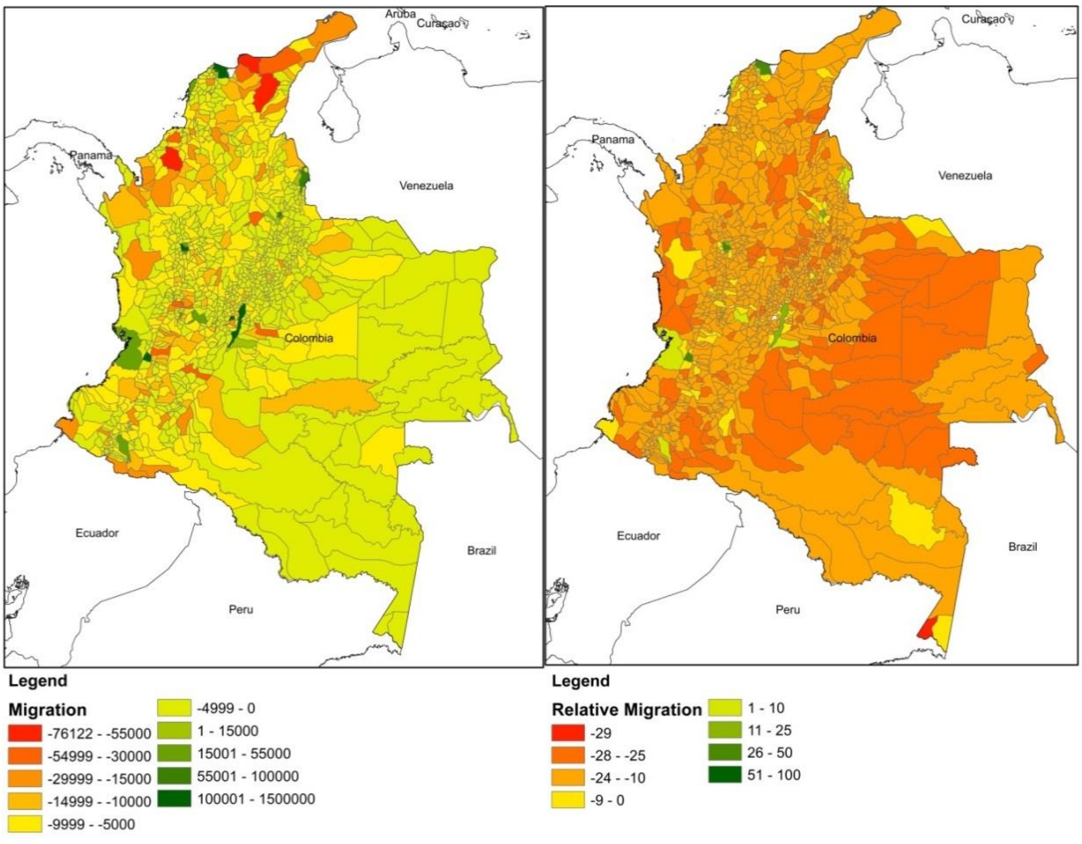


Figure 28: Migration change maps for Colombia: left: migration right: relative migration.

In table 13 it can be seen that the migration values in Colombia were similar to the global migration values; more than 75% had a negative regional net migration. This is also visible in the two histograms in figure 29 for absolute and relative migration.

Table 13: Summary statistics of regional migration in Colombia in persons and in percentage.

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Migration	-227	63,987	-76,100	-5,950	-3,420	-1,910	1,450,000
Relative Migration	-19	18	-29	-24	-22	-18	524

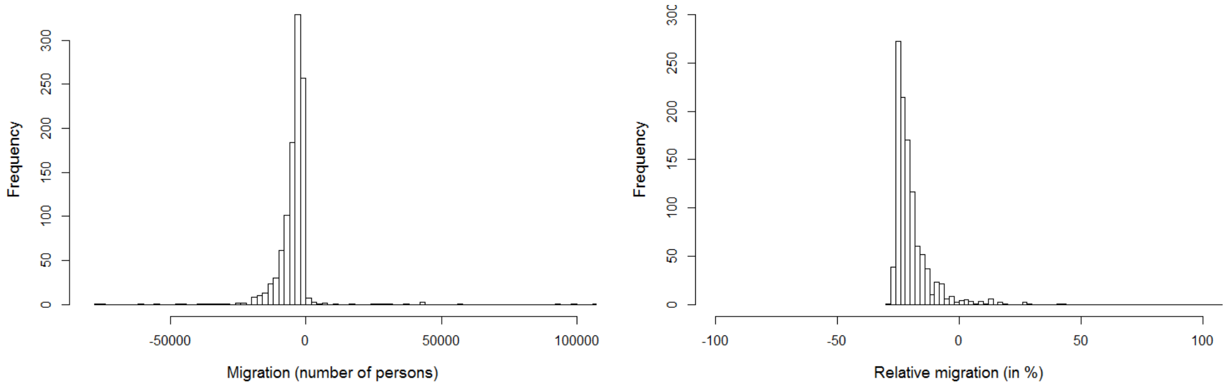


Figure 29: Histogram migration Colombia. left: migration. six regions had values above 100,000 persons (not shown) and right: relative migration.

Linear correlation

Figure 30 shows the scatter plots for all the land cover variables versus the migration variables for Colombia. Visually, no relation can be derived. In addition, the correlation coefficients are shown in figure 31. Moderate negative correlations were found between the cropland change and migration variables with the highest correlation between the relative variables [r = -0.42]. This is in contrast with the hypothesis that a positive regional migration is related to cropland cover increase in the region.

A weak negative relation was found between the variables forest change and cropland change [r = -0.26]. This indicates that an increase in forest cover was related to a cropland cover increase and a decrease in forest to an increase in cropland cover.

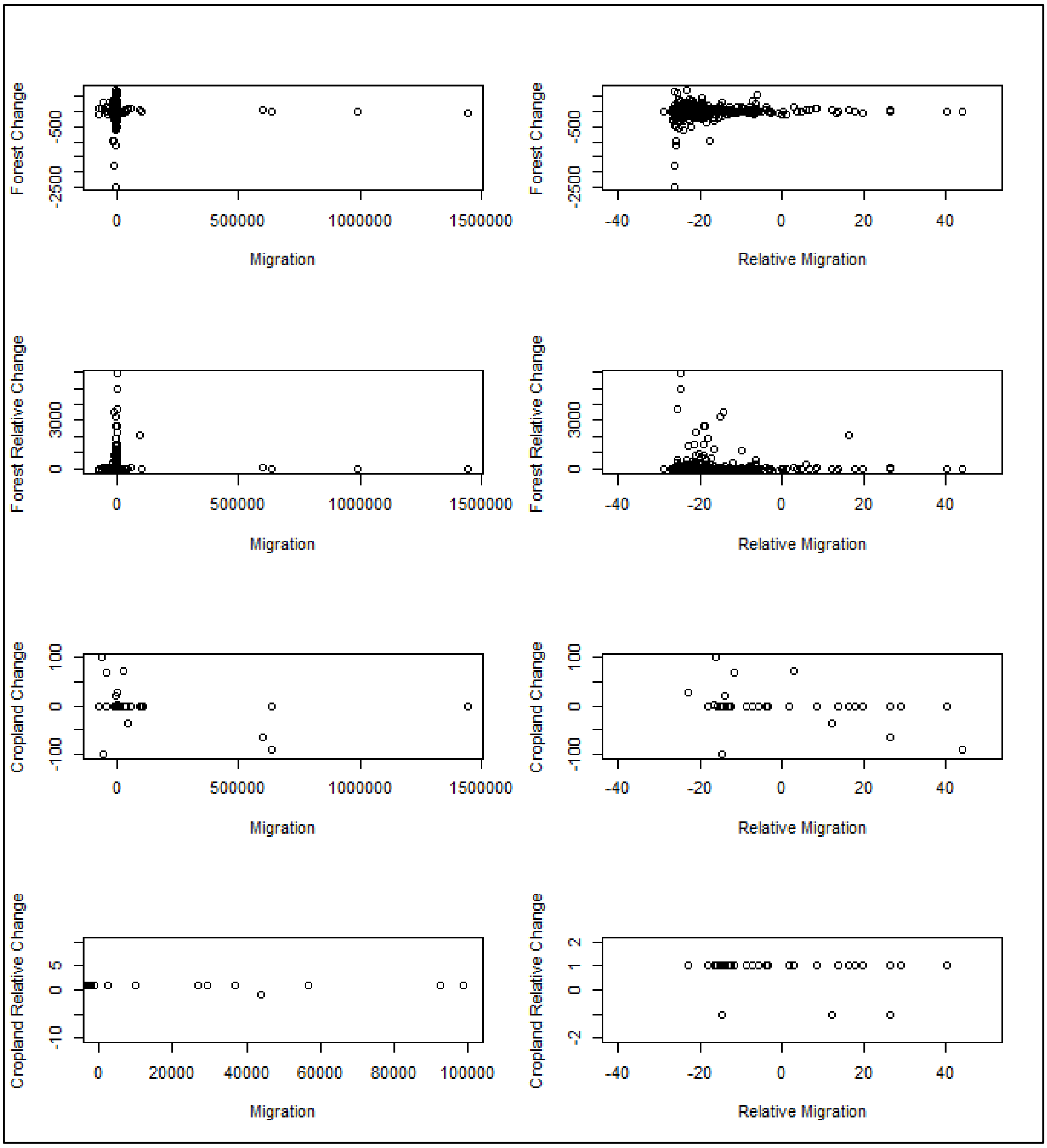


Figure 30: Scatter plots of the regional land cover change variables to the migration variables in Colombia 2000-2010

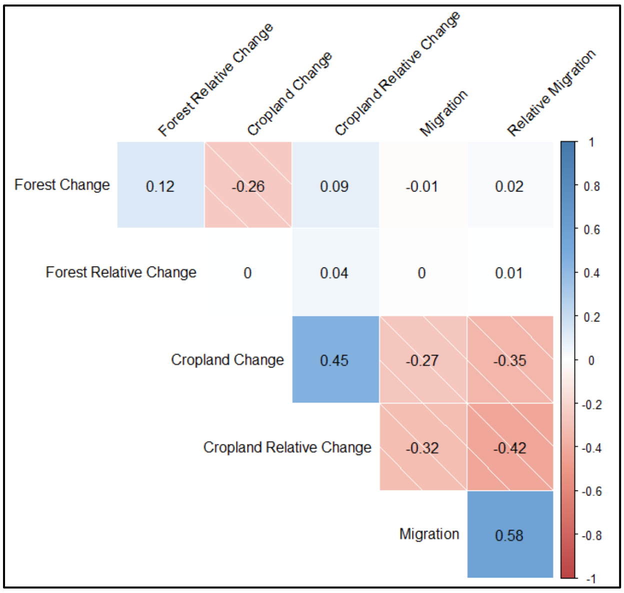


Figure 31: Pearson's correlation coefficients for all the combinations of land cover change and migration variables.

Regression Trees

Figure 32 shows the first result of a regression three for predicting forest change in Colombia. In contrast to the results for the tropics, migration is now included as a split. The validation of the three in figure 33 shows that the first two and the fourth split decrease the relative error slightly. Consequently, no splits where pruned as they were all found to be important enough. The first split is based on forest area in 2001. According to the three, if forest 01 is more than 8636 km^2, 24 cases, i.e. 3% of the data enter the rightmost branch which is subsequently split based on migration. If migration is below -4276 (net outmigration), then forest change is on average -867 km^2. This contradicts the hypothesis that positive migration is related to negative forest change. Consequently, for Colombia hypothesis 1 can be rejected.

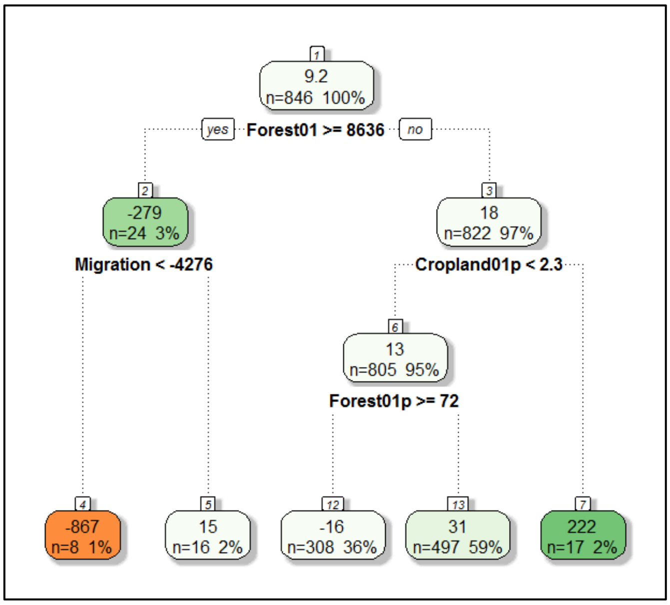


Figure 32: Regression tree for predicting forest change with land cover and migration explanatory variables in Colombia.

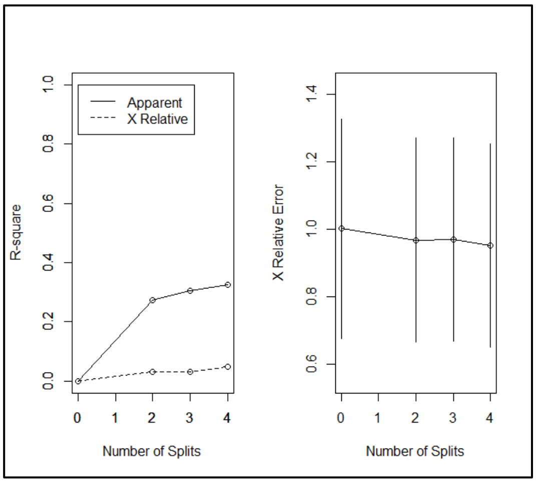


Figure 33: Validation of the forest regression tree in figure 35. Left: the improvement of r-square with each additional split. Right: cross-validation relative error.

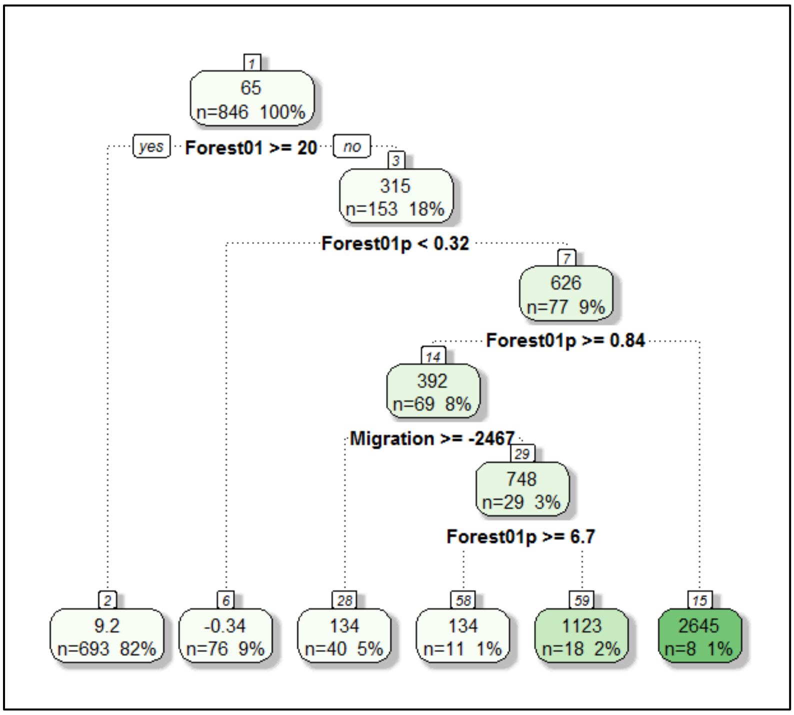


Figure 34: Regression tree for predicting **relative** forest change with land cover and migration explanatory variables in Colombia

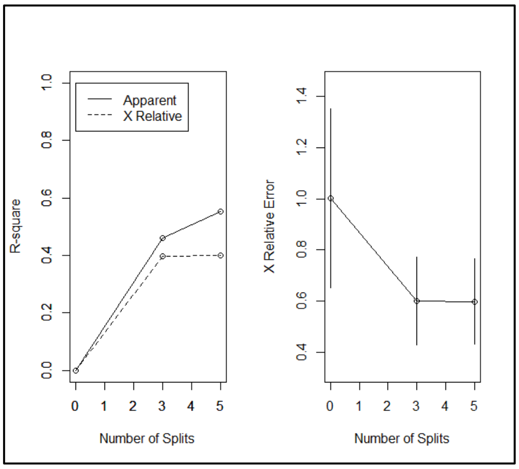


Figure 35: Validation for relative forest change regression tree in Colombia

The regression tree for relative forest change is presented in figure 34 and the validation in figure 35. Also here, the pruned three did not lose any splits, implying that al splits were judged to be important. Again the forest area in 2001 was used in the first split. The migration split here comes third and is splits 8% of the data into two groups. If migration in a region was below -2467 persons, the resulting subset had a mean relative forest change of 134%. If migration was above 2467 persons, the subset was split another time by the forest percentage in the region in 2001. If this was below 6,7% the resulting mean value of the subgroup was also 134%. If the forest percentage in the region in 2001 was above 6,7% the mean value of the resulting subset was 1123% forest increase, which is much higher. In summary, if a net outmigration was large and the forest percentage in 2001 was below 6,7 % the forest change was more positive. Although this does not provide evidence for the hypothesis that in migration is related to deforestation, it is for opposite of it: outmigration is related to reforestation.

In figure 36 and 37 can be seen that the prediction tree for cropland change only includes the significant splits using explanatory variable urban in 2001. If the area of urban in 2001 in a region was below 1,5 km^2 the mean cropland increase was 43 km^2. When the urban area in a region in 2001 was

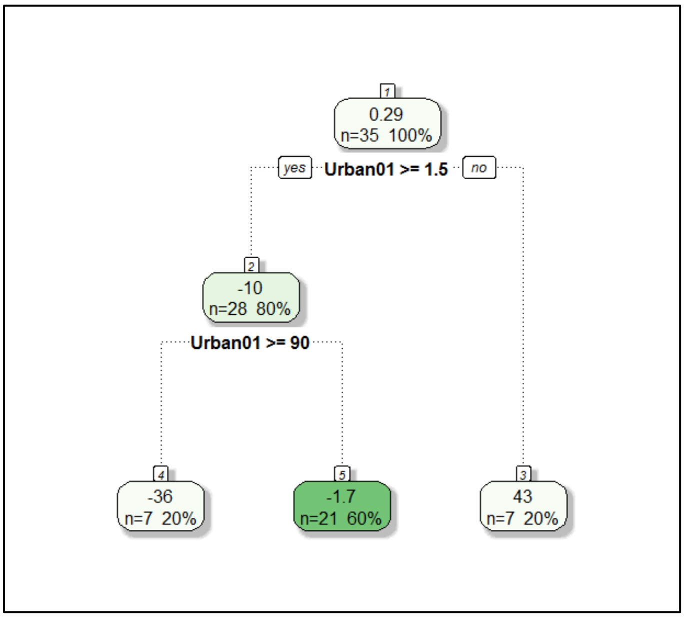


Figure 36: Regression tree for relative cropland change in Colombia

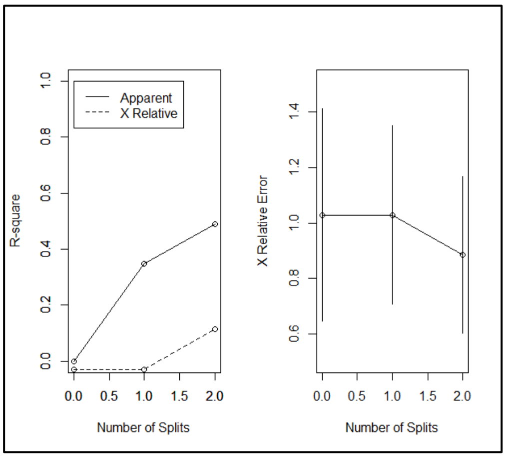


Figure 37: Validation of the cropland change regression tree in Colombia

above 1.5 km^2, the cropland change was negative, depending on whether the urban area was above 90 km^2 more negative. This suggests that an increase in urban land area rivalled croplands, resulting in a decrease of cropland cover.

The relative cropland change is shows only one split which is not significant as it increases the relative error (figure 38 & 39).

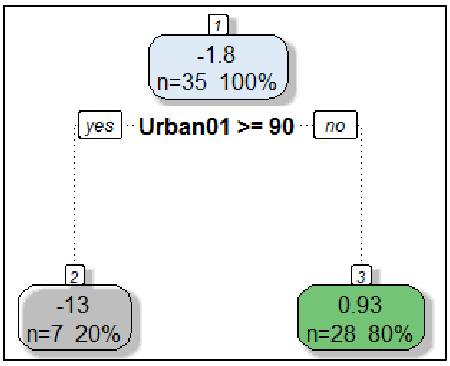


Figure 38: Regression tree for cropland change in Colombia

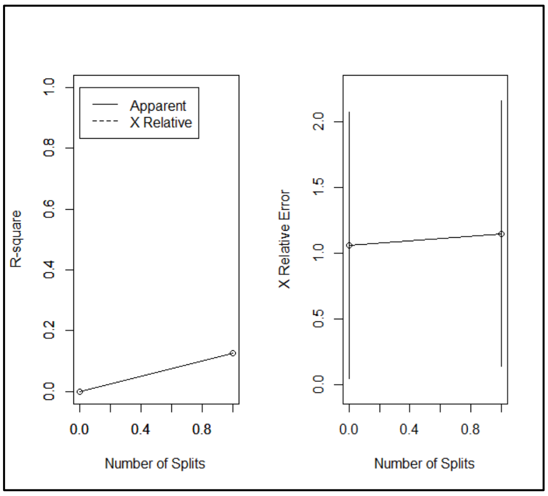


Figure 39: Validation of regression tree relative cropland change in Colombia

3.3.2 Indonesia

Data exploration

Forest Cover

In the global analysis, the regional level that was used for Indonesia was level 2, resulting in 441 regions. The same regions were used for this national analysis. Table 14 presents the area sizes of the regions. The high standard deviation indicates that the variation of region sizes is considerably large.

Table 14: Descriptive statistics of the area sizes of the regions in Indonesia in km^2.

Level	Number of Regions	Mean	Std.dev	Quantile Areas				
				Min	25%	Median	75%	Max
2	441	5,036	7,270	12	984	2,490	5,570	52,200

In 388 of the 441 regions forest cover change occurred between 2001 and 2010. The spatial distribution of the regional forest change in Indonesia is shown in in figure 40. Spatially distinct patterns of both absolute forest cover gain and loss are present. While on Sumatra, Java and Papua deforestation and reforestation between 2001 and 2010 was significantly strong, the other islands

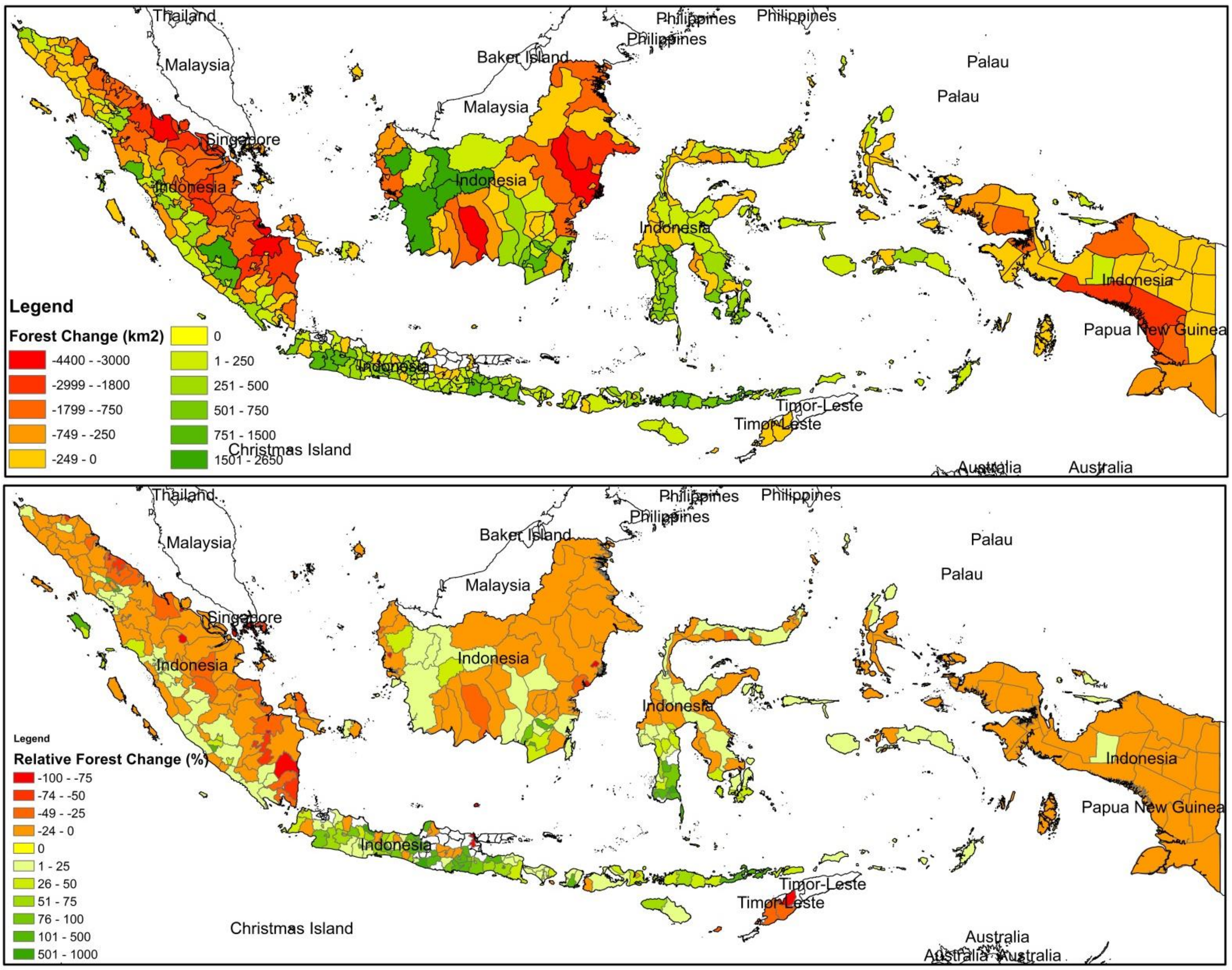


Figure 40: Regional forest change 2001- 2010 in Indonesia. Upper map: the absolute net change in km^2. Map below: forest change relative to the forest area in the region in 2001.

with smaller regions show mainly a moderate increase in forest cover. However, the map below in figure 40 shows that relative forest cover changes are less extreme for the larger regions. This shows that the absolute values were influenced by the size of the regions.

The descriptive statistics of regional forest cover change in Indonesia between 2001 and 2010 (table 15) and the histograms in figure 41 reveal that half of the regions had a values between -69 and +136 km^2, while the range is much wider from -4,400 to +2640. Also, the relative forest cover change variable has a majority of the data close to zero, indicating both the absolute and relative forest change was small in most regions.

Table 15: Descriptive statistics regional forest change Indonesia in km^2 and in percentage.

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Forest Change (km^2)	-61	690	-4,400	-69	0	136	2,640
Relative Forest Change (%)	34	166	-100	-2	0	14	1,720

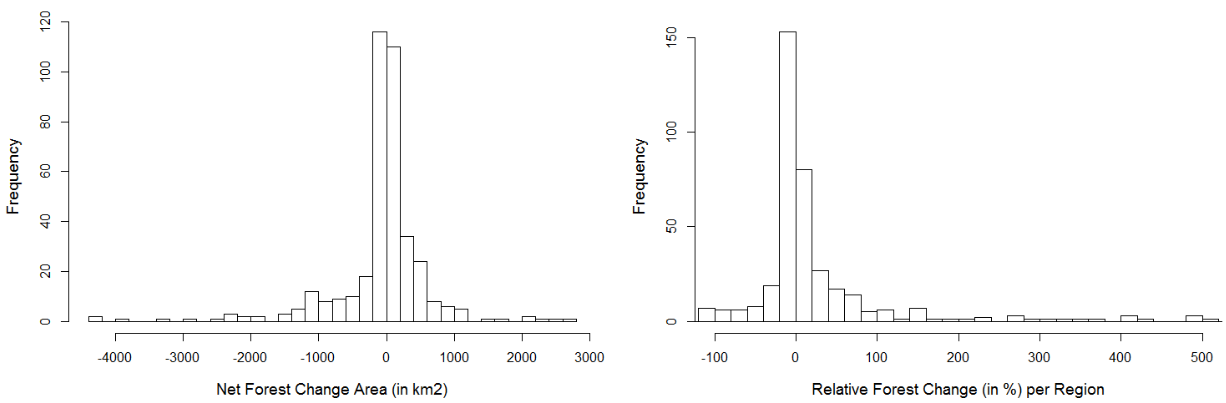
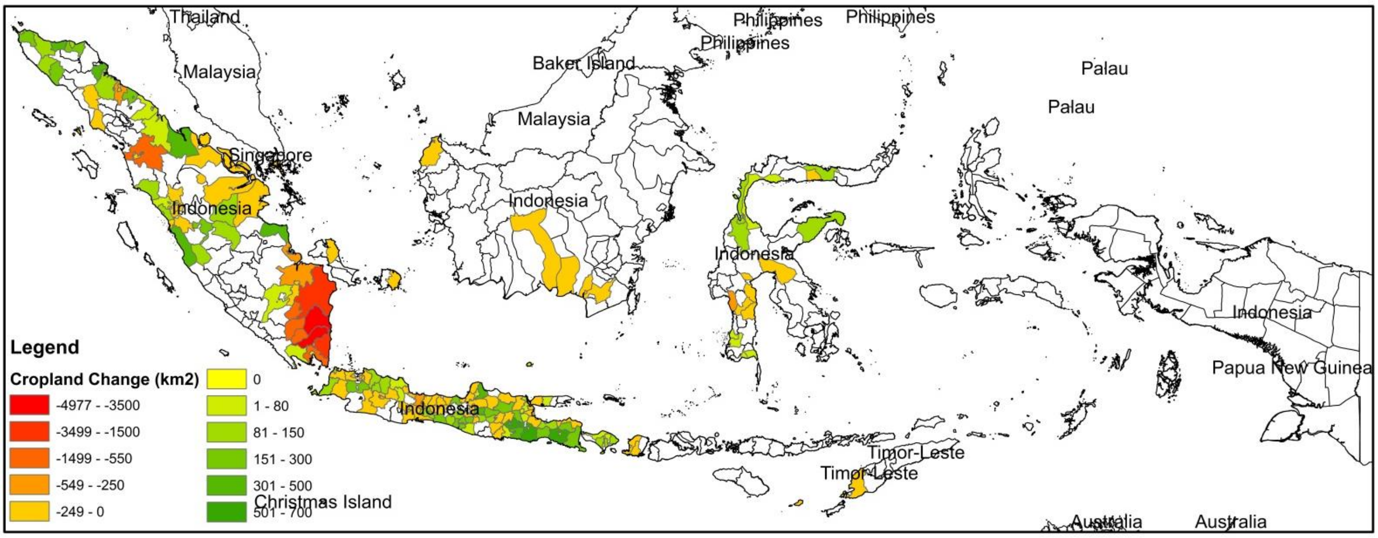


Figure 41: Histograms regional Forest change Indonesia. left: forest change. right: relative forest change. 7 regions had values above 500% (not shown).

The relative forest cover change (figure 41, right histogram) is more right-skewed as values cannot be smaller than -100% which is total deforestation, but can be much higher than 100%, especially when the forest area in a region in 2001 is low. As a result the mean of the relative forest cover is positive (+34%) while the average absolute forest change was negative with -61 km^2.

Cropland Cover

From the 441 regions in Indonesia, 184 experienced a change in cropland areas between 2001 and 2010. Figure 42 shows that both cropland increase and decrease occurred. The relative cropland change was more extreme than the absolute cropland change.



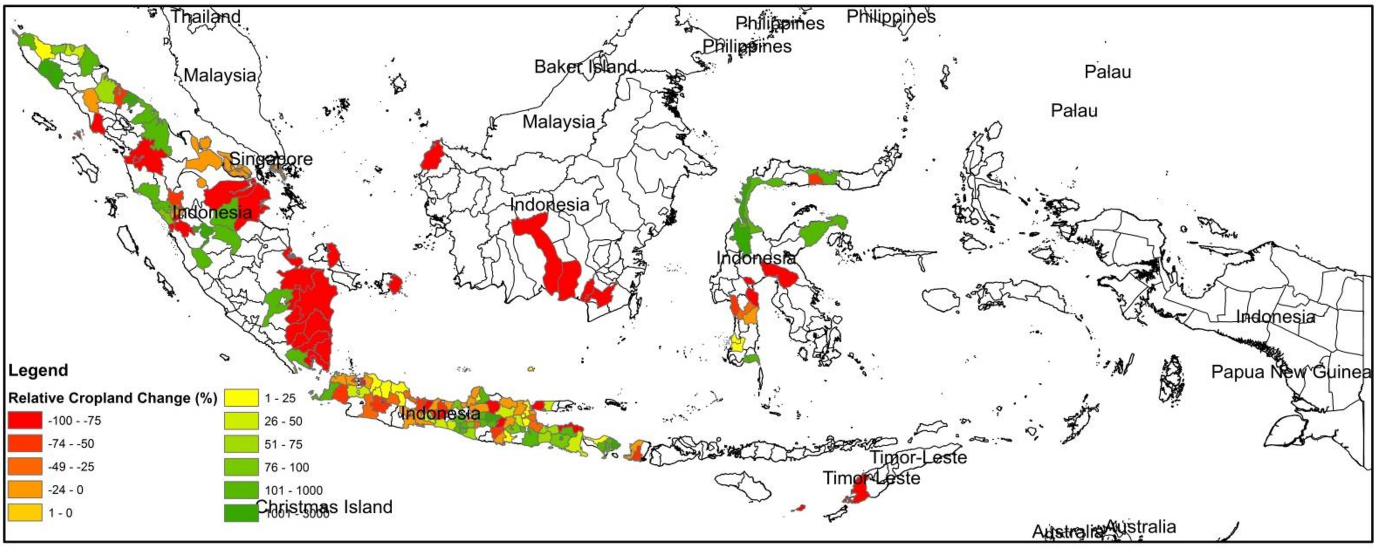


Figure 42: regional cropland change 2001- 2010 in Indonesia: upper map: the absolute net change in km^2. Map below: cropland change relative to the cropland area in the region in 2001.

The descriptive statistics of regional cropland cover change in Indonesia (table 16) shows a median of zero and a mean of -52, implying that on average the cropland area declined.

Table 16: descriptive statistics of regional cropland change Indonesia in km^2 and in percentage

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Cropland Change (km^2)	-52	538	-4,980	-85	0	100	704
Relative Cropland Change (%)	5	140	-100	-72	0	6	1,200

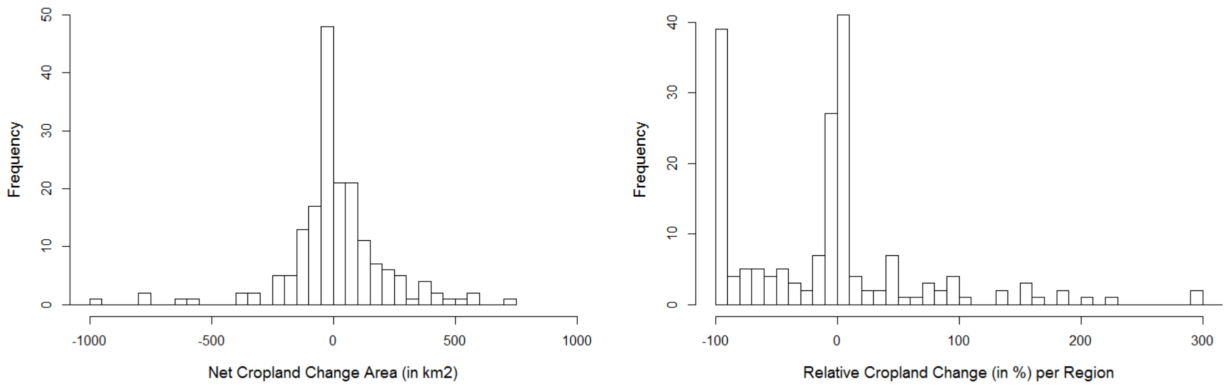


Figure 43: Histograms of regional cropland change in Indonesia. Left: absolute change and Right: relative change.

The histogram for absolute cropland change values (figure 43) has a peak slightly to the left of zero. The relative values show an edge peak for -100%.

Migration

The migration maps (figure 44) show that at the aggregated level most regions are characterized by outmigration. The map showing regional relative migration shows mainly values between -22 and -15 persons (red). There were some hotspots have high immigration. It seems that there was a country wide outmigration with most people having the same destinations in mind. The few regions with net in

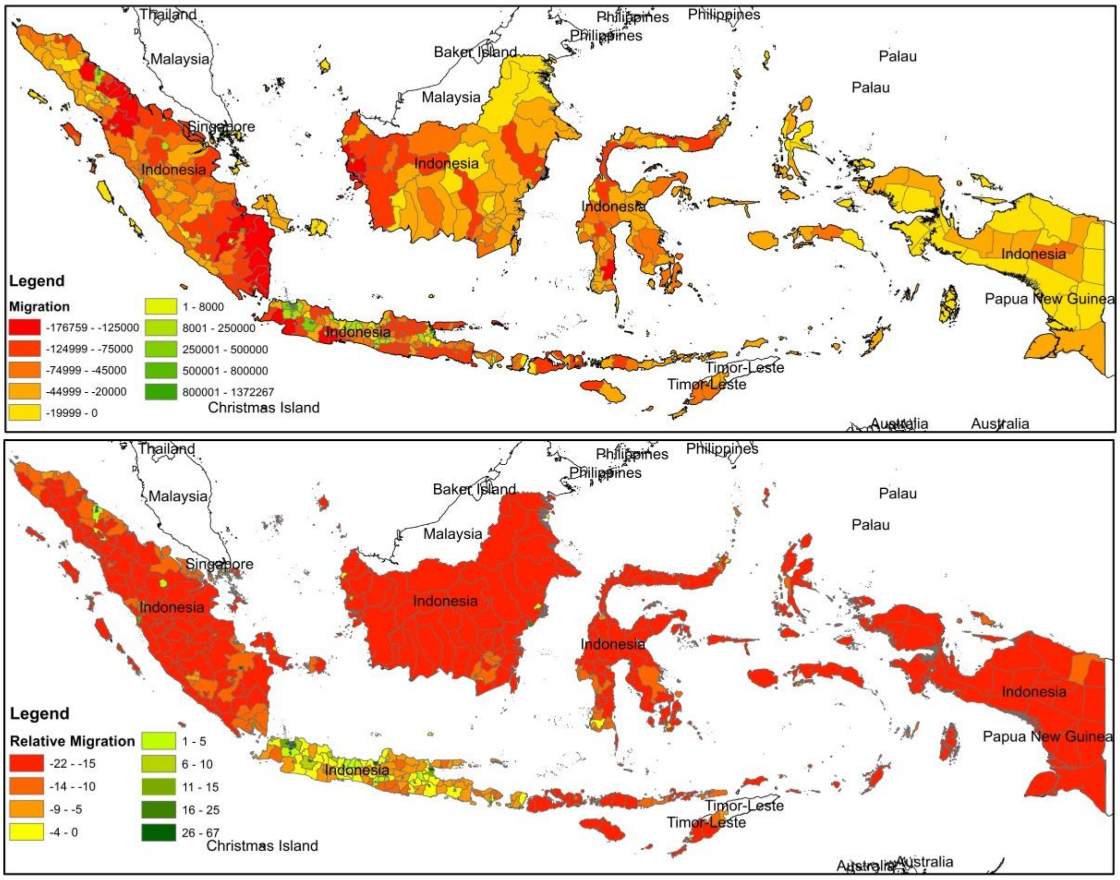


Figure 44: Regional migration 2001 – 2010 in Indonesia. Upper map: migration in persons. Map below: migration relative to the population in the region in 2001.

migration are location in Java, in which Jakarta in the west and Surabaya in the east had the most extreme migration values.

The descriptive statistics in table 17 also show that more than 75% of the regions had a negative migration value. The mean and median for both variables are negative.

Table 17: Descriptive statistics of regional migration in Indonesia in persons and in percentage

	Mean	Std.dev	Quantile Areas				
			Min	25%	Median	75%	Max
Migration (persons)	-3,500	165,711	-177,000	-56,900	-29,400	-10,600	1,370,000
Relative Migration (%)	-9	165,711	-22	-18	-14	-4	66

The histograms of regional migration in Indonesia (figure 45) also show that the frequency of regions with a negative value is higher than the frequency of regions with a positive value. The positive regional migration values include more extreme values; the highest in migration of 1,370,000 persons in Surabaya is much larger than the largest out migration of -177,000 persons.

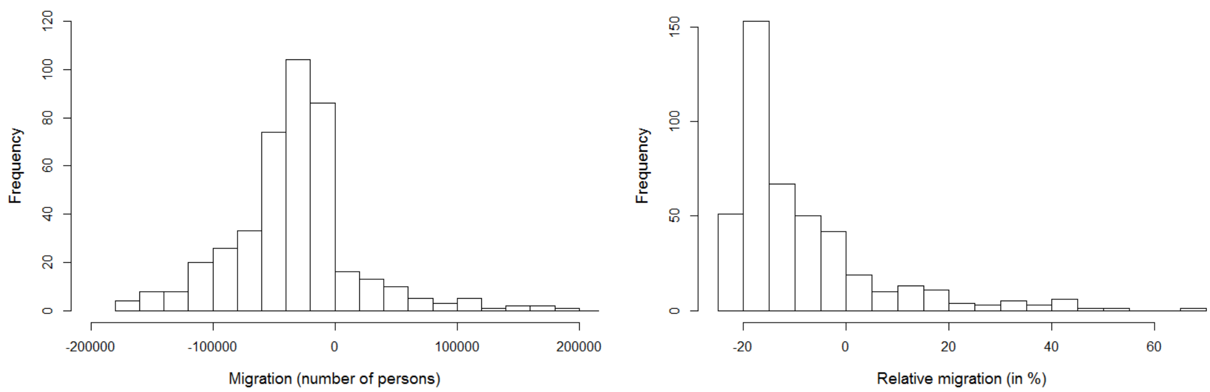


Figure 45: Histograms regional migration Indonesia. Left: absolute migration. Right: relative migration.

Linear Correlation

In figure 46 scatter plots are presented for the regional land cover change variables versus the migration variables. Some weak negative linear relation can be detected visually in the second row for relative forest change versus migration and in the bottom plots for relative cropland cover change is versus migration. However, the Pearson's correlation coefficients between these two variables (figure 47) are all negligible.

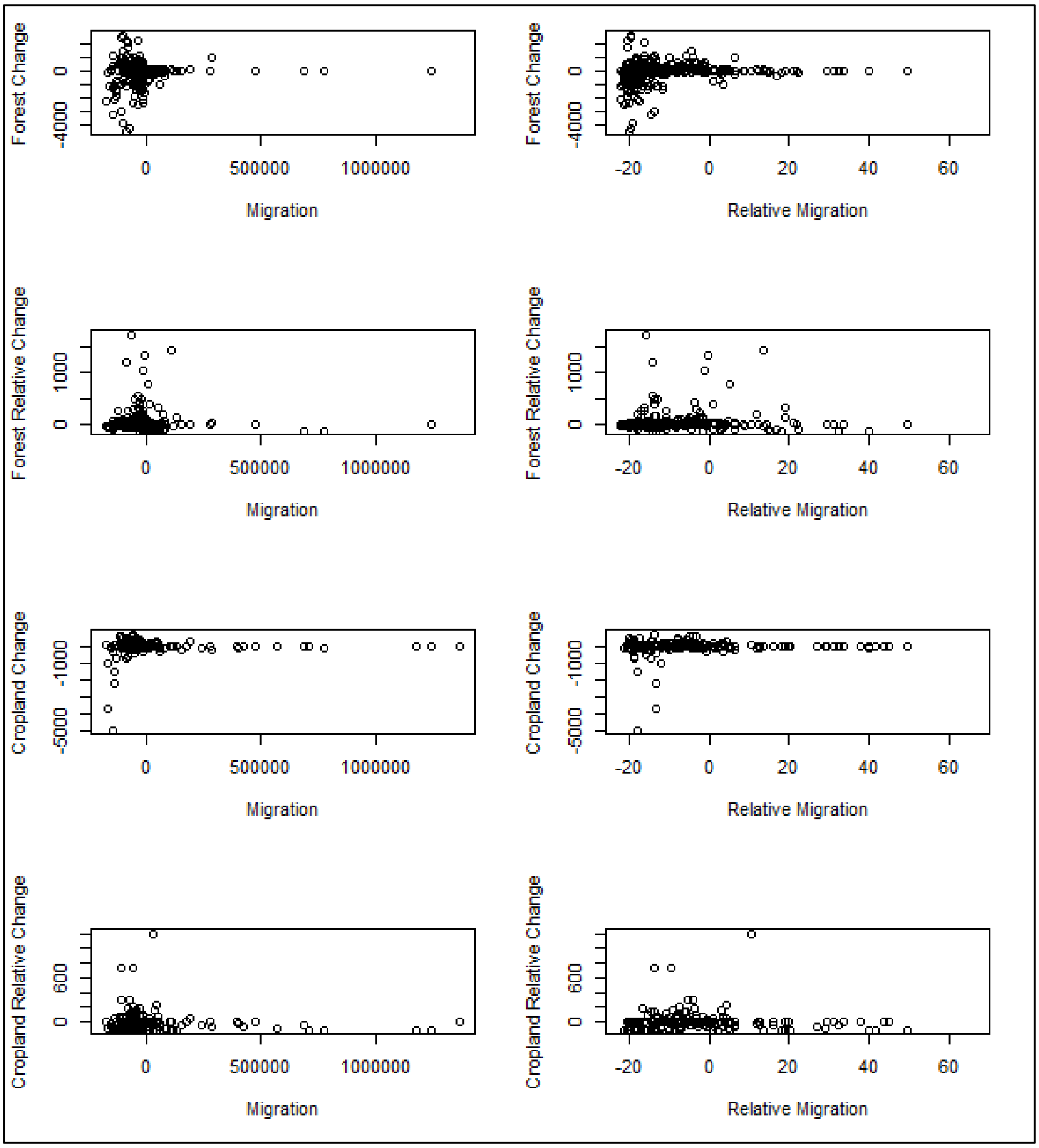


Figure 46: Scatter plots for all land cover change variables versus all migration variables for Indonesia.

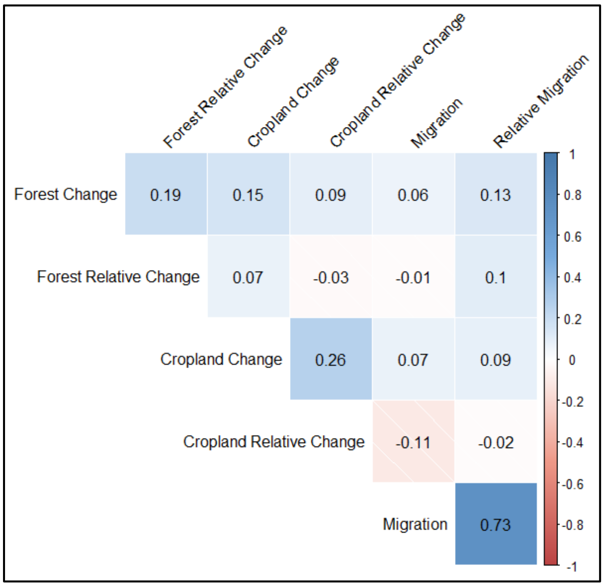


Figure 47: Correlation coefficients for all possible combinations between forest cover change and migration variables.

Regression Trees

The first split in the tree for explaining the variance in the regional forest cover change in Indonesia is based on the variable forest cover in the region in 2001. A second split used migration for variance reduction in regional forest change. Also, further in the tree the regional migration variables are used to split the data into more homogenous groups. However, as can be seen in the validation of the tree (figure 47) only the first split decreased the relative error and thus is not significant. As a result, pruning the tree left tree with only one split (figure 50). If the forest area in a region was larger than 8822 km^2 the mean of the subset of the forest change values was -648 km^2. When it was above 8822 km^2 the mean forest change in the resulting subgroup was +50 km^2.

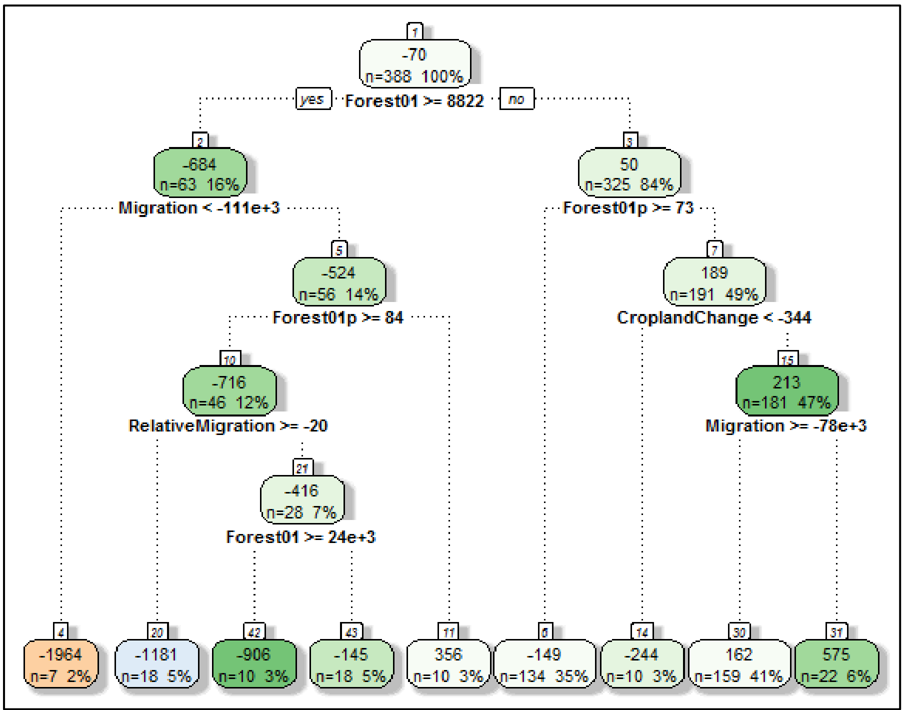


Figure 48: Regression tree for explaining forest change(in km^2) Indonesia using migration and land cover as explanatory variables.

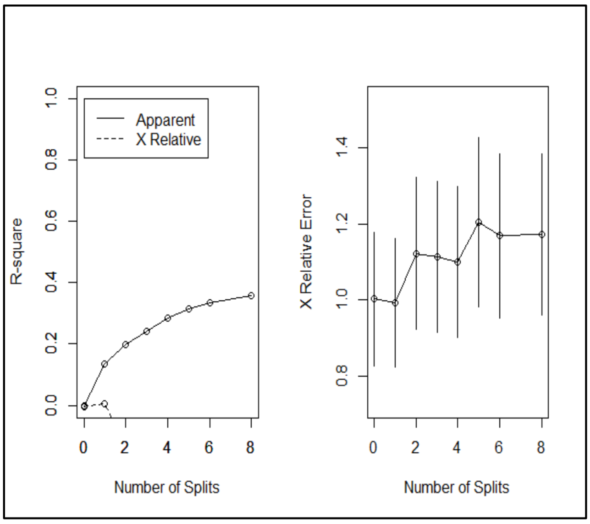


Figure 49: Validation of the tree explaining forest change in Indonesia (figure 48)

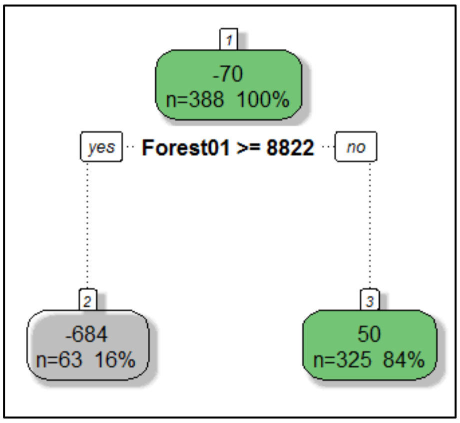


Figure 50: Pruned tree explaining forest change in Indonesia

Figure 51 presents the produced regression tree for regional relative forest change in Indonesia. The first split was based on the forest cover area in a region in 2001. This split contributed the most to the R-square improvement and a decrease in the relative error (figure 49). As a result, pruning the tree led to a tree using only this to explain relative forest change (figure 50). If forest cover in 2001 was larger than 82 km^2 the forest change was predicted to be lower (9,9 km^2 on average per administrative unit) and when forest cover in 2001 was smaller, then the predicted forest cover change was 272 km^2. This implies that a larger forest area in 2001 in the regions was related to lower forest cover change increase the decade 2001 – 2010.

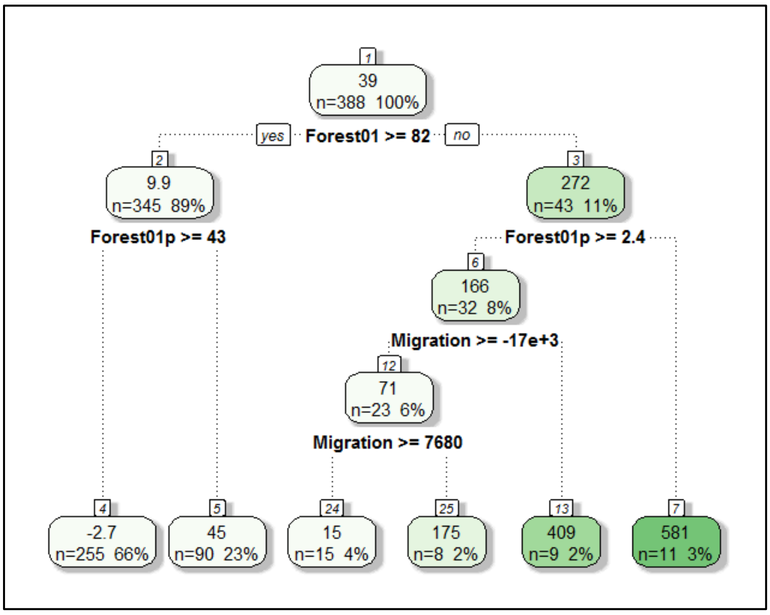


Figure 51: Tree explaining regional relative forest cover change (in km^2) in Indonesia.

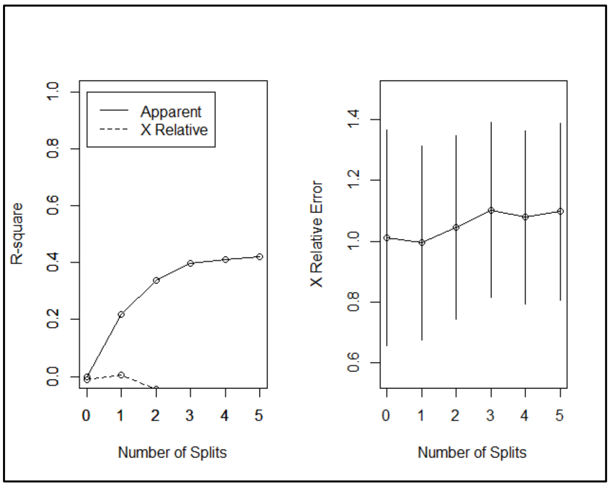


Figure 52: Validation of the regression tree for relative forest change in Indonesia.

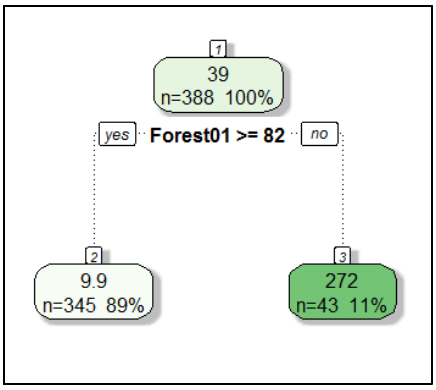


Figure 53: Pruned tree explaining relative regional forest cover change

The produced tree to explain cropland cover change in Indonesia (figure 54) used the variable migration for the first split. This split was found to be significant as it caused a decrease in the relative error (figure 55). Also the second and last split based on the variable cropland in 2001 was found to be significant. Consequently, both splits were retained in the pruned tree. The first split is based on a value of regional migration of -140,000 persons. If the regional migration was below this value, the cropland change was predicted to be -1534 km^2 and when it was above the value it was -154 km^2 or 43 km^2 depending on whether the cropland area in the region was below above or below 562 km^2 respectively. Unfortunately, this could not support the hypothesis that a positive regional migration was related to an increase of cropland cover in Indonesia as the split has such a large negative number. A large net regional outmigration was found to be related to a cropland cover decrease.

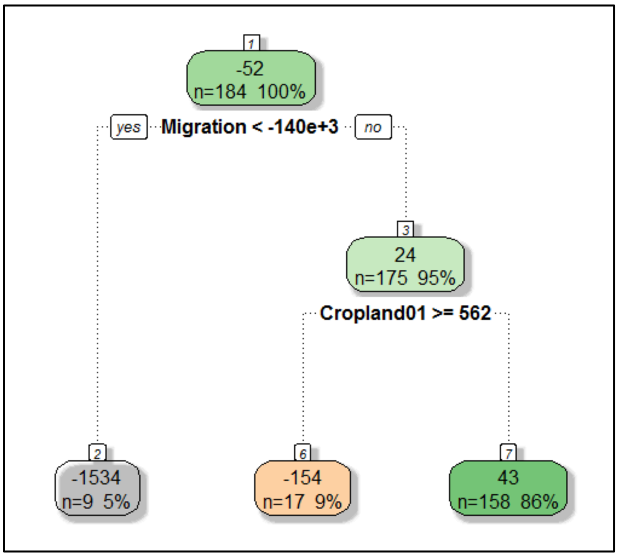


Figure 54: Regression tree explaining regional cropland change (in km^2) in Indonesia.

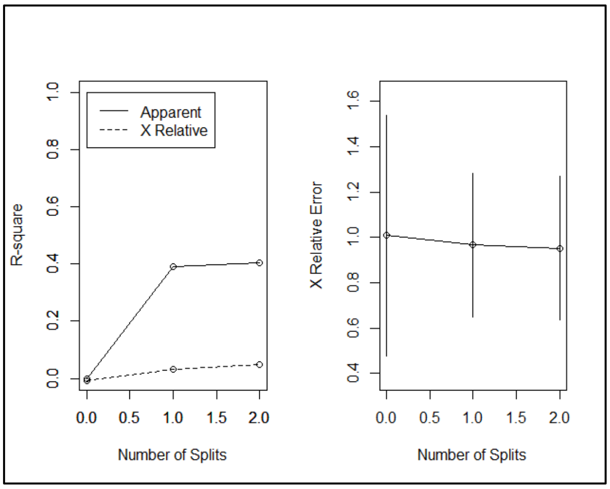


Figure 55: Validation of the tree for region cropland cover change in Indonesia.

The last tree produced for land cover change in Indonesia was for the relative cropland cover change (figure 56). Although the migration variables were used as explanatory variables, the validation (figure 55) shows that the splits in the tree did not decrease de relative error, which indicated that they were insignificant.

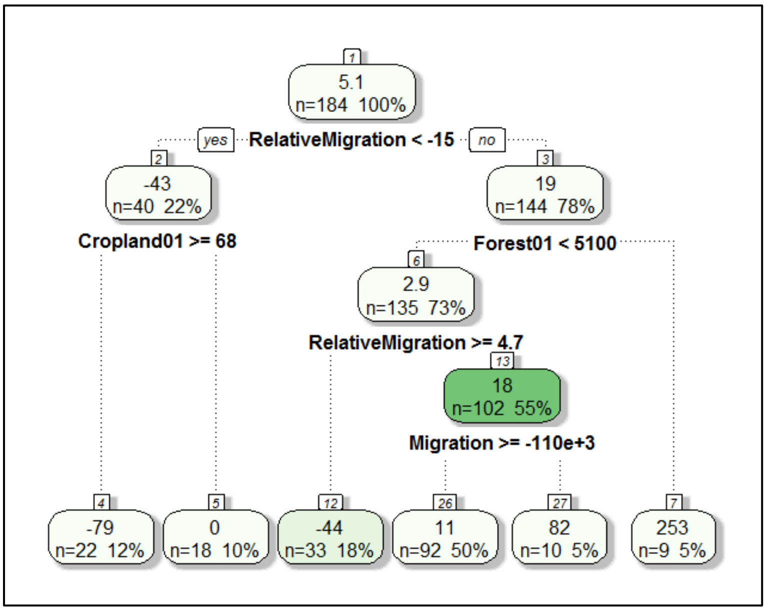


Figure 56: Regression tree explaining regional relative cropland cover change (in km^2) Indonesia.

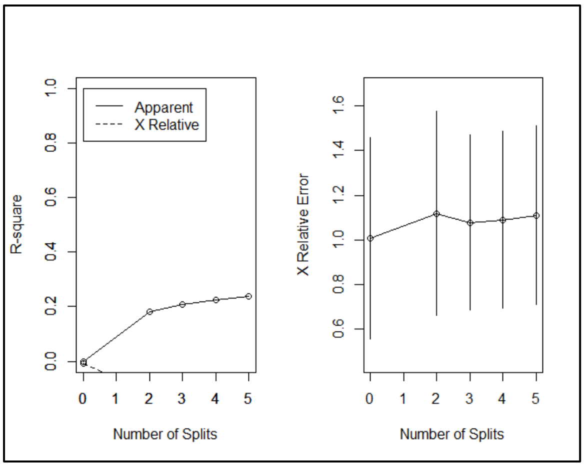


Figure 57: Validation the regression tree explaining relative cropland cover change in Indonesia.

In summary, from the decision trees in this section could be concluded that for Indonesia no significant prove could be found that support the hypotheses about positive migration being related to land cover change. However, a relation was found between large out migration and cropland cover decrease. The only other meaningful split found was based on forest cover in 2001, explaining forest cover change between 2001 and 2010.

4. Discussion

In this chapter the results from chapter 3 will be interpreted and related to existing literature. First the used datasets will be discussed in section 4.1. In 4.2 the results of the statistical analysis of global relations between land cover change and migration will be discussed. In section 4.3 the results of the statistical analysis of national relation between land cover and migration will be discussed and compared to the results of the global analysis. The final section 4.4 is a discussion of the chosen scale in this study.

4.1 Main Characteristics of the Spatially Explicit Datasets

4.1.1 Land Cover Data

The global land cover data have uncertainties which need to be acknowledged when trying to detect changes. Firstly, the data exploration (section 3.1.1) shows remarkable total change values between 2001 and 2010 for land cover types urban land and wetlands. Since the urban land cover change derived from the MODIS raster land cover data was little and had a negative total change contrary to expectations (table 6), it was deemed not to be usable for this study. The unexpected decline in urban land could be a result of urban concentration. People live closer together leaving thinly populated urban areas (Grau & Aide, 2008). However, more likely is that the selected spatial resolution of the applied Modis data is too low for urban change detection. Studies about urban change is usually executed with a high resolution data (Seto et al., 2011; Yang, Xian, Klaver, & Deal, 2003). As a result, the third hypothesis concerning the relation between urban land use and migration could not be tested.

The decline in total wetlands in the tropics between the 2000 and 2010 MODIS data (table 6) is also remarkable. In principle, wetlands are a known problem in global land cover classifications systems. For example, Friedl et al. (2010) describe that the wetlands class is most commonly confused in land cover classification in the Modis data. Accordingly, the accuracy of this land cover class is typically the lowest amongst all land cover classes. Other studies also discuss challenges for classifying wetlands as a result of climatic conditions of the year(Junk et al., 2012; Salari, Zakaria, Nielsen, & Boyce, 2014). Locations where, according to the MODIS data, larger areas of wetland disappeared are the Pantanal wetlands along the border between Brazil and Bolivia, which is under pressure of economic activities, Dolok islands in Papua and South Vietnam and along the river Nile in Sudan and in Zambia next to the Lake Bangweulu. In a comparison between the MODIS and the GLC 2000 global land cover map, the largest difference occurred for the class wetlands (Giri, 2005). However, since not only a relatively low number of cells in the MODIS data are wetland but also this class is not tested in the hypothesis in this study, it is not taken as a major problem to continue.

The cropland class shows the expected large increase over the studies decade (table 6). The total increase in agricultural is the largest change among the classes in the tropics in 2010 followed by the negative total amount of forest change, which is why these two classes are included in the study.

Secondly, Year to year variability is high because of phenology and disturbances such as fire, drought and insect infestation which is why it is not advised to compare the years of the MODIS data (Friedl et al., 2010). In appendix 1 this is shown in scatter plots of year versus count of MODIS pixels of the particular land cover class. Although the total changes seem to be in line with the general trend, this effect can still have occurred.

4.1.2 Migration Data

The migration data show that most regions have a small negative net-migration and some regions have a very high positive net migration, which mostly include a large city (figures 10 & 11, table 9). This can be explained by the significant rural – urban migration streams that typically occur in the tropics (DeFries et al., 2010; Grau & Aide, 2008).

The bimodal distribution around the peak of -15% and 0% (figure 11) in the relative migration results from a combination of a majority of regions that have a small net negative migration and a majority of regions that have almost no migration. As mentioned before the small negative majority is probably the global trend of rural outmigration to urban areas. The peak around zero indicates that net migration is relatively little in most regions. Furthermore, since net migration values are used even large migration streams can have net value of zero when in and outmigration occurs both. Also, migration over short distances within a region cannot be accounted for with the sub-national data while these short-distance movements can be significant.

Some areas have extreme high net migration values: Maharashtra in India, which is mentioned in a census report as one of the highest immigration regions of India (CHAKRAVORTY, 2001). However the value calculated in this study is much higher than in the report, which could be an artefact or because of the difference in year. In Madagascar the regions Antananarivo had a net immigration of 739,465 people. Here, an urban-rural migration and Chinese immigrants are trends (Tremann, 2014).

The lack of sub national observed population in 2010 data input is a substantial concern for the reliability of CIESIN migration dataset (CIESIN, 2011). The used 2010 population inputs were largely an extrapolation of 1990–2000 subnational trends in combination of the World Population Prospects 2008 country-level estimates. Furthermore, the migration data are estimates based on the use of impute natural increase rates applied to the year 2001. Another factor of uncertainty is missing values.

4.2 Relationships between Migration and Land Cover Changes in the Tropics

In general, no significant linear relations could be found between the land cover and migration variables (figures 12, 13 & 14). Furthermore, a non-linear approach using decision trees did not find any significant splits based on the migration data to predict the land cover change. Forest change and cropland change were explained by the amount of the land cover type in 2001 rather than by migration variables.

Most likely interactions with factors beyond migration play a role in the relation between migration and land cover change. Lambin et al. (2001) stress that population dynamics alone cannot explain global land cover change as local policies and economy have a major impact which should therefore always be included in an analysis. Furthermore, many studies also find that biophysical characteristics per region strongly influence the relations between population and land cover change (Bonilla-Moheno, Aide, & Clark, 2012; Etter, McAlpine, Wilson, Phinn, & Possingham, 2006). As a result, interactions with policy and biophysical variables should be included in the model. However, it is rather difficult to include more factors in a global study since sufficient datasets are not available for all factors. Although data for biophysical factors such as elevation (Aster GDEM), climate (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) and ecoregions are available for the world, global policies and economic factors are less present especially with a finer resolution than country level.

The CART method that was used is a suitable method for this study as it is able to deal with all kind of data, whereby no distribution assumptions are made. In addition, the method is not affected by outliers, which were often present in the used datasets. CART places outliers in isolated nodes. Also favourable is that it can reveal interactions in the dataset, such as that when outmigration in a region is large; the average cropland change is also determined by the cropland area in the region in 2001.

However, a disadvantage of the model is that it only decides which split is best at each step, not overall. As a result some overall interaction may not be found unless it is purposely added to the data frame. Another weakness of the method is that overfitting can occur as it can produce very complex trees to explain complex data. The fact that after pruning the trees often solely the root was left show that the trees where over fitted.

4.3 Comparison to Result on a National scale

For Colombia, a moderate negative correlation was found between relative cropland change and migration which contrasts the hypothesis that positive migration is related to a cropland cover increase. The correlations between the forest change and migration variables are negligible. Furthermore, the correlation between the different land cover variables was a weak negative, suggesting that cropland cover increase is related to forest cover increase in Colombia. The regression trees explaining forest change contradicted hypothesis 1 which was that that positive migration is related to negative forest change. Also a negative relation between cropland cover change and urban cover in 2001 was detected with the regression trees.

Sánchez-Cuervo & Aide (2013) concluded for Colombia that on different scales; country, biome and ecoregion, different relations could be found between woody vegetation cover and demographic, socio-economic and environmental drivers. The presence of armed illegal favours caused deforestation on strategic locations for illegal practices because the forest is used as a cover. On the other hand, areas rich in minerals sources were more likely to be deforested. Etter et al. (2006) found that accessibility was an important variable for explaining deforestation in Colombia at both regional and national scale.

According to Wicke, Sikkema, Dornburg, & Faaij, (2011), Indonesia is mostly characterized by deforestation. Furthermore, it has a strong rural-urban migration (Liu & Yamauchi, 2014). This is in line with the observations in the data exploration where mostly net migration is visible (figure 42) except for regions with large cities.

Remarkable in the data exploration for land cover change in Indonesia was the high number of 100% cropland loss. The regions with the largest absolute cropland cover loss of these regions are located on Sumatra. These croplands turned into the land cover class Cropland/Natural vegetation mosaic indicating that probably they were abandoned because of unfavourable circumstances for cropland such as soil degradation or because of outmigration to cities (Döös, 2002). With the regression trees in this study a relation between large net regional outmigration and cropland cover decrease was found.

Similar to the global study, no linear correlations were found between land cover change and migration variables in Indonesia. The produced regression trees for all land cover change variables could provide evidence to support the hypotheses of a positive relation between positive migration and land cover change. In line with the global analysis, a relation was found between the forest stock in 2001 in the region and the forest cover change.

4.4 Resolution

The sub national regions were chosen to mitigate the uncertainties in the migration data. Since the migration estimations were based on census data with subnational units, aggregating to similar units seems to makes sense. A disadvantage of this unit is that the size varies widely over regions. Although with the harmonization of the two level regions reduced the range of the sizes, there was still large variability in the sizes. When comparing the relative migration and land cover variables to the absolute ones, it shows that the size of the regions influenced the absolute variables. Large regions can have higher values. This is also shown by the regression trees using solely the particular land cover type to explain the change in that land cover type. For this reason the use of a grid with equal size units could improve the quality of the research. However, a higher certainty of migration data is required, which does not exist yet. In the future this will be produced. Furthermore, the chosen aggregated resolution does not show what happened within the region. A large dynamic within a region can when summed up result in a small change value.

Also, it seems that chosen resolution of land cover data (10 x 10 km) was too coarse as pixels were assigned to the dominant LC class, which took away a lot of detail. However, since the aggraded to subnational resolution regions resolution was used it would probably not improve this study substantially. When better-quality migration data will be available, using higher resolution land cover data is recommended.

Furthermore, it could be that migration is not an important predictor of land cover change at regional resolution not because it is a too coarse level but because larger scale processes play a role. For example the study of Birch-Thomsen et al. (2010) showed that increased food demand of a growing population on the Bellona Island, Solomon Islands, was covered by import. Another well-known example of export of soy beans from Brazil to China (Nepstad et al., 2006). The population increase in Chinese cities leads to increased soy production in Brazil and not to deforestation in China.

5. Summary and Conclusion

This study's aim was to answer the following questions.

RQ1 Which spatially explicit datasets of demographic dynamics and land cover changes exist and what are their main characteristics and metadata?

RQ2 What are the relationships between human migration and land cover changes if analysed over the tropics?

To answer this research question the following hypotheses shall be tested:

H1: Deforestation is positively related to immigration.
H2: Agricultural land cover increase is positively related to immigration.
H3: Urban land cover expansion is positively related to immigration.

RQ3 Are relationships stronger or different when considering selected countries and finer resolution data?

It was possible to do an exploratory analysis about global relations between migration and land cover change because of increasing available data. However, the data is still limited in terms of certainty, resolution and themes. The current existing CIESIN migration dataset is uncertain especially for the decade 2000 – 2010. However, global land cover data was only available for this period. The land cover change data also are not optimal as post-classification comparison had to be used.

No linear correlations were found on a global scale. Also the non-linear regression tree could not find significant contributions of migration explaining land cover change. As a result none of the hypotheses can be accepted. Only the stock of the land cover type in the region in 2001 could explain the change of the land cover type. In the national scale study for Indonesia a relation between outmigration and cropland cover decrease was found with a regression tree. From this can be concluded that is important to compare different scales when researching relation between migration and land cover change, as they can be different.

It could be that no relation exist because of for example global economy with export between countries. However, it could also be that interactions exits with additional factors beyond migration such as biophysical regions and local policies, which would be necessary to include finding a relation between migration and land cover change. Furthermore, the used resolution of subnational regions with varying sizes probably influences the results. However, with the current data it was not considered appropriate to work with finer resolution data. The uncertainty of both the MODIS land cover, but especially the CIESIN migration data probably influenced the results.

This study was explorative and showed that with current available data no relations could be found. However, since in the future better quality data will be available it could be improved by using a higher resolution with equal unit sizes. Furthermore it is recommended to include more factors that might have an interaction with the relation between migration and land cover change such as biophysical regions and policies.

References

Aide, T. M., Clark, M. L., Grau, H. R., López-Carr, D., Levy, M. A., Redo, D., … Muñiz, M. (2013). Deforestation and Reforestation of Latin America and the Caribbean (2001-2010). *Biotropica*, *45*, 262–271. doi:10.1111/j.1744-7429.2012.00908.x

Aide, T. M., & Grau, H. R. (2004). Ecology. Globalization, migration, and Latin American ecosystems. *Science (New York, N.Y.)*, *305*, 1915–1916. doi:10.1126/science.1103179

Bartholomé, E., & Belward, A. S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*. doi:10.1080/01431160412331291297

Birch-Thomsen, T., Reenberg, A., Mertz, O., & Fog, B. (2010). Continuity and change: Spatiotemporal land use dynamics on Bellona Island, Solomon Islands. *Singapore Journal of Tropical Geography*, *31*, 27–40. doi:10.1111/j.1467-9493.2010.00383.x

Bonilla-Moheno, M., Aide, T. M., & Clark, M. L. (2012). The influence of socioeconomic, environmental, and demographic factors on municipality-scale land-cover change in Mexico. *Regional Environmental Change*, *12*, 543–557. doi:10.1007/s10113-011-0268-z

Bontemps, S., Defourny, P., Bogaert, E. V, Arino, O., Kalogirou, V., & Perez, J. R. (2011). *GLOBCOVER 2009 - Products Description and Validation Report*. Retrieved from http://due.esrin.esa.int/globcover/LandCover2009/GLOBCOVER2009_Validation_Report_2.2.pdf

Carr, D. (2009). Population and deforestation: why rural migration matters. *Progress in Human Geography*. doi:10.1177/0309132508096031

Carr, D. L. (2004). Proximate population factors and deforestation in tropical agricultural frontiers. *Population and Environment*. doi:10.1023/B:POEN.0000039066.05666.8d

CHAKRAVORTY, C. (2001). *Census of India 2001*.

CIESIN. (2011). *Migration and Global Environmental Change, MR4: Estimating net migration by ecosystem and by decade: 1970–2010*. Retrieved from http://www.ciesin.org/documents/foresight-2011-oct.pdf

Clark, W. A. V, & Hosking, P. L. (1986). *Statistical methods for geographers* (Vol. 518). Wiley New York.

Congalton, R. G., Gu, J., Yadav, K., Thenkabail, P., & Ozdogan, M. (2014). Global land cover mapping: a review and uncertainty analysis. *Remote Sensing*, *6*(12), 12070–12093.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., … van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*. doi:10.1038/387253a0

DeFries, R. S., Rudel, T., Uriarte, M., & Hansen, M. (2010). Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience*.

Domencich, T. A., & McFadden, D. (1975). *Urban Travel Demand-A Behavioral Analysis*.

Döös, B. R. (2002). Population growth and loss of arable land. *Global Environmental Change*. doi:10.1016/S0959-3780(02)00043-2

Etter, A., McAlpine, C., Wilson, K., Phinn, S., & Possingham, H. (2006). Regional patterns of agricultural land use and deforestation in Colombia. *Agriculture, Ecosystems and Environment*, *114*(2-4), 369–386. doi:10.1016/j.agee.2005.11.013

Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., … Snyder, P. K. (2005). Global consequences of land use. *Science (New York, N.Y.)*, *309*, 570–574. doi:10.1126/science.1111772

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X., Muchoney, D., strahler, A. H., … Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing Of Environment*, *83*, 287–302. doi:10.1016/S0034-4257(02)00078-0

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182. doi:10.1016/j.rse.2009.08.016

Gasparri, N. I., & Grau, H. R. (2009). Deforestation and fragmentation of Chaco dry forest in NW Argentina (1972-2007). *Forest Ecology and Management*, *258*(6), 913–921. doi:10.1016/j.foreco.2009.02.024

Giri, C. (2005). Global Land Cover Mapping and Characterization: Present Situation and Future Research Priorities. *Geocarto International*. doi:10.1080/10106040508542334

Grau, H. R., & Aide, M. (2008). Globalization and land-use transitions in Latin America. *Ecology and Society*, *13*. doi:10.1057/9780230603554

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. *Science (New York, N.Y.)*, *319*, 756–760. doi:10.1126/science.1150195

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., … Williams, N. S. G. (2009). A global synthesis of plant extinction rates in urban areas. *Ecology Letters*, *12*, 1165–1173. doi:10.1111/j.1461-0248.2009.01372.x

Hansen, M. C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S. a, Tyukavina, a, … Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science (New York, N.Y.)*, *342*, 850–3. doi:10.1126/science.1244693

Hartemink, A. E., Veldkamp, T., & Bai, Z. (2008). Land cover change and Soil fertility decline in tropical regions. *Turkish Journal of Agriculture and Forestry*, *32*, 195–213.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. doi:10.1002/joc.1276

Huijun, G., Padoch, C., Coffey, K., Aiguo, C., & Yongneng, F. (2002). Economic development, land use and biodiversity change in the tropical mountains of Xishuangbanna, Yunnan, Southwest China. *Environmental Science & Policy*. doi:10.1016/S1462-9011(02)00093-X

Junk, W. J., An, S., Finlayson, C. M., Gopal, B., Květ, J., Mitchell, S. A., … Robarts, R. D. (2012). Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. *Aquatic Sciences*. doi:10.1007/s00027-012-0278-z

Kauppi, P. E., Ausubel, J. H., Fang, J., Mather, A. S., Sedjo, R. A., & Waggoner, P. E. (2006). Returning forests analyzed with the forest identity. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(46), 17574–17579. doi:10.1073/pnas.0608343103

Klein Goldewijk, K., & Ramankutty, N. (2004). Land cover change over the last three centuries due to human activities: The availability of new global data sets. *GeoJournal*. doi:10.1007/s10708-004-5050-z

Kok, K., & Veldkamp, A. (2001). Evaluating impact of spatial scales on land use pattern analysis in Central America. *Agriculture, Ecosystems & Environment*. doi:10.1016/S0167-8809(01)00185-2

Lambin, E. F., & Geist, H. (2006). *Land-use and Land-cover change: Local Processes and Global Impacts*. *Land Use Policy* (Vol. 12). doi:10.1016/0264-8377(95)00005-X

Lambin, E. F., Geist, H. J., & Lepers, E. (2003). DYNAMICS OF LAND-USE AND LAND-COVER CHANGE IN TROPICAL REGIONS. *Annual Review of Environment and Resources*, *28*, 205–241. doi:10.1146/annurev.energy.28.050302.105459

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., … Xu, J. (2001). The causes of land-use and land-cover change: Moving beyond the myths. *Global Environmental Change*, *11*, 261–269. doi:10.1016/S0959-3780(01)00007-3

Lawrence, P. J., & Chase, T. N. (2010). Investigating the climate impacts of global land cover change in the community climate system model. *International Journal of Climatology*, *30*, 2066–2087. doi:10.1002/joc.2061

Lele, N., Nagendra, H., & Southworth, J. (2010). Accessibility, demography and protection: Drivers of forest stability and change at multiple scales in the Cauvery Basin, India. *Remote Sensing*, *2*, 306–332. doi:10.3390/rs2010306

Liu, Y., & Yamauchi, F. (2014). Population density, migration, and the returns to human capital and land: Insights from Indonesia. *Food Policy*, *48*, 182–193.

Loveland, T. R., Zhu, Z. L., Ohlen, D. O., Brown, J. F., Reed, B. C., & Yang, L. M. (1999). An analysis of the IGBP global land-cover characterization process. *Photogrammetric Engineering and Remote Sensing*, *65*(9), 1021–1032. Retrieved from <Go to ISI>://000082486400006

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics* (pp. 105–142). doi:10.1108/eb028592

Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-being: Synthesis*. *Ecosystems* (Vol. 5). doi:10.1196/annals.1439.003

Nepstad, D. C., Stickler, C. M., & Almeida, O. T. (2006). Globalization of the Amazon soy and beef industries: Opportunities for conservation. *Conservation Biology*. doi:10.1111/j.1523-1739.2006.00510.x

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., … Hayes, D. (2011). A large and persistent carbon sink in the world's forests. *Science (New York, N.Y.)*, *333*, 988–993. doi:10.1126/science.1201609

Ramankutty, N., Foley, J. A., & Olejniczak, N. J. (2002). People on the land: changes in global population and croplands during the 20th century. *Ambio*, *31*, 251–257. doi:10.1579/0044-7447-31.3.251

Redo, D. J., Aide, T. M., & Clark, M. L. (2012). The Relative Importance of Socioeconomic and Environmental Variables in Explaining Land Change in Bolivia, 2001–2010. *Annals of the Association of American Geographers*. doi:10.1080/00045608.2012.678036

Reidsma, P., Tekelenburg, T., van den Berg, M., & Alkemade, R. (2006). Impacts of land-use change on biodiversity: An assessment of agricultural biodiversity in the European Union. *Agriculture, Ecosystems & Environment*. doi:10.1016/j.agee.2005.11.026

Ricketts, T. H., Daily, G. C., Ehrlich, P. R., & Michener, C. D. (2004). Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 12579–12582. doi:10.1073/pnas.0405147101

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., … Foley, J. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, *14*. doi:10.1038/461472a

Rudel, T. K., Coomes, O. T., Moran, E., Achard, F., Angelsen, A., Xu, J., & Lambin, E. (2005). Forest transitions: Towards a global understanding of land use change. *Global Environmental Change*, *15*(1), 23–31. doi:10.1016/j.gloenvcha.2004.11.001

Salari, A., Zakaria, M., Nielsen, C. C., & Boyce, M. S. (2014). Quantifying Tropical Wetlands Using Field Surveys, Spatial Statistics and Remote Sensing. *Wetlands*, *34*(3), 565–574.

Sánchez-Cuervo, A. M., & Aide, T. M. (2013). Consequences of the Armed Conflict, Forced Human Displacement, and Land Abandonment on Forest Cover Change in Colombia: A Multi-scaled Analysis. *Ecosystems*, *16*, 1052–1070. doi:10.1007/s10021-013-9667-y

Schneider, A., Friedl, M. A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*. doi:10.1088/1748-9326/4/4/044003

Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. *PLoS ONE*, *6*. doi:10.1371/journal.pone.0023777

Therneau, T. M., & Atkinson, B. (2012). rpart: Recursive Partitioning. R package version 3.1-51. *Http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm*.

Therneau, T. M., Atkinson, B., & Ripley, B. D. (2006). rpart: Recursive Partitioning. *Rpart Package Manual*. Retrieved from http://www.mayo.edu/hsr/Sfunc.html

Tremann, C. (2014). Temporary Chinese migration to Madagascar: local perceptions, economic impacts, and human capital flows. *African Review of Economics and Finance*, *5*(1), 7–16.

Tsegaye, D., Moe, S. R., Vedeld, P., & Aynekulu, E. (2010). Land-use/cover dynamics in Northern Afar rangelands, Ethiopia. *Agriculture, Ecosystems and Environment*, *139*, 174–180. doi:10.1016/j.agee.2010.07.017

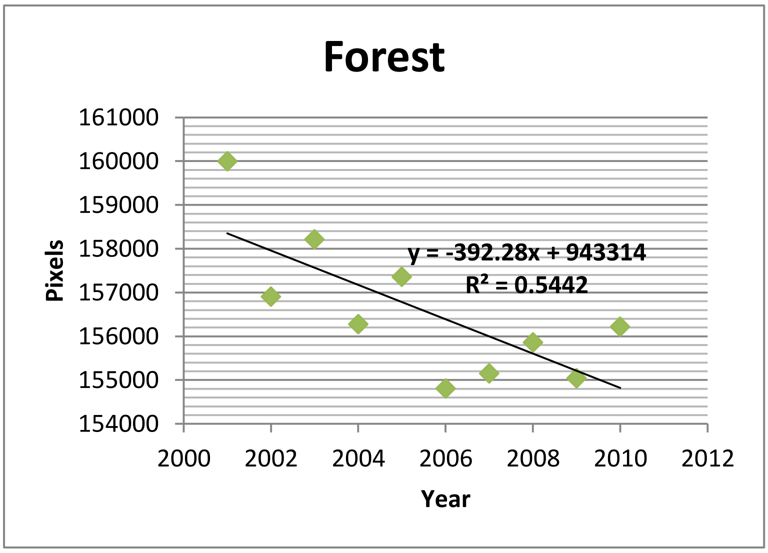
Van der Werf, G. R., Morton, D. C., DeFries, R. S., Olivier, J. G. J., Kasibhatla, P. S., Jackson, R. B., … Randerson, J. T. (2009). CO2 emissions from forest loss. *Nature Geoscience*. doi:10.1038/ngeo671

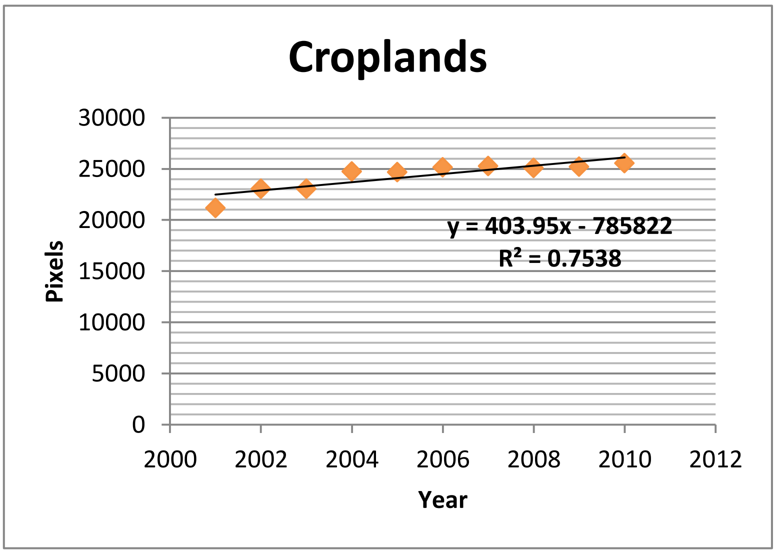
Wicke, B., Sikkema, R., Dornburg, V., & Faaij, A. (2011). Exploring land use changes and the role of palm oil production in Indonesia and Malaysia. *Land Use Policy*, *28*(1), 193–206. doi:10.1016/j.landusepol.2010.06.001

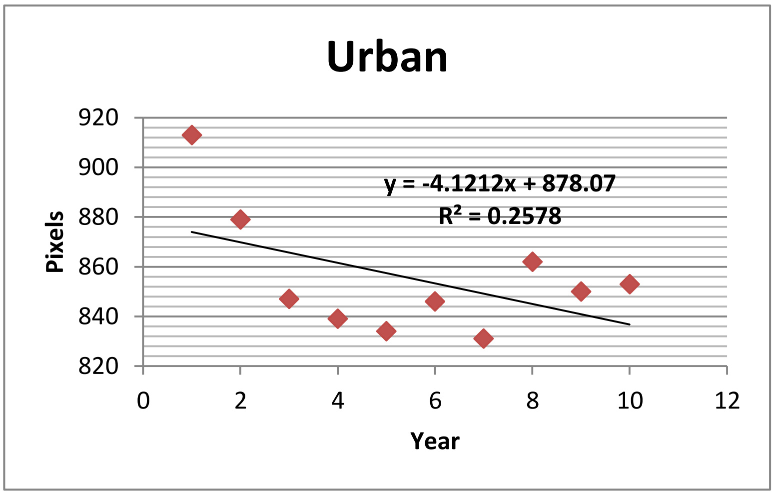
Yang, L., Xian, G., Klaver, J. M., & Deal, B. (2003). Urban Land-Cover Change Detection through Sub-Pixel Imperviousness Mapping Using Remotely Sensed Data. *Photogrammetric Engineering Remote Sensing*, *69*(9), 1003–1010. doi:10.14358/PERS.69.9.1003

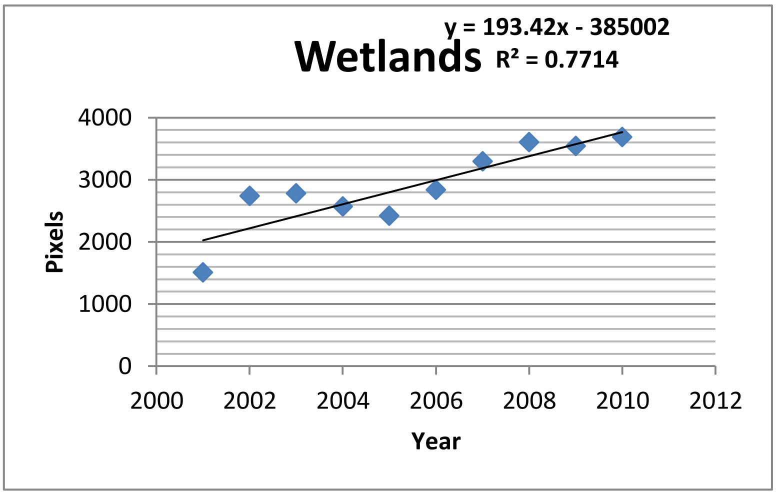
Appendix 1

Plots of the annual counts of pixels in the MODIS land cover data per land cover type for forest, cropland, urban an wetlands. It can be seen here that the inter-annual variability exist especially for forest and urban.






Urban
920
900
880
860
840
820
Pixels
y = -4.1212x + 878.07
R² = 0.2578
0
2
4
6
8
10
12
Year


Wetlands
y = 193.42x - 385002
R² = 0.7714
4000
3000
2000
1000
0
Pixels
2000
2002
2004
2006
2008
2010
2012
Year