Masters Program in Geospatial Technologies

SPATIO-TEMPORAL DATA MODELING IN RESPONSE TO DEFORESTATION MONITORING

(A CASE STUDY OF SMALL REGION IN RIAU PROVINCE, INDONESIA)

DIAN NURAINI MELATI

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

UNIVERSITAT JAUME·I

ifgi
Institute for Geoinformatics
University of Münster

Supported by:
Education and Culture
Erasmus Mundus

Spatio-Temporal Data Modeling in Response to Deforestation Monitoring

(A case study of small region in Riau Province, Indonesia)

Supervisor:

Prof. Dr. Edzer Pebesma (University of Münster, Germany)

Co-supervisors:

Prof. Mário Caetano (New University of Lisbon, Portugal)

Prof. Filiberto Pla (University Jaume I Castellon, Spain)

Declaration of originality

This is to certify that this thesis is completely my own work without help of other persons, unless clearly acknowledged (including citation of published and unpublished sources). The thesis has not been submitted before to other universities or to any other institution.

Signed _____

Date ____28/02/2012_____

ABSTRACT

Indonesia with large amount of area covered by tropical forest faces a critical problem of deforestation. A lot of forested areas were converted into other coverage influenced by human activities. Therefore, deforestation monitoring and forest prediction have to be done in order to manage the sustainability of forest. To monitor deforestation, this research has analyzed the trend of forest cover in the study area by combining NDVI differencing and image classification to describe the forest cover change. In order to do that, Landsat images acquired in different time (1996, 2000, and 2005) have been chosen as input. NDVI differencing has been conducted by doing normalization of one image to another image initially. Subsequently, thresholds to identify the change and no change have been carried out separately for decrease and increase part. Apart from that, image classification was applied using supervised classification. Eventually, land cover change detection has been performed by combining NDVI differencing and image classification. It has been proved by the research that forest in study area has decreased by 6% during 1996-2005.

In order to forecast future forest cover, three models were chosen to get the best model for prediction. These models are Stochastic Markov Modal, Cellular Automata Markov (CA_Markov) Model, and GEOMOD. To measure the best model among them, Kappa index was employed to validate the simulation. As the result, GEOMOD performed the highest Kappa. Therefore, GEOMOD was implemented to model forest cover in 2015. The result of GEOMOD implementation revealed that forest cover will be decreased by 12% during 2005-2015.

Keywords: Forest, Deforestation, Monitoring, Prediction, NDVI differencing, Image Classification, Landsat, Normalization, Supervised Classification, Stochastic Markov Model, CA_Markov Model, GEOMOD, Kappa index, Validate

ACKNOWLEDGEMENTS

I would like to say Alhamdulillah and special thanks to Allah for the protection, grace and praise during my course and my life until now.

My sincere gratitude goes to European Union through Erasmus Mundus Scholarship for the opportunity to pursue my studies. I express my appreciation to Prof. Dr. Marco Painho, Prof. Dr. Werner Kuhn, Dr. Christoph Brox, Prof. Dr. Joaquín Huerta and all the staff members of UNL, ifgi and UJI for the facilities and coordination the program.

It is my great pleasure to acknowledge Prof. Dr. Edzer Pebesma, Prof. Mário Caetano and Prof. Filiberto Pla for the discussions, inputs and comments during my thesis work. Their competence and expertise drew my thesis more valuable. I am also grateful to Prof. Dr. Pedro Cabral, Prof. Dr. Ana Cristina Costa and Prof. Dr. Angela Schwering for giving magnificent courses.

I would like to acknowledge my colleagues in The Agency for the Assessment and Application of Technology (BPPT), Jakarta-Indonesia to allow me temporarily leaving my work for continuing my study. My thanks also go to all my friends during my stay and study for the friendship and international environment.

Last but not least, I remain indebted to my mother for her support and pray during my stay abroad and also to my sisters and all families who always encouraged me to accomplish my studies. Special gratitude goes to my husband for his patience and moral support especially for taking care of me during my last semester.

TABLE OF CONTENTS

ABSTRACT ..... **i**
ACKNOWLEDGEMENTS ..... **ii**
TABLE OF CONTENTS ..... **iii**
INDEX OF FIGURES ..... **v**
INDEX OF TABLES ..... **vi**

CHAPTER I **1**
INTRODUCTION **1**
1.1. Background 1
1.2. Statement of the Problem 2
1.3. Study Area 3
1.4. Aim and Objectives 5
1.5. Research Hypotheses 5
1.6. Research Limitations 6
1.7. Materials and Tools 6
CHAPTER 2 ..... **8**
CONCEPTS AND DEFINITIONS **8**
2.1. Remotely Sensed Data Implementation for Forest Monitoring 8
2.2. Land Use and Land Cover Change (LUCC) Detection 10
2.3. Forest and Deforestation 11
2.4. Land Use and Land Cover Change Modeling 12
CHAPTER 3 ..... **14**
LAND COVER CHANGE DETECTION **14**
3.1. Spatial Data Preparation **15**
3.1.1. Image Pre-Processing 15
3.1.2. Image Processing 16
3.1.2.1. Normalized Difference Vegetation Index (NDVI) differencing 16
3.1.2.2. Image Classification 24
3.1.3. Image Post-Processing 26
3.2. Land Cover Change Detection Analysis **28**
CHAPTER 4 ..... **31**
FOREST COVER MODELING **31**
4.1. Calibration **31**

4.1.1. Stochastic Markov 31
4.1.1.1. Stochastic Markov Process 32
4.1.1.2. Stochastic Markov Model 33
4.1.2. Cellular Automata Markov (CA_Markov) 35
4.1.2.1 Cellular Automata Markov (CA_Markov) Process 36
4.1.2.2 Cellular Automata Markov (CA_Markov) Model 38
4.1.3. GEOMOD 40
4.2. Validation 46
4.3. Discussions 48
CHAPTER 5 49
FUTURE FOREST COVER MODELING 49
5.1. Modeling 49
5.2. Discussions 53
CHAPTER 6 54
CONCLUSIONS AND RECOMMENDATIONS 54
6.1. Conclusions 54
6.2. Recommendations 55
REFERENCES 56
APPENDICES 60
Appendix A-Parameters of Normalization 60
Appendix B-Basic Terminologies of Accuracy Assessment 62
B.1. Overall Accuracy Assessment 62
B.2. Kappa 63
Appendix C-Accuracy Assessment 65
C.1. Accuracy Assessment of Classified Image in 1996 65
C.2. Accuracy Assessment of Classified Image in 2000 66
C.3. Accuracy Assessment of Classified Image in 2005 67
Appendix D-Quantity Prediction Using GEOMOD 68
D.1. Quantity Prediction of Forest Cover in 2005 68
D.2. Quantity Prediction of Forest Cover in 2015 69

INDEX OF FIGURES

Figure 1.1 Forest cover change by major island group (1990-2000) 3
Figure 1.2 Forest condition in study area 4
Figure 3.1 Flow chart of land cover change detection 14
Figure 3.2 Landsat TM Composite of study area in 1996: (a) Composite of band 3, 2, 1 and (b) Composite of band 5, 4, 3 16
Figure 3.3 NDVI images of (a) 1996, (b) 2000 and (c) 2005 17
Figure 3.4 Histograms of NDVI images. The histograms were extracted from (a) 1996, (b) 2000, and (c) 2005 NDVI images and predicted NDVI histograms for (d) 2000 NDVI (Y) predicted from 1996 NDVI (X), (e) 2005 NDVI (Y) predicted from 1996 NDVI (X), and (f) 2005 NDVI (Y) predicted from 2000 NDVI (X) 19
Figure 3.5 Histograms of NDVI differencing between (a) 1996 and 2000, (b) 1996 and 2005, (c) 2000 and 2005 20
Figure 3.6 The dissemination of probability density functions within NDVI differencing 21
Figure 3.7 The selection of optimal c value 22
Figure 3.8 Change and no change images between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005 23
Figure 3.9 Land cover types in study area (a) 1996, (b) 2000 and (c) 2005 26
Figure 3.10 Distribution of training samples for accuracy assessment 27
Figure 3.11 Land cover change in study area 29
Figure 3.12 Proportion of forest cover change in the study area 29
Figure 3.13 Forest cover change images between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005 30
Figure 4.1 Flow chart to model forest covers change 31
Figure 4.2 Flow chart of Stochastic Markov Model 32
Figure 4.3 Example of Markov Chain 32
Figure 4.4 Conditional Probability Images 34
Figure 4.5 Projected Forest cover in 2005 using Stochastic Markov Model 35
Figure 4.6 Flow chart of Cellular Automata Markov Model 35
Figure 4.7 Two-dimensional cellular automata 36
Figure 4.8 Distance Images of each land cover in 2000 (unit: meters) 39
Figure 4.9 Suitability Images of each land cover in 2000 40
Figure 4.10 Projected Forest cover in 2005 using CA_Markov Model 40
Figure 4.11 Flow chart of GEOMOD 41
Figure 4.12 Distance Images of driving factors (2000) (unit: meters) 43
Figure 4.13 Suitability Images of driving factors (2000) 44
Figure 4.14 Aggregated Suitability Images 45
Figure 4.15 Projected Forest cover in 2005 using GEOMOD 45
Figure 4.16 Prediction error map 2005 of (a) Stochastic Markov, (b) CA_Markov, (c) GEOMOD 47
Figure 5.1 Distance Images of driving factors (2005) (unit: metres) 50
Figure 5.2 Suitability Images of driving factors (2005) 51
Figure 5.3 Aggregated Suitability Images (2005) 52
Figure 5.4 Projected Forest cover in 2015 52
Figure A.1 Distribution of training samples for normalization between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005 60
Figure A.2 Relationship between NDVI values of (a) 1996-2000, (b) 1996-2005 and (c) 2000-2005 61

INDEX OF TABLES

Table 1.1 Data used in the research 7
Table 2.1 Characteristics of Landsat 7 ETM+ 9
Table 3.1 Thresholds of decrease and increase for different NDVI differencing 23
Table 3.2 Land cover types in study area 24
Table 3.3 The area of each land cover 28
Table 4.1 Transition Probability Matrix between Forest and Non Forest 34
Table 4.2 Transition Area Matrix between Forest and Non Forest 34
Table 4.3 Factor Weights for the Suitability Images 41
Table 4.4 Categorization of Kappa Index 46
Table 4.5 Kappa Index for each model 47
Table 5.1 Specification of GEOMOD parameters 49
Table B.1 Confusion matrix between reference and classified map 62
Table B.2 Contingency Table of *J* Categories in which inputs are the proportions in study area 63
Table B.3 The Proportion correct classification regarding a simulation capability to identify correctly location and quantity 64

CHAPTER I

INTRODUCTION

1.1. Background

Forest as important natural resources along with large quantity of biodiversity has a significant role to keep the balance of carbon existence. It has an obvious position regarding the climate change effect due to the greenhouse gases (GHGs) increment especially CO_2 which shares a big contribution as a source of emissions (Metz et al., 2007). Forest can be as a sink of carbon emission as well as a source of carbon emissions. As a sink, forest absorbs and stores large amount of carbon because it needs CO_2 for photosynthesis process. However, forest also can be a source of carbon emissions in some cases such as illegal logging, peat fires, and deforestation. In synthesis report, which had been performed by International Panel on Climate Change (IPCC), forestry contributed 17.4% of GHG emissions after energy supply and industry (Pachauri, 2007) and one of the reasons is tropical deforestation.

Tropical deforestation has been the mainly source of CO_2 flux rising in term of land use change in the past two decades (Denman et al., 2007). Alteration of forest to agriculture is the primary reason of deforestation; it is about 13 million ha per year (FAO, 2005). Furthermore, FAO revealed that net change of forest was declined around 7.3 million ha per year that actually it was declined around 8.9 million ha per year from 1990 to 2000. They emphasized in Global Forest Resources Assessment 2005 that tropical forest in Brazil and Indonesia shared the most deforestation in 2000-2005. The way to reduce emission from deforestation should be performed in order to minimize the effect of CO_2 rising caused by the land use change of forest. Formally, UNFCC (2001) defines “deforestation as the direct human-induced conversion of forested land to non-forested land”; furthermore, it comprises the conversion of forest into agriculture, pasture, water reservoirs, and urban areas and not taking into account the region in which the trees are taken away due to the harvesting or logging issues except these activities are continued by substituting the logged area to other land uses (Schoene, D, 2007).

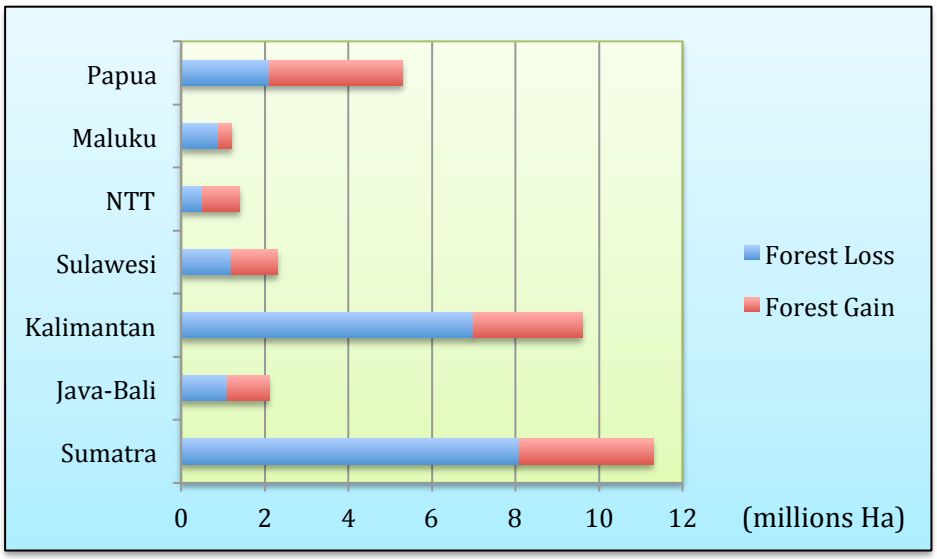
The decreased of forest cover can be assessed using remote sensing technology which is integrated with field measurements (De Fries et al., 2006).

Nowadays, a large number of remotely sensed data can be found with varying characteristics either in term of spatial or spectral characteristic and also whether active or passive system. In analysis of changing, temporal issue of the data has the essential role for change monitoring. The time-series ability provided by satellite remote sensing supplies the trend of land use/land cover (Holmgren, 2008). This capacity to capture data periodically brings a benefit to monitor the change of forest. In order to reach the goal of this study, this research used remotely sensed data acquired in different time and applies some images processing techniques to analyze deforestation. Having deforestation information, modeling for future can be performed to forecast what is going to be happened in the future. This information is important in order to conserve and sustain the management of forest to support Reducing Emissions from Deforestation and Forest Degradation (REDD) Program. The case study of this research is a part of tropical forest in Indonesia. It will be done in a part of Sumatran forest which is located in Riau Province.

1.2. Statement of the Problem

Indonesia is covered by forest around more than 90 million hectares; almost covers 46 percent of land area in Indonesia (FAO, 2007). It brings benefits in making economic profits for all Indonesians and to manage the biodiversity and all environmental aspects but almost 30 percent of this area are not forested due to degradation and deforestation reaching 2 million hectares per year (World Bank, 2006). Moreover, in the same report by analyzing using remote sensing data, it exposed that Indonesia lost about 20 million hectares in period 1990 and 2000, which is containing 8 million hectares of Sumatra's forest and 6.9 million hectares of Kalimantan's forest. These phenomena are dominantly caused by illegal activities such as excessive cutting of forest, cutting in protected area and also having forest fire in the late 1990s (FAO, 2007).

Most of the large forestland can be found in Kalimantan, Sumatra, and Papua. The figure below shows the forest cover change in period 1990-2000 by major island in Indonesia. Sumatra had the largest forest loss among the other islands and followed by Kalimantan, it is because of economically accessible area of Sumatra (World Bank, 2006). Nonetheless, forest gain was also achieved due to reforestation activities led by the Government.



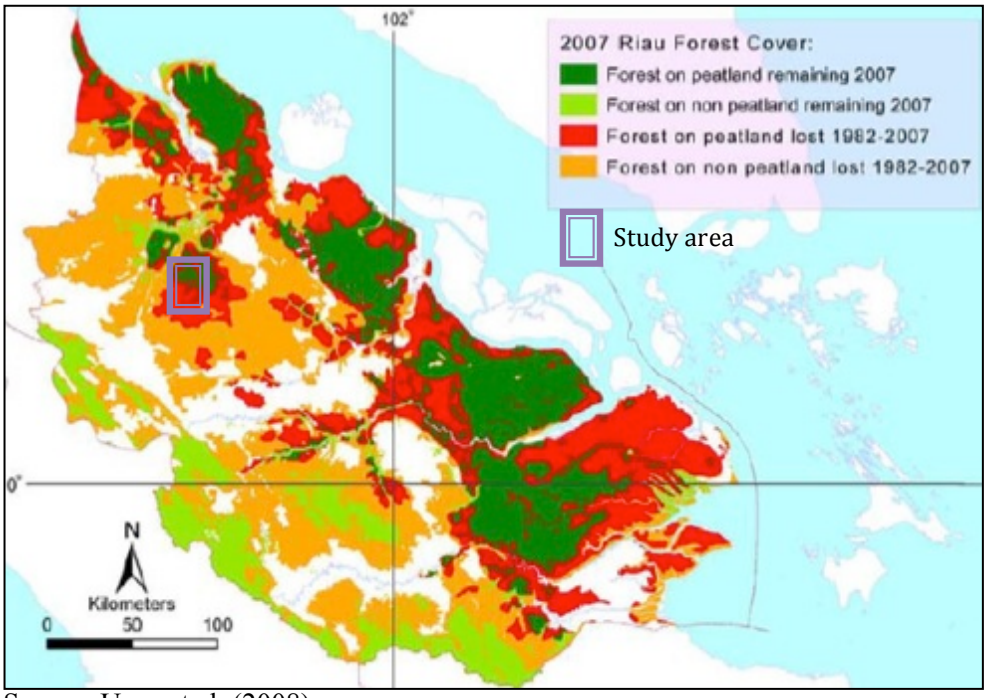
Source: World Bank Analysis (2006)

Figure 1.1 Forest cover change by major island group (1990-2000)

Riau as a province in Indonesia contributed the largest deforestation compared to other provinces in Sumatra Island. In 1982-2007, forest in Riau was decreasing about 65% (4,166,381 ha) and shrinking from 6,420,499 ha to 2,254,118 ha (Uryu et al., 2008). Other assessment conducted by government found that the deforestation rate in period 2003-2006 is 158.130,6 Ha per year (MoF, 2008). Having deforestation issues, it is necessary to monitor and control the forest cover in this area by monitoring and predicting the forest cover in the future so that will support a good decision for related stakeholders in managing the forest sustainability. In order to achieve the goal of this research, it is focused on such particular area in Riau that has been experienced with deforestation.

1.3. Study Area

The study area is located in a forest area around Rokan Hulu District, Riau Province, Indonesia. However, the study area does not cover the whole district because remotely sensed data used in this research are not clear enough of cloud cover so it was decided to select particular area which covers forest area and has the least cloud cover and also the important thing is that this area experienced deforestation as shown in figure 1.2 based on World Wild Foundation report (Uryu et al., 2008). Rokan Hulu district consists of 85% land and 15% water area and wetlands. The location of study area is located between the coordinates 0°58'27" N – 1°10'N and 100° 35'31" E - 100°44'10" E.



Source: Uryu et al. (2008)

Figure 1.2 Forest condition in study area

The climate of Rokan Hulu is commonly tropical wet with an average maximum temperature of 31° C- 32° C. Rokan Hulu is the fourth district in Riau Province, which has the highest amount of rain each year around 136 days. The average annual rainfall is about 227.06 mm based on the data calculated in 2009. The large number of rainy days throughout the year is located around the area of Pasir Pengarayan and the least number of them is located in around Kota Tengah (Centre Statistics Agency of Rokan Hulu, 2009).

The results of population registration in Rokan Hulu area in 2009 recorded that its population was about 449,454 people consisting 229,653 people (51.09%) of males and 449,454 people (48.91%) of females. Apparently, male population was balance with female population. In addition, Rokan Hulu population density in 2009 was about 60 inhabitants per square kilometer (Centre Statistics Agency of Riau, 2010).

The annual rate of economic growth in Rokan Hulu was recorded at 6 %. Obviously, agricultural sector gave significant contribution for Gross Domestic Product in Rokan Hulu every year. In addition, transport and mining are also becoming fundamental and potential sectors for economic production in Rokan Hulu. The commodities of Rokan Hulu come from agriculture, fisheries, plantation, and

industrial sector. Oil palm plantations in Rokan Hulu have the most widespread area when compared to other districts in Riau Province. Based on the research conducted by the government of Riau Province, it explained that preserved forest in Rokan Hulu has the most extensive area, which is 67574.05 Ha or 29.53 % of the total area of all forests in Riau (Research and Development Agency of Riau Province, 2011).

1.4. Aim and Objectives

This research is aiming to analyze forest cover change by applying image processing of remotely sensed data in different time and model the future forest cover in order to monitor deforestation. Remotely sensed data were classified to derive forest cover as the main object of this research. Furthermore, future forest cover modeling was carried out by conducting some techniques that are used to model land use land cover change (LUCC).

In order to achieve the goal, some objectives of this research were built as follows:

1. To identify and quantify the forest cover in study area using remotely sensed data
2. To analyze forest cover change in study area within 1996-2005
3. To investigate and validate the future forest cover model using several techniques
4. To model future forest cover in 2015

1.5. Research Hypotheses

These hypotheses have been developed to deal with the research's objectives:

1. Remotely sensed data can perform forest cover change analysis
2. Forest cover has been decreased in the study area
3. Future forest cover can be predicted by using land use land cover change (LUCC) models

Based on these hypotheses, research questions can be drawn as follows:

1. How capable is remotely sensed data could perform forest cover change analysis?
2. What kind of techniques could be used to analyze forest cover change?
3. What is the trend of forest cover change in study area?

4. What kind of method could be applied to forecast forest cover change?
5. Which method could be appropriate to predict future forest cover?
6. Is the selected method suitable enough to perform prediction?
7. How will be the forest cover of study area in 2015 based on simulation?

1.6. Research Limitations

In order to carry out this research, remotely sensed data are the important raw data to achieve the goal, which is forest cover modeling. Remotely sensed data have been used to derive land cover in study area and determine the forest cover as the main object in this research. Multi-temporal Landsat imageries have been selected to generate land cover types because of the free availability to download these data provided by several institutions. Moreover, the spatial resolution of Landsat which is 30m is still reasonable to generate land cover types. However, the determination of land cover would be improved by using the higher spatial resolution of some imagery such as QuickBird with 60cm and IKONOS with 1m in spatial resolution. Landsat as an optical remote sensing is sensitively influenced by atmospheric condition so that it makes difficult to find data without cloud cover in study area, which is located in tropical area. To deal with areas covered by cloud, some secondary data were used to validate the existing land cover in the particular area such as reference map and image obtained from Google Earth. The other limitation that came up in this research was the use of reference map. Ideally, reference map as the tool to validate the land cover classification should coincide with the Landsat acquisition, but in this research reference map in 1997 was used instead of 1996 as Landsat acquisition was in 1996 and reference map 2005 is not available. Therefore, some adjustments using Google Earth image has been carried out.

1.7. Materials and Tools

Materials used in this research are remotely sensed data and land use maps in the study area for different time acquisition. Remotely sensed data occupied are Landsat TM and Landsat ETM+ because these data are available for free to download. For analyzing in 2005, Landsat TM acquired in 2005 was chosen instead of Landsat ETM+ because Landsat ETM+ acquired after May 2003 has been operating in SLC-Off mode. For further data used in this research are explained in table 1.1.

In this research, implementation incorporated spatial data in vector and raster types for analyzing. Vector data analyzing along with image processing has been done by using ArcGIS 10. In addition, to model forest cover change using 3 models which are Stochastic Markov, Cellular Automata Markov and GEOMOD Model, IDRISI Taiga has been employed as well as to forecast the future forest cover.

Table 1.1 Data used in the research

No.	Data	Acquisition	Format Data	Source
1.	Landsat 5 TM Path/Row 127/059	July 28th, 1996	Geotiff	http://www.glcf.umiacs.umd.edu/
2.	Landsat 7 ETM+ Path/Row 127/059	April 26th, 2000	Geotiff	http://www.glcf.umiacs.umd.edu/
3.	Landsat 5 TM Path/Row 127/059	June 19th, 2005	Geotiff	http://www.glcf.umiacs.umd.edu/
4.	Landuse Maps	1997 and 2000	jpeg	http://www.dephut.go.id
5.	Google Earth Image	unknown	jpeg	Google Earth

CHAPTER 2

CONCEPTS AND DEFINITIONS

2.1. Remotely Sensed Data Implementation for Forest Monitoring

The definition of remote sensing as stated by Lillesand and Kiefer (2000) is "the science and art of obtaining useful information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation". Remotely sensed data inform the coverage in land by the reflectance of objects in the ground, which are denoted as digital number. It implies that different object will perform different reflectance. Based on the reflectance difference, one can recognize varying land cover types.

Initially, forest cover maps were collected by doing terrestrial survey in the ground and reporting them as the country's forest cover to Food and Agriculture Organization of the United Nations (FAO) as the international agency which organizes forest cover data collection from over the world (Fuller, 2006). Apparently, the methodology to gather the data became unstandarized. In order to make it harmonized, remotely sensed data have been constituted internationally to monitor forest cover (DeFries et al., 2006). It was proposed to achieve the sustainable forest management by using remotely sensed data which have multi temporal availability to investigate frequently. Without using remotely sensed data, it would be difficult and not efficient to do survey in the field. Therefore, the developing of digital image processing has been rapidly increasing due to the importance of these data. Several methods for image recognition to derive land cover type can be employed and it is quite helpful to get better accuracy.

Implementation of remotely sensed data for forest monitoring has been widely carried out by several countries such as Brazil supported by National Institute for Space Research known as INPE with a broad study about Amazon, some countries in Central Africa and Southeast Asia as well (Fuller, 2006). The least requisite, Landsat which has 30m in spatial resolution has been suggested as the remote sensing data for this purpose (GOFC-GOLD, 2009). These data are recommended because of the high spatial resolution and the free accessibility in many sources such as NASA and

University of Maryland (DeFries et al., 2006). Although the higher spatial resolution with less than 5m will give better recognition, this data will be very costly.

Landsat has been evolved in term of its sensor since the first launch in 1972; they are Multispectral Scanner (MSS), Thematic Mapper (TM) and ETM+ (Enhanced Thematic Mapper Plus). Their capabilities are improved from one to another respectively. Landsat brings benefits due to the moderate spatial resolution and also the capability to capture a particular area periodically with a 16-day temporal resolution (NASA, 2008). The latest generation of Landsat, which is Landsat 7 ETM+ launched in 1999, has 8 bands and each of them has its characteristics in term of sensitivity to object's reflectance (as explained in table 2.1.).

Table 2.1 Characteristics of Landsat 7 ETM+

Band Number	Wavelength (μm)	Spatial Resolution	Band Characteristics
1	0.45-0.52	30	To investigate water body penetration, soil and vegetation determination
2	0.52-0.6		Reflects the healthy vegetation and extents within blue and red chlorophyll absorption bands
3	0.63-0.69		Very helpful to differentiate vegetation due to red chlorophyll absorption band and also to map soil and geological coverage
4	0.76-0.9		Correspond to vegetation biomass and well performed for vegetation types discrimination. In addition, it could identify soil moisture.
5	1.55-1.75		It is very sensitive to the amount of plant's water and benefits for drought studies. Moreover, it can differentiate cloud, snow and ice.
6	10.5-12.5	60	Corresponds to heat emitted by the surface. It is important for geothermal and soil moisture studies as well as plants stress analysis.
7	2.08-2.35	30	To differentiate geological formation.
8	0.52-0.9	15	Can be fused with other band to improve detection.

Source: Biradar et al. (2003)

2.2. Land Use and Land Cover Change (LUCC) Detection

Land use and land cover are commonly used to define the coverage and its activity of earth's surface. Knowing land use and land cover are very essential because this research needs to determine forest cover as the principal point to be analyzed and it is subtracted from land use and land cover information. Land cover is defined as the cover type of which lands that are being occupied, while land use is related to human activities in the particular land (Anderson et al., 1976). This paper also mentioned that remotely sensed data could not inform the human activity, thus local knowledge and ancillary data are necessary to formulate land use information in particular area. Another definition stated that land cover is the elements of surface in the earth and also the contained surface such as soil, the plant and animal, topography, surface and groundwater, and also structures created by human; and on the other hand, land use is defined as intention of human to utilize the land for some purposes (Lambin et al., 2003).

Several robust techniques have been developed to recognize land use and land cover from remotely sensed data. By these means, classification of land use and land cover for different time can perform change analysis. Land Use and Land Cover Change (LUCC) has become a major concern for international agencies such as Intergovernmental Panel on Climate Change (IPCC) because it influences the interchange of greenhouse gases between the environment and atmosphere (IPCC, 2003).

Detection of LUCC using remote sensing technology has been widely implemented for several researches. Pu, et al. (2008) applied Compact Airborne Spectrographic Imager (CASI) hyperspectral image to monitor saltcedar coverage. NDVI differencing and classification using Principal Component Analysis (PCA) have been performed to analyzed change detection. In this study, NDVI differencing performed better than classification technique. Moreover, the use of novel method in analyzing change/no change by counting threshold separately from decrease and increase part contributed high accuracy which was 93.04% comparing to PCA technique which contributed 91.56%. Another example, Hayes & Sader (2001) applied Landsat Thematic Mapper in three different acquisitions to monitor forest clearing and regrowth using NDVI differencing, PCA, and RGB-NDVI change detection by conducting radiometric normalization previously. In order to do accuracy

assessment, reference samples by visual interpretation of Landsat TM color composite have been conducted and proved that RGB-NDVI gained the highest overall accuracy around 85%.

2.3. Forest and Deforestation

Many definitions of forest are established based on classification categories in typical and ecosystem function, such as the forest as a natural forest, an artificial forest, forest conservation, a forest conversion, and a forest production. The definition of forest in this study is related to a land cover characteristic with trees as the domination of its vegetation type. FAO has defined forest as "land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agriculture or urban use" (Schoene et al., 2007). Mostly the canopy of natural forest in Indonesia reached this forest definition.

FAO's definition of the forest is similar to the definition of the United Nations Environment Programme (UNEP). UNEP defines "forest is a land area of more than 0.5 ha, with a tree canopy cover of more than 10 percent, which is not primarily under agriculture or other specific non-forest land use. In the case of young forest or regions where tree growth is climatically suppressed, the trees should be capable of reaching a height of 5 m in situ, and of meeting the canopy cover requirement" (Schoene et al., 2007).

Meanwhile, deforestation process associates with logging activities in the forest cover and land use change in forest that takes place permanently for a variety of other commercial purposes. According to the definition of deforestation generated by the FAO, deforestation is "the conversion of forest to another land use or the long-term reduction of the tree canopy cover below the minimum 10 percent threshold" and additionally, UNFCC defines deforestation as "the direct human-induced conversion of forested land to non-forested land" (Schoene et al., 2007). Deforestation is not related to logging activities in forest temporally and controllability, because the trees in the forest from its activities may grow back or replanted. It is correlated to only after the land is permanently converted to non-forest functions, such as industry, settlement, and other commercial activities (IPCC, 2000). Therefore, from remotely sensed data used in this research with moderate resolution,

deforestation has been determined by interpretation based on land cover change from a forest area to a non-forest area in Rokan Hulu District, Riau Province, Indonesia.

2.4. Land Use and Land Cover Change Modeling

The existences of land use land cover change (LUCC) models are very huge for implementation with different assumptions related to the particular study (Pontius & Malanson, 2005). Some of the models are Conversion Land Use and its Effect (CLUE); Slope, Land use, Exclusion, Urban extent, Transportation, and Hill shade (SLEUTH); GEOMOD; Cellular Automata Markov; Stochastic Markov; Land Use Scanner; and Land Transformation Model (LTM) (Pontius et al., 2008; Koomen & Borsboom-van Beurden, 2011). Land use change modeling is important to recognize the pattern along with process of changing and could bring benefit for policymakers to understand the chance of future conditions with different scenarios (Koomen et al., 2007). Models are implemented with digital map from beginning event to create transition for forecasting what will happen in the future in the same area (Pontius et al., 2008).

In addition, land use change model comes up with transformation and allocation, transformation begins with recent land use and conducts the transition into future forest cover by taking into account the transformation probability of neighboring land use, while allocation deals with allocating the possible location of particular land use considering its characteristics (Koomen et al., 2007). Basically, calibration and validation in LUCC modeling have been considered separately as two different processes that should be carried out. Calibration is performed by determining the factors used for model; these factors should be based on information derived from some points in time (t1) or before which extrapolation starts (Pontius & Malanson, 2005). For instance, Cellular Automata Markov (CA_Markov) model uses two maps from two different time, time0 and time1, to predict time2. Meanwhile, validation is conducted in order to compare how well the prediction of time2 from reference map in time2, reference map is trusted as the map showed the recent land cover.

Several researches regarding land use land cover change modeling have been done with employing different kind of models. Pontius and Malanson (2005) compared two models, CA Markov and GEOMOD to predict land change in central

Massachusetts, USA. It incorporated land use change from 1971 to 1985 to predict land use in 1999 and found that GEOMOD gave more benefit to model. Echeverria et al. (2009) performed GEOMOD and spatially explicit linier regression to detect the drivers of deforestation between 1976 and 1999 and predict the deforestation in 2010 and 2020 in southern Chile. Another example related to forecast deforestation has been done by Rashmi and Lele (2010), it applied GEOMOD by combining road and settlement as the drivers along with forest cover maps in 1973, 1992, and 2000 to predict deforestation in 2015.

CHAPTER 3

LAND COVER CHANGE DETECTION

The following sections are explaining step by step in deriving land cover change detection. Initially, land cover should be generated from remotely sensed data. In order to do that, image pre-processing need to be conducted then all processing related to remotely sensed data could be performed. In this research, land cover derivation has been performed by applying supervised classification using some training samples. Apart from that, NDVI differencing was carried out to detect land cover change. The result of NDVI differencing which is image showing change and no change has been combined with classified image in order to generate what kind of land cover has been change whether increase or decrease in particular area. The flow chart below describes the way to determine land cover change.

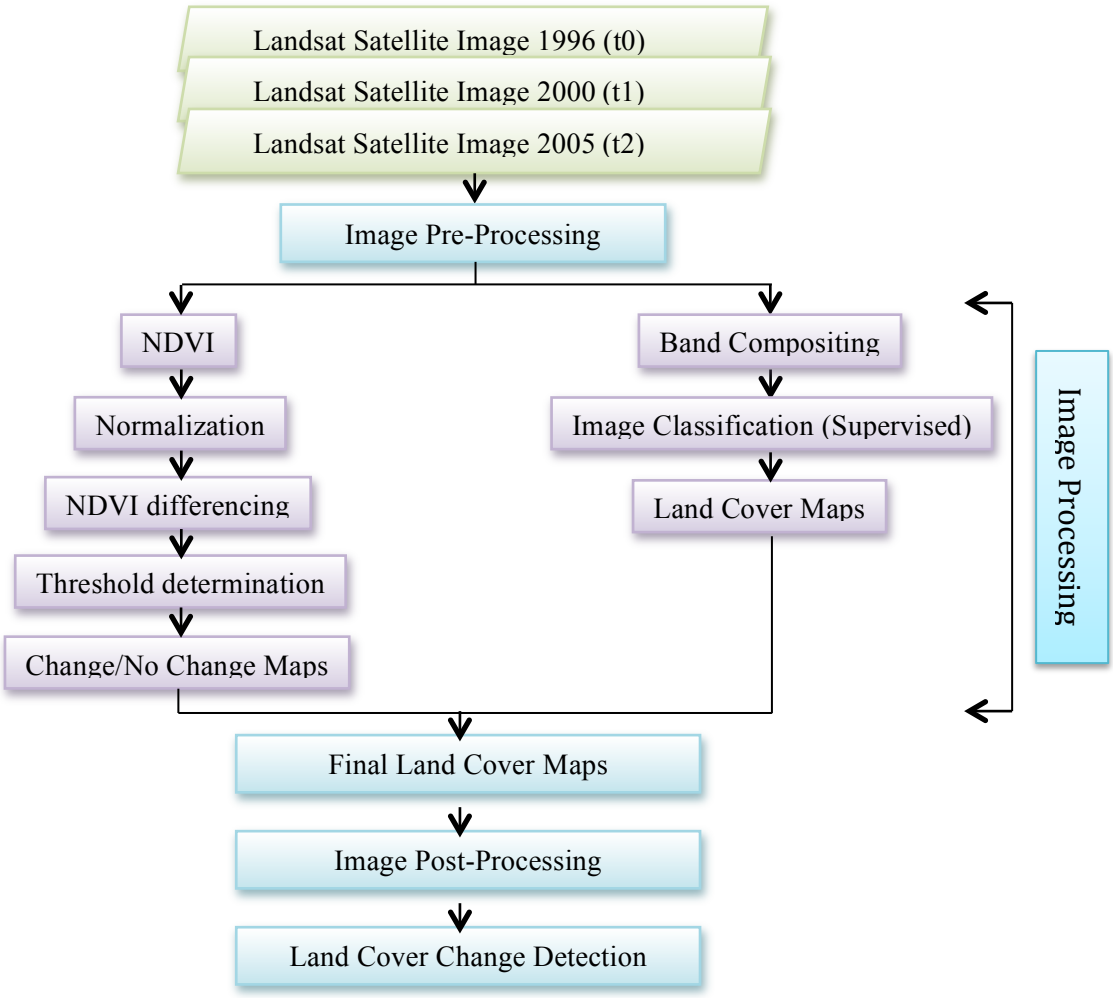


Figure 3.1 Flow chart of land cover change detection

3.1. Spatial Data Preparation

3.1.1. Image Pre-Processing

Initially, pre-processing of remotely sensed data should be done before doing further analysis. Pre-processing would be performed in regard to geometric and radiometric correction. Geometric correction aims at correcting the position error due to earth rotation, whereas radiometric correction aims at correcting the error due to atmospheric effect (Lillesand & Kiefer, 2000). In fact, recent remote sensing data have been corrected in the ground station before making them available to the users. As we can find in many sources providing Landsat data, each of Landsat imagery was attributed with different level showing the degree of correction such as Level 1R, 1G and 1T. Landsat data acquired in this research have been corrected in Level 1T for 1996 image as well as 2005 image and Level 1G for 2000 image. Level 1G expresses systematic geometric and radiometric correction from data acquired by the sensor. On the other hand, Level 1T illustrates systematic geometric and radiometric correction involving Digital Elevation Model (DEM) and ground control point for topographic accuracy (USGS, 2010).

Another pre-processing that needs to be prepared is band compositing. It is done by combining three bands as a color composite band with respect to spectral reflectance signature to make easier in visual interpretation. Compositions are organized by incorporating three bands and visualizing them as Red, Green, and Blue (RGB). Band Composition could be presented either as False Color Composite (FCC) or True Color Composite (TCC). True Color Composite is achieved by combining band 3, 2 and 1 as it performs the true color of some land covers as found in the ground. As the purpose of this research is to analyze forest cover, false color composite of band 543 was chosen because this combination is very advantageous for vegetation studies (Quinn, 2001) and very helpful to distinguish forest cover from other coverage. Practically, band 5 as Middle Infrared (MIR) is specified into Red gun; band 4 as Near Infrared (NIR) is specified into Green gun; and band 3 as Red is specified into Blue gun. Interpretation will be considered to spectral reflectance signature. Figure 3.1 shows band composites as TCC and FCC of study area in 1996.

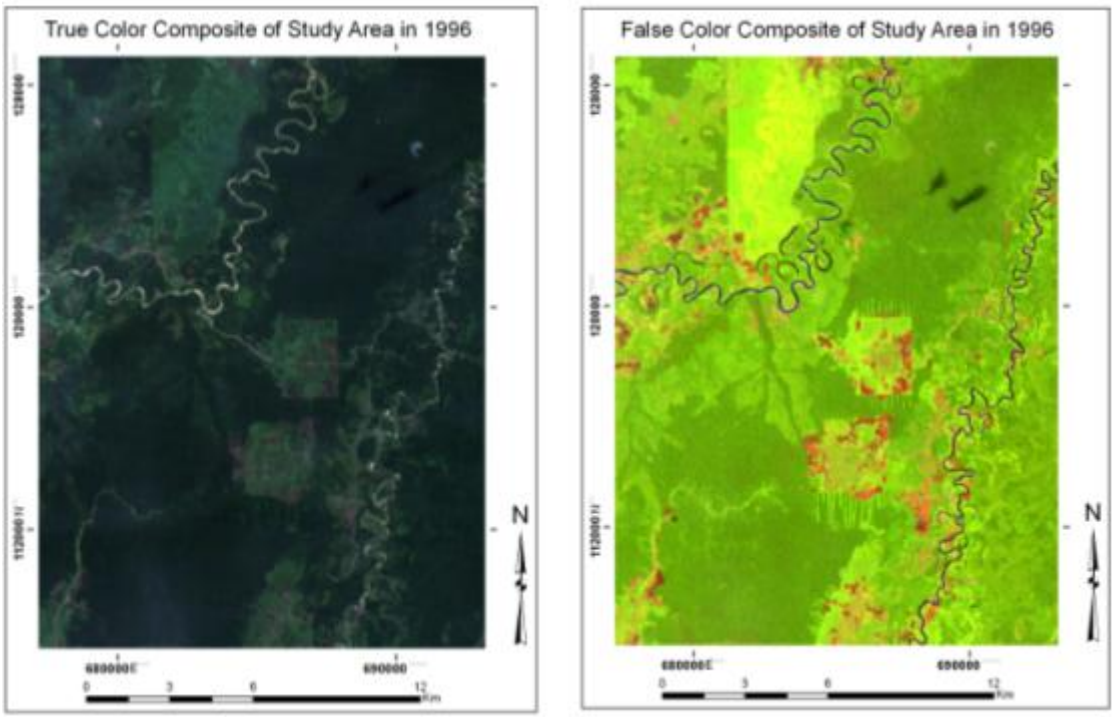


Figure 3.2 Landsat TM Composite of study area in 1996: (a) Composite of band 3, 2, 1 and (b) Composite of band 5, 4, 3

In the composite of band 5, 4, 3 as we can see above, red color expresses soil because MIR specified into red gun gives higher reflectance for soil. Then, green color expresses vegetation because NIR specified in green gun gives a higher reflectance for vegetation. Eventually, the rest band indicates water surface.

3.1.2. Image Processing

Image processing has been carried out to determine object of the research in order to differentiate forest cover among other covers along with to detect the change. Determination has been conducted by combining Normalized Difference Vegetation Index (NDVI) differencing and image classification. The further explanation will be described in the following sections.

3.1.2.1. Normalized Difference Vegetation Index (NDVI) differencing

NDVI differencing to analyze change and no change of several land cover types occupies NDVI (Pu et al., 2008). NDVI is vegetation index performing the greenness of particular area by incorporating near infrared band and red band because of the vegetation characteristics, which absorb visible light for photosynthetic process and reflect near infrared light (NASA, n.d.). The difference of spectral reflectance

between red and near infrared could identify the greenness easily (ESRI, 2011). NDVI is considerably influenced by phenology so that healthy vegetation will contribute high value of NDVI because of the high relectance in near infrared light, and comparatively low reflectance in red light (CCRS, n.d.). NDVI is formulated as follows :

$$NDVI = (NIR - R)/(NIR + R)$$

This range of NDVI's value is between -1 and +1. A zero value indicates that there is no vegetation (ERDAS, 2009).

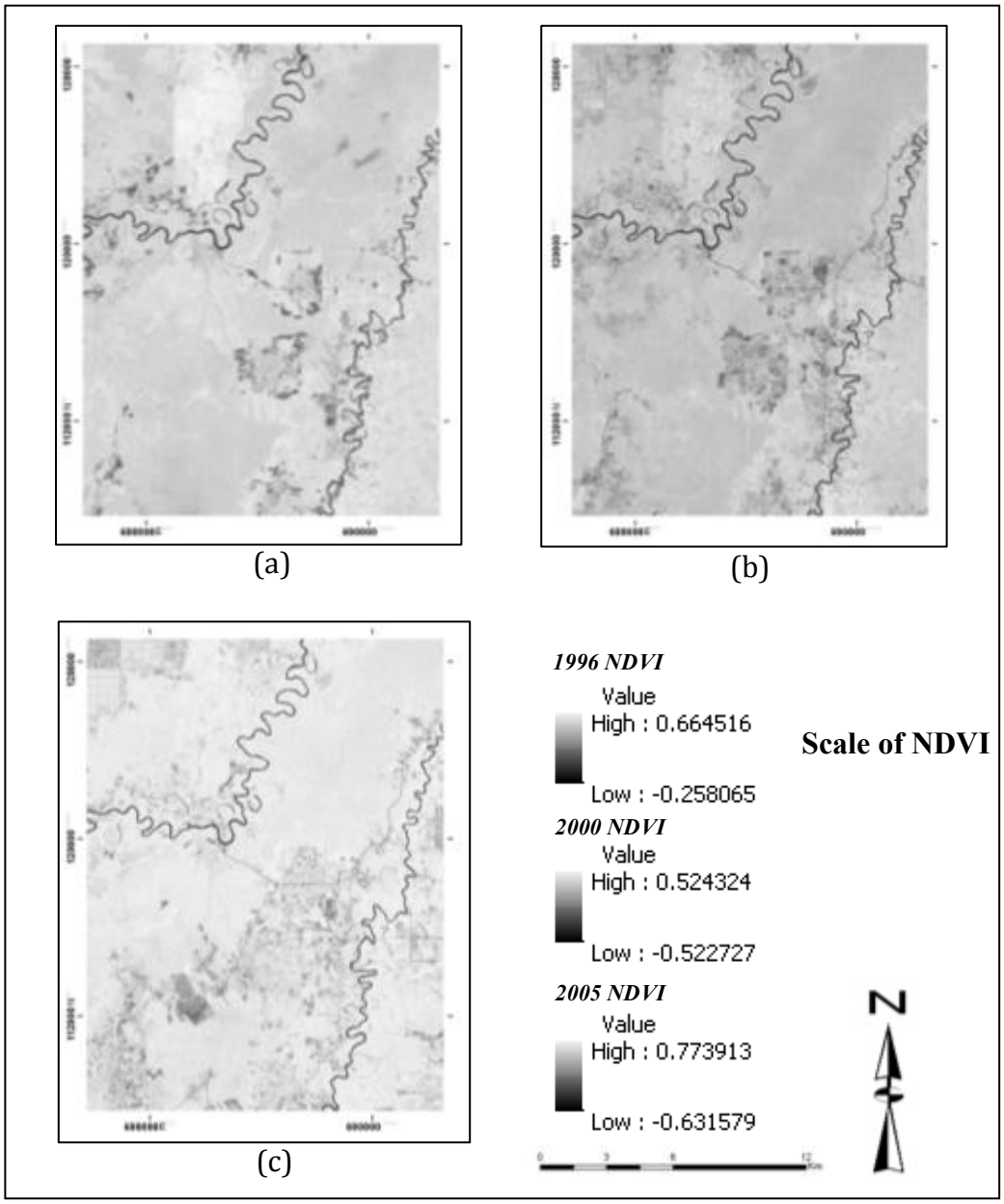


Figure 3.3 NDVI images of (a) 1996, (b) 2000 and (c) 2005

Moreover, low value (less than or equal to 0.1) indicates bare land, rock, sand and snow; value between 0.2 and 0.3 indicates shrub and grass; high value from 0.6 to 0.8 indicates dense vegetation (ESRI, 2011). The use of NDVI in vegetation studies has been widely implemented and achieved high accuracy. By considering the pair NDVI images, NDVI differencing has been conducted. However, due to the difference of acquisition from one image to another image, normalization one to another image should be done otherwise it will influence the radiance differences (Pu et al., 2008). It can be caused by atmospheric condition in different acquisition. This kind of difference will bring bias in extracting change and no change information. Instead of informing the true land cover change it will inform the wrong one because it formulates the difference pixel, which actually does not indicate the true land cover. In this case, the same land cover is presented with quite different pixel and causes the difference.

Such kind of normalization is done in order to normalize radiometric value into similar level for both of images acquired in different time (Pu et al., 2008). Pu et al., (2008) suggested applying normalization between two NDVI images using linear function instead of applying normalization into the original bands (the red and NIR band) then performing NDVI images. Both NDVI images are assumed to have linear function. Then, $y = ax + b$ is formulated in which x and y are digital numbers of NDVI image X and Y, while a and b are coefficients which can be obtained from the digital numbers as a number of samples derived from unchanged land cover. In this research, 50 samples have been subtracted between two NDVI images (as described in Appendix A). As the results, three linear regression equations have been achieved as follows,

$y = 1.0057\text{x} - 0.2621, \text{r}^2 = 0.87$ for NDVI image pair 2000(Y) – 1996(X)…(1)

$y = 1.4389\text{x} - 0.1304, \text{r}^2 = 0.82$ for NDVI image pair 2005(Y) – 1996(X)…(2)

$y = 1.3053\text{x} + 0.2999, \text{r}^2 = 0.90$ for NDVI image pair 2005(Y) – 2000(X)…(3)

From this equation, normalization has been performed. These three equations were used to predict NDVI image of 2000 from NDVI image of 1996 using equation (1). The predictive NDVI image of 2005 can be estimated either from NDVI image of 1996 using equation (2) or NDVI image of 2000 using equation (3). After applying the equation of normalization above, histograms of predictive NDVI can be extracted.

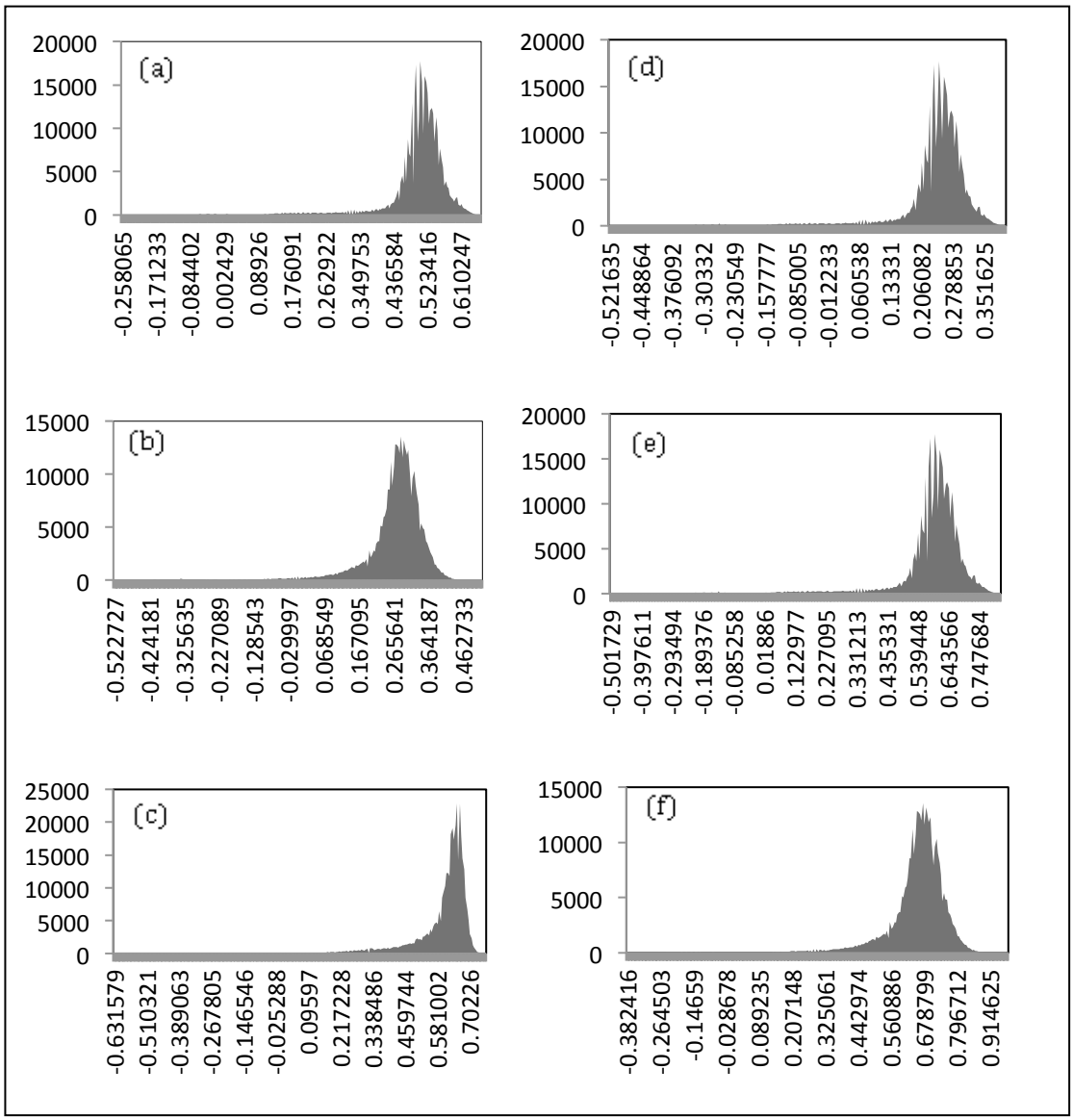


Figure 3.4 Histograms of NDVI images. The histograms were extracted from (a) 1996, (b) 2000, and (c) 2005 NDVI images and predicted NDVI histograms for (d) 2000 NDVI (Y) predicted from 1996 NDVI (X), (e) 2005 NDVI (Y) predicted from 1996 NDVI (X), and (f) 2005 NDVI (Y) predicted from 2000 NDVI (X)

From the figures above, we can notice that the histograms of predictive NDVI images of 2000 (d) and 2005 (e) have similar histogram of 1996 NDVI image because they are actually representing the distribution of coverage in 1996 which have been normalized corresponding to 2000 NDVI and 2005 NDVI, respectively. Another histogram which is predictive NDVI image of 2005 (f) has similar shape with histogram of 2000 NDVI image because it represents the distribution in 2000 which was normalized corresponding to 2005 NDVI image. Having these predictive images, the NDVI differencing could be achieved. The NDVI differencing has been done by subtracting the predicted NDVI from the original NDVI. For instance, to analyze the

change between 2000 and 1996, subtracting the predicted NDVI of 2000 (by 1996 NDVI) from actual NDVI of 2000 should be done. It was continued to derive another NDVI differencing. NDVI differencing between 2005 and 1996 was generated by subtracting predicted NDVI of 2005 (by 1996 NDVI) from the actual 2005 NDVI. The last one, NDVI differencing between 2005 and 2000 was generated by subtracting predicted NDVI of 2005 (by 2000 NDVI) from the actual 2005 NDVI.

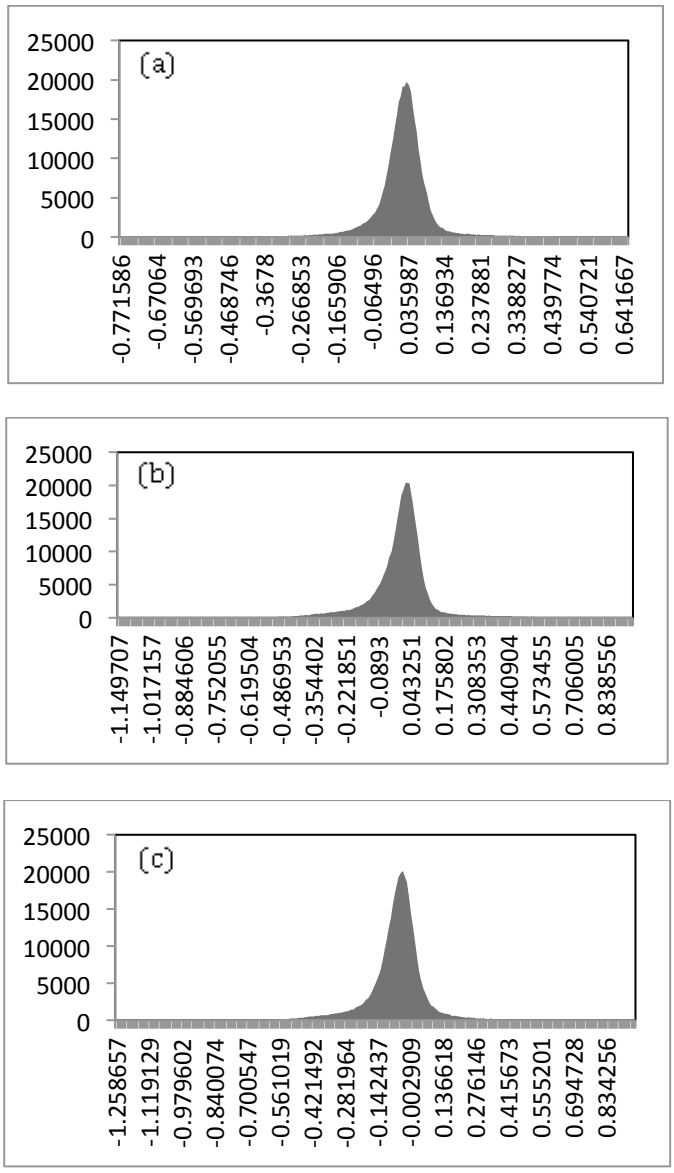
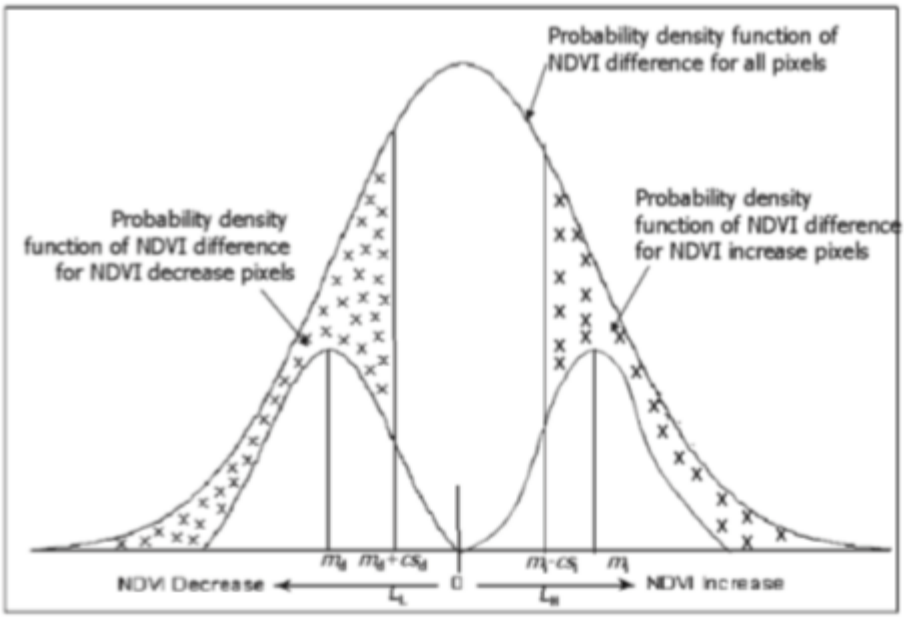


Figure 3.5 Histograms of NDVI differencing between (a) 1996 and 2000, (b) 1996 and 2005, (c) 2000 and 2005

The final results of NDVI differencing are representing the level of difference in which there is a part that does not identify a change, as well as a part that identifies a change. In order to recognize land cover change, they need to be processed further

to get the information of change and no change. Hence, threshold based on mean and standard deviation (SD) showing the limit of change and no change should be determined. In this research, assigning the threshold has been defined by dividing the histogram of NDVI differencing into two parts, decrease part and increase part, then mean and standard deviation of respective part were involved in assigning the threshold as shown in figure 3.6. From the figure, L_L and L_H show the threshold of the low and high threshold; and d and i indicate the decrease and increase, respectively. The change due to decrement will be reached when some particular pixels are less than $m_d + CS_d$ and the change due to increment will be reached when some particular pixels are greater than $m_i - CS_i$. Afterwards, the other pixels located between them $m_d + CS_d \leq x \leq m_i - CS_i$ imply no change. C itself is the coefficient of change/no-change, which can be assigned by choosing the optimal c through the highest kappa or overall accuracy index which performs the best change and no change condition as reference data (Pu et al., 2008). C is defined by a set of coefficient (from -1.8 to 0.4) as shown in figure 3.7. By calculating the threshold with different c for both $m_d + CS_d$ and $m_i - CS_i$, then several maps with different threshold have been created.

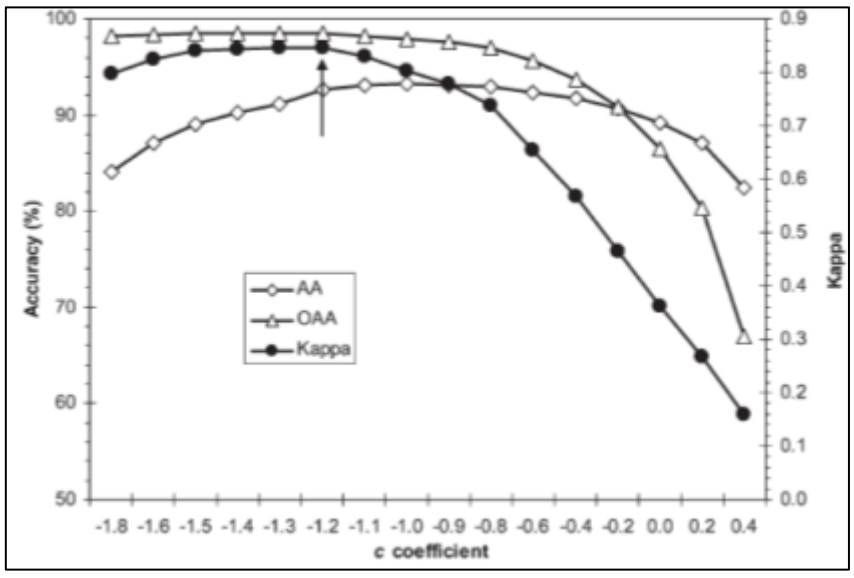


Source: Pu et al. (2008)

Figure 3.6 The dissemination of probability density functions within NDVI differencing

After producing several maps with different c, one image that shows the best change and no change in study area should be selected. It is supposed to compare

these maps using ground data to measure the accuracy either using Kappa or overall accuracy index (Appendix B explains about these terminologies) in order to define the optimal c that can show the best change and no change image. However, due to the lack of ground data in this research, optimal c has been selected by considering the best performance in showing change and no change within study area in respect to land cover map obtained by supervised classification. The tendency that has been found from several maps in different c is the higher the c value, the more noise can be found; and on the other hand, the smaller the c value, the more information has been lost. After having investigation, it was defined the optimal c value and thresholds of decrease and increase for different NDVI differencing as described in table 3.1. The change between 1996 and 2000 is decreasing when the pixels less than 0.0891 and increasing when the pixels greater than 0.1313 while pixels between -0.0891 and 0.1313 indicate no change. It is also working with the same way for analyzing the change between 1996 and 2000 and from 2000 to 2005. The change and no change as the result of NDVI differencing are showing the change of overall land cover types in study area. Consequently, combining the change map and classified image needs to be done in order to get the information of land cover change detection



Source: Pu et al. (2008)

Figure 3.7 The selection of optimal c value

Table 3.1 Thresholds of decrease and increase for different NDVI differencing

NDVI Differencing	NDVI Decrease		NDVI Increase		Optimal C Value	Threshold	
	Mean	SD	Mean	SD		Decrease	Increase
1996 - 2000	-0.0254	0.0637	0.0735	0.0578	-1	-0.0891	0.1313
1996 - 2005	-0.0888	0.1088	0.0756	0.0866	-0.9	-0.1867	0.1535
2000 - 2005	-0.1409	0.1045	0.0155	0.0723	-1	-0.2454	0.0878

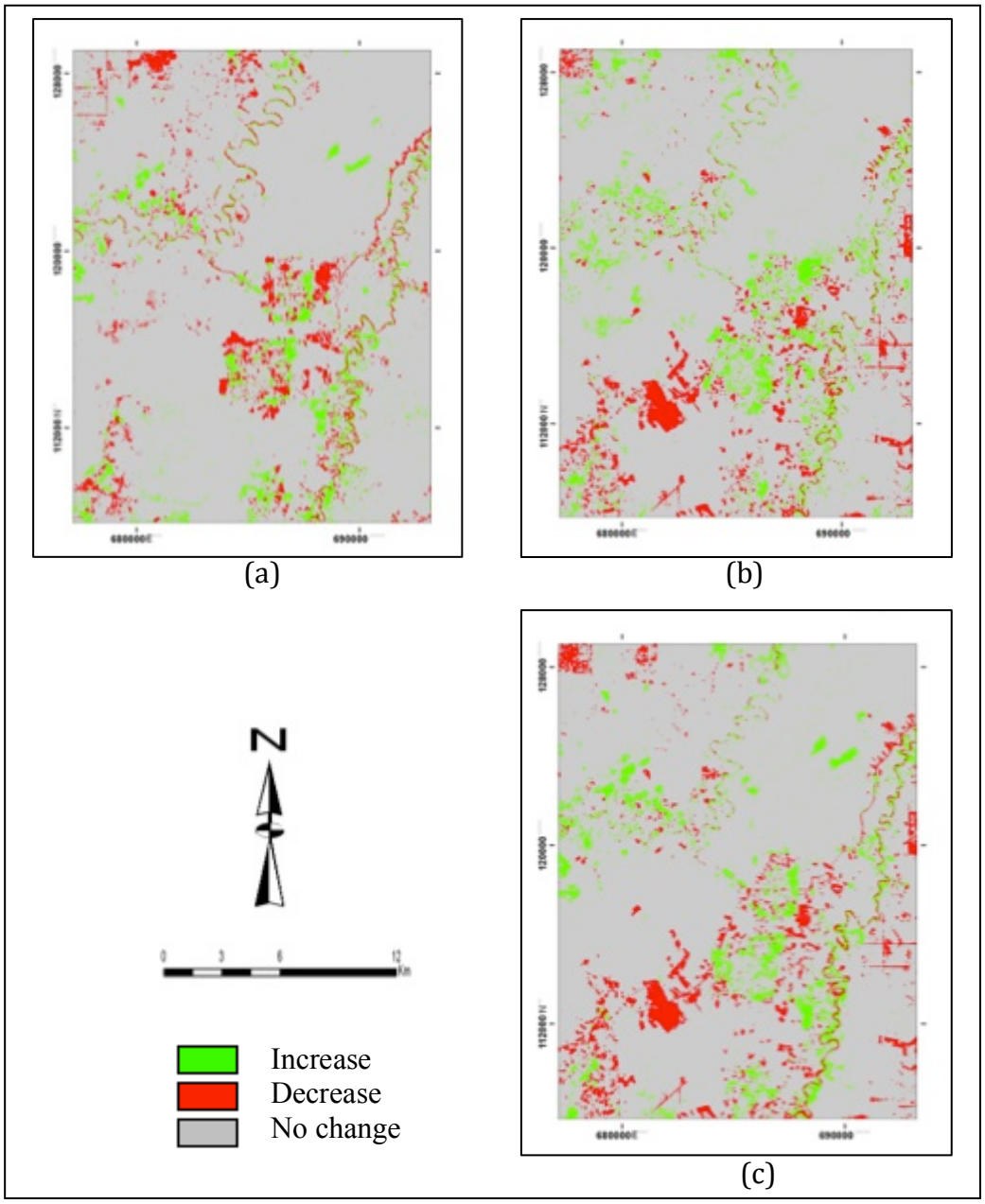


Figure 3.8 Change and no change images between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005

3.1.2.2. Image Classification

Image classification employs digital image processing and classifies the groups of pixels based on spectral information into particular themes (CCRS, n.d.). Basically, there are two kinds of classifications, supervised and unsupervised classification. Supervised classification considers training samples to assign the class of pixels group and unsupervised does not consider (CCRS, n.d.). In this research, supervised classification has been chosen. Training samples were defined by assistance of some ancillary data such as land use map and image from Google Earth in the study area. Google Earth image is selected because of its higher spatial resolution. After checking, the image provided by Google Earth in the study area has a spatial resolution of 15m. These ancillary data are quite helpful to assign the coverage based on spectral recognition because it ensures the type of class in particular area regarding the existing land use in the field.

Table 3.2 Land cover types in study area

Land Cover Type	Description
Forest	Areas occupied by natural forest, which are categorized as tropical rainforest.
Cropland	Areas used for agricultural activities.
Shrubs	Areas covered by bushes or shrubs.
Water bodies	Comprise river and ponds.
Fallow land	Comprises open space, bare land and cropland that is not planted.

In this research, training samples that have been used to generate land cover in 1996 are 13samples of forest, 15samples of cropland, 8samples of shrubs, 14samples of water bodies, 11samples of fallow land, 2samples of shadow, and 2samples of cloud cover; training samples in 2000 are 16samples of forest, 21samples of cropland, 14samples of shrubs, 13samples of water bodies, and 9samples of fallow land; and the last training samples in 2005 are 16samples of forest, 17samples of cropland, 10samples of shrubs, 13samples of water bodies, and 9samples of fallow land. After creating those signatures, maximum likelihood classifier was chosen to generate the land cover type for the whole image. Maximum likelihood assigns a pixel grouped in

a particular class based on the probability (ERDAS, 2009). This classification also takes into account the variance and covariance of the signature file for each class to determine the pixel belongs to a particular class (ESRI, n.d.). Based on the result of classification, there are five categories of land cover types that have been identified in this research (as explained in table 3.2).

The final land cover map was derived by combining the classified map and change map in order to improve the categorization due to the lack of ground data. The base map used is classified map of 1996 as the result of supervised classification then land cover map in 2000 was derived by combining the base map of 1996 in which NDVI differencing between 1996 and 2000 does not identify any change and supervised map of 2000 in which NDVI differencing identifies a change. Using similar approach, land cover in 2005 was derived by combining the improved map of 2000 in which NDVI differencing between 2000 and 2005 does not identify any change and supervised map of 2005 in which NDVI differencing identifies a change.

These final land cover maps have been obtained by implementing generalization. Generalization process in the remote sensing analysis and mapping technique is related to the determination of minimum mapping unit (MMU). The MMU is defined as the smallest area that can be mapped from the transformation process in a geometric analysis of satellite imagery (Knight & Lunetta, 2003). CORINE land cover map in 2000 applied MMU as 25 ha (Gómez & Milego, 2005; Pekkarinen, Reithmaier, & Strobl, 2009). In addition, Brazil PRODES was using 6.25 ha but now it uses 1 ha as MMU, India national forest monitoring uses 1 ha, and GMES forest monitoring uses 0.5 ha (GOFC-GOLD, 2009). Knight and Lunetta (2003) stated that possible MMU for land cover map generated from Landsat is 8100 m^2 (0.81 ha) or 90 x 90 m (3 x 3 pixels). Hence, 1 ha has been applied for generalization in this research where all the units smaller than 1 ha will be merged into the neighboring land cover.

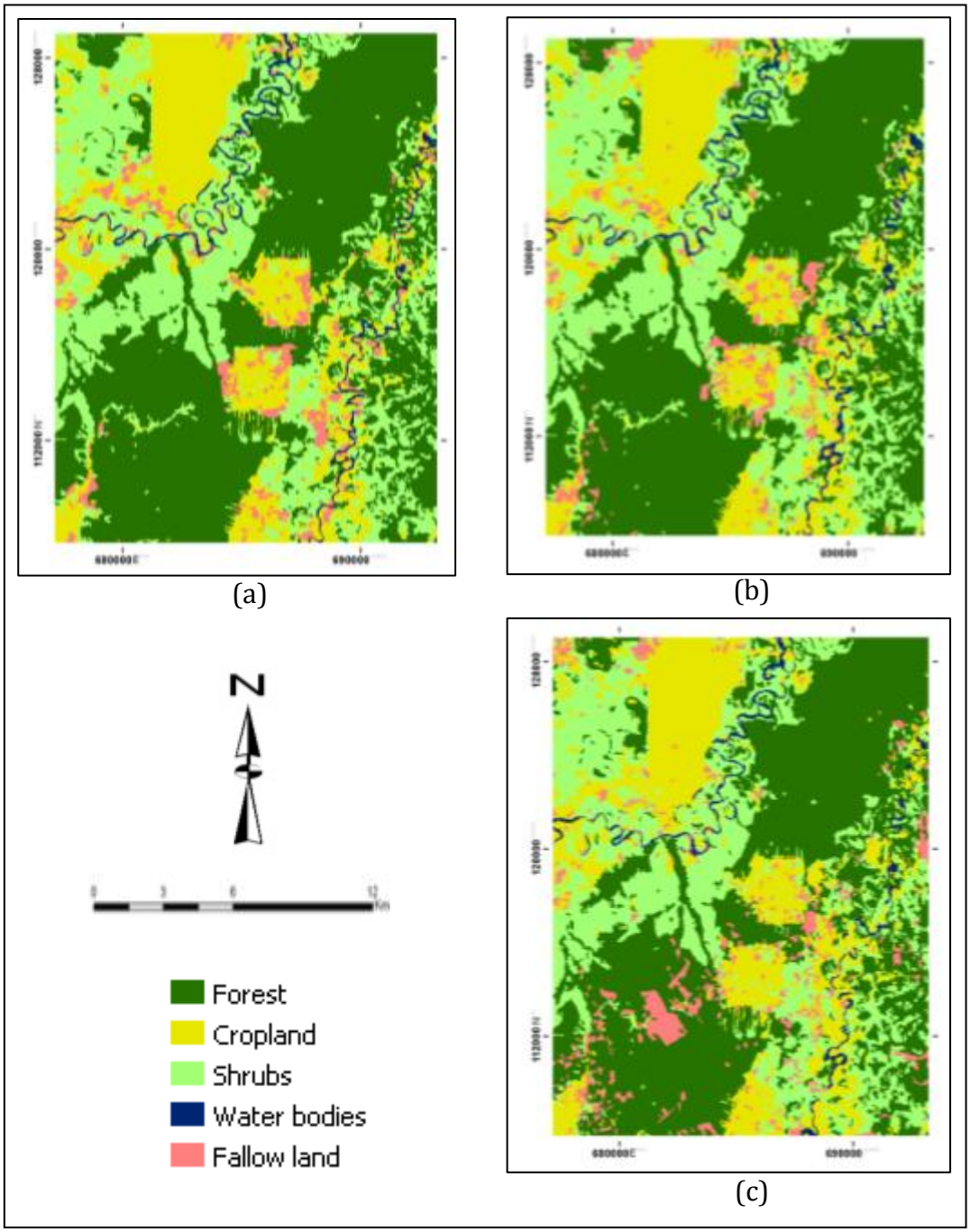


Figure 3.9 Land cover types in study area (a) 1996, (b) 2000 and (c) 2005

3.1.3. Image Post-Processing

The last step, which is post-processing, has been done by doing accuracy assessment. Accuracy assessment performs the number of likelihood some categories that are mapped comparing into ground data (reference data), it is important to define the level of error provided by land cover map for further analysis (Eastman, 2009). Some datasets representing the categories in map and ground data should be chosen to perform the accuracy assessment. Performing accuracy assessment are commonly presented by using error matrix, which is a square array presenting column and row

for identifying a number of sample sets (such as pixels or polygons) within classified categories and reference data, column represents the number of categories as reference data and row represents the number of categories as classified data (Congalton, 1991). In other word, error matrix is also known as confusion matrix describing omission error (producer's accuracy) and commission errors (user's accuracy) (Rashid, Lone, & Romshoo, 2011). Meanwhile, the overall accuracy can be calculated from the total correct of categories (the sum of major diagonal in error matrix) divided by total number of pixels used as sample data; whereas omission error measures the percentage of particular class in reference map that has been correctly identified and commission error measures the percentage of the correct pixels corresponds to the total pixels that has been classified (Congalton, 1991). Omission error is derived by dividing the correct number of particular class by the total number of pixels for that class in reference data, whereas commission error is derived by dividing the correct pixel in particular class by total number of pixels that have been classified into that class.

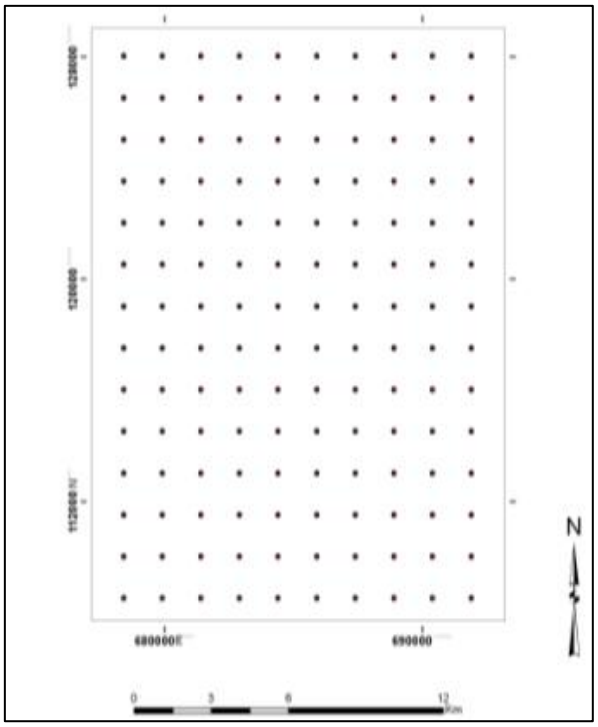


Figure 3.10 Distribution of training samples for accuracy assessment

In this research 140 samples have been carried out to assess the accuracy for classified image in 1996, 2000 and 2005. These samples have been assigned systematically around the study area as shown in figure 3.10. In order to conduct the accuracy assessment, Google Earth image, color composite Landsat, and reference

land cover maps have been used. Some adjustment have been performed to assign the existing land cover by considering spatial reflectance of color composite Landsat to Google Earth image and some information from reference maps although this reference maps were not so helpful because they have slightly different nomenclature from classified image conducted in this research. As the results, classified image in 1996 attained accuracy assessment around 82%, 76% for classified image in 2000, and 74% for classified image in 2005. The detailed accuracy assessment can be seen in Appendix C.

3.2. Land Cover Change Detection Analysis

Land cover change analysis was performed by considering the change among land cover types in different time. Having classified images that have been generated as explained in section above, area of each land cover types were investigated to understand the behavior of land cover change as shown in table 3.3. Each of land cover types has its own trend of changing as shown in figure 3.11. It reveals that human impact has been contributing in changing of land cover.

Table 3.3 The area of each land cover

Land cover	Area (Ha)		
	1996	2000	2005
Forest	15669.63	15501.51	14701.50
Cropland	6744.87	6717.06	7270.83
Shrubs	9708.03	9867.96	9692.82
Water bodies	621.54	692.64	578.97
Fallow land	1445.40	1410.30	1945.35

Based on the observation, it indicates that forest cover tends to decrease from the initial year to the latest year. There are three possibilities to analyze deforestation, which are between 1996-2000, 2000-2005, and 1996-2005. It is obvious that deforestation between 1996-2000 reached 1%, 2000-2005 reached 5%, and overall between 1996-2005 reached 6%. The other types such as cropland and fallow land note a significantly increasing trend especially in 2005, it indicates that probably there were clearing activities in this study area. In this case, human activities contribute a big influence in converting the land. It is also proved based on interpretation that

apparently fallow land was found around cropland. It indicates that this fallow land could be a temporary fallow land, which was being unplanted due to rotation system. The last land cover types, shrubs and water bodies indicate a slight change from time to time.

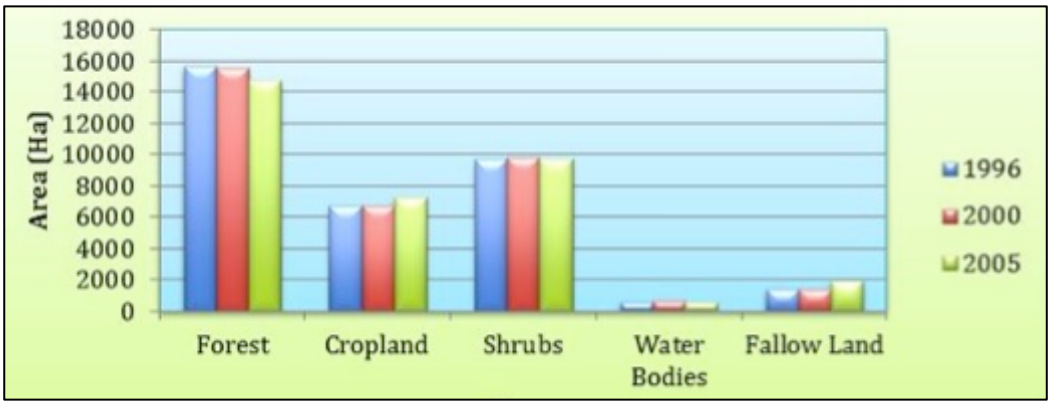


Figure 3.11 Land cover change in study area

As focus of this research is forest cover, it was performed in figure 3.12 showing the different proportions of loss, gain, and persistence of forest cover compare to the total area. During this period, the loss of forest (deforestation) is higher than the gain of forest (reforestation). In period 2000-1996, the loss of forest reached about 1% and the gain reached 0.5%; in period 2005-2000, the loss reached about 3% and gained about 1%; in period 2005-1996, the loss reached 4% and gained about 1%. It implies that between 1996 until 2000 there was just a slight change but starting from 2000 to 2005, a lot of disturbance caused high deforestation.

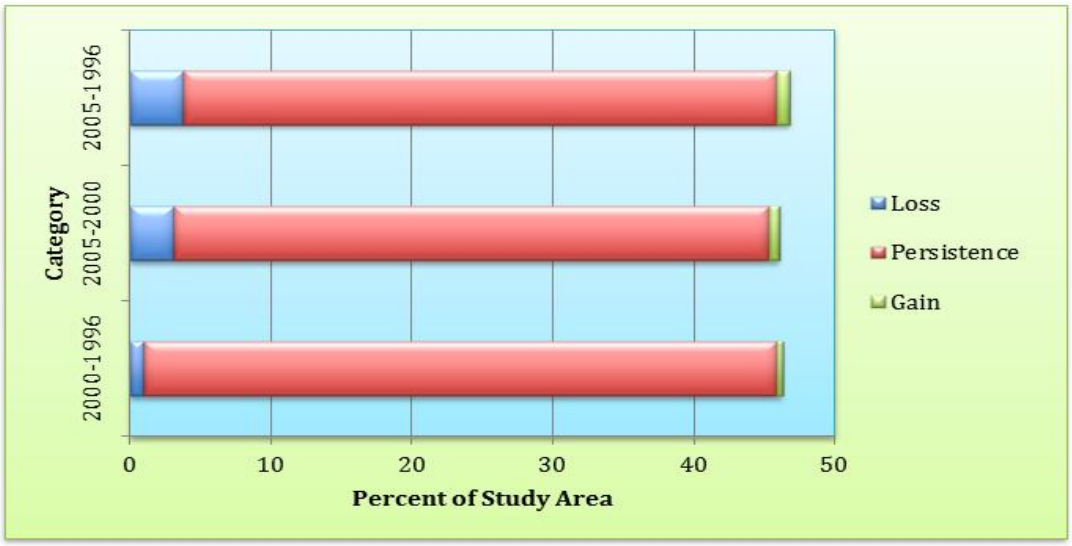


Figure 3.12 Proportion of forest cover change in the study area

The spatial distributions of forest cover change during the time period in 1996-2000, 2000-2005, and 1996-2005 are shown in figure 3.13. It is obvious that the loss of forest between 1996 and 2000 is the least compare to other image pairs. The two image pairs show similar pattern due to the influence of a significant change in 2005. As we can see in the land cover map of 2005, we can find a lot of forested area in southwest part were opened which probably would be cropland. In this sense, human activities have a big influence in forest cover change.

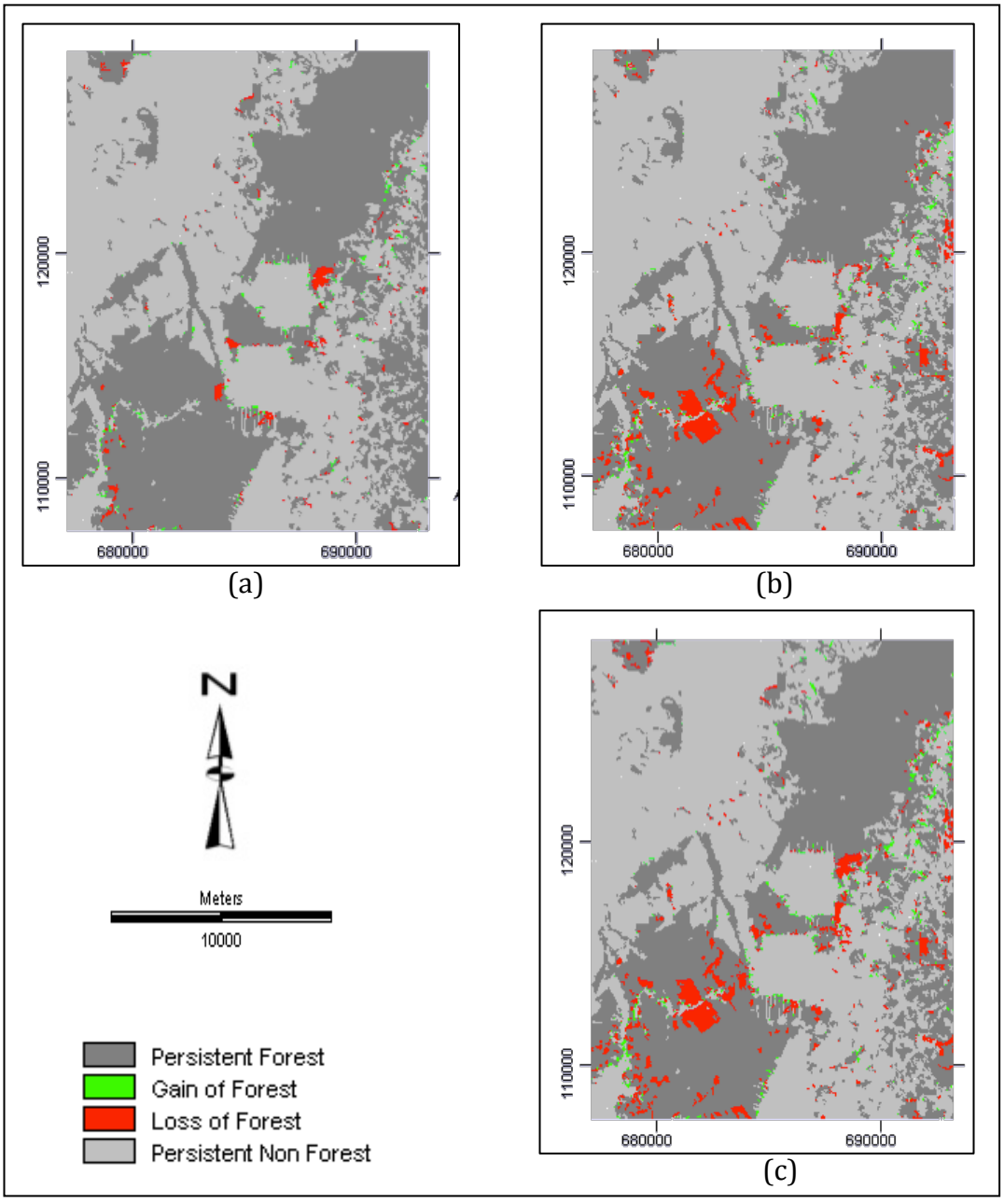


Figure 3.13 Forest cover change images between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005

CHAPTER 4
FOREST COVER MODELING

Forest Cover Modeling has been done by doing separately calibration and validation using three models and then validating them in order to compare the capability of three models in predicting the future. Validation has been done by employing Kappa index. Eventually, the best model that showed the highest Kappa was chosen to forecast the future forest cover in 2015. The detailed explanation will be described in the following sections.

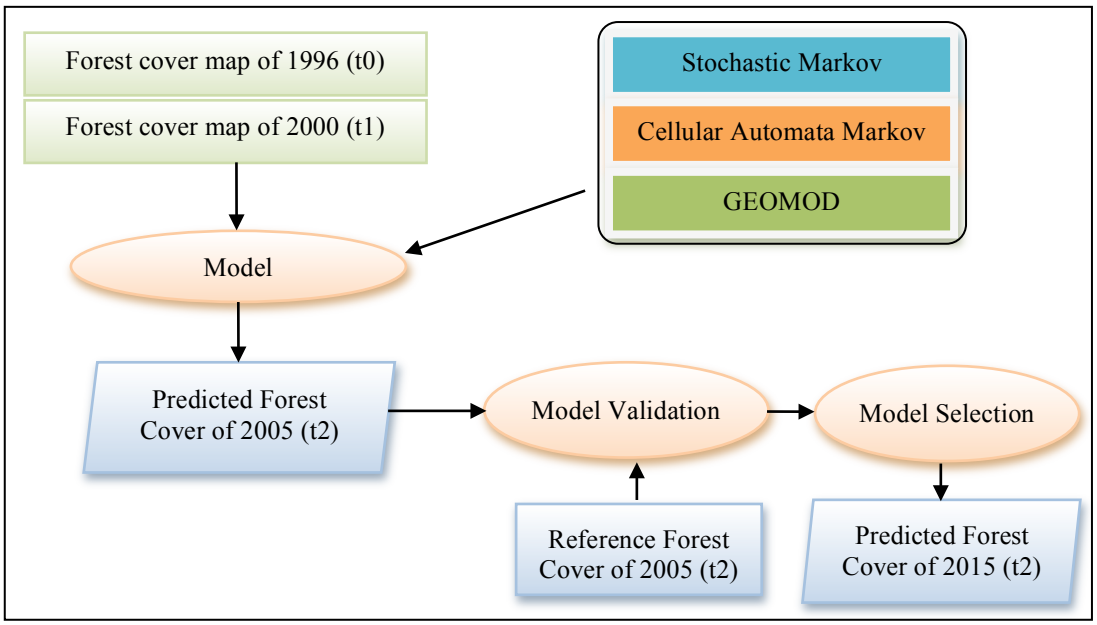


Figure 4.1 Flow chart to model forest cover change

4.1. Calibration

Calibration of some factors that are incorporated for models has been performed using three models; they are Stochastic Markov Model, Cellular Automata Model and GEOMOD. In order to focus on object of the research, land cover maps have been simplified into forest and non-forest. The following sections will explain the details of each model.

4.1.1. Stochastic Markov

In this process, markovian model has been applied to investigate land cover change in the present time (t1) and the time before (t0) to predict land cover in the future (t2). The details of markov process and how the model works are performed in the sections below consecutively.

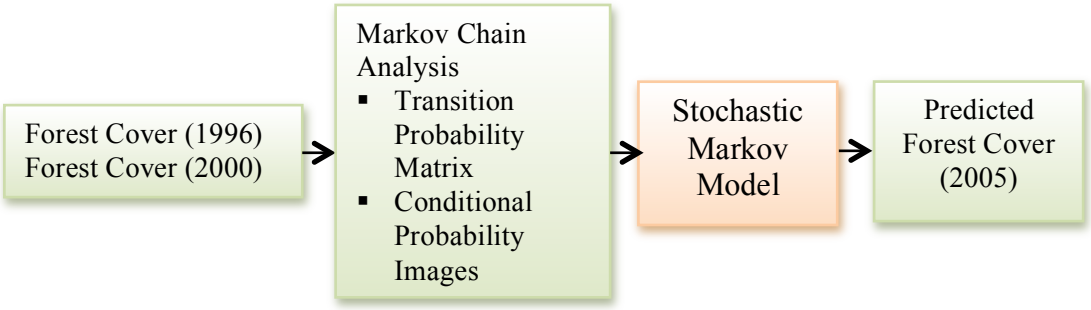


Figure 4.2 Flow chart of Stochastic Markov Model

4.1.1.1. Stochastic Markov Process

Markov model that works as a chain is known as markov chain. Markov chain as a stochastic process used in this model incorporates every single category as the state of a chain (Weng, 2002). Stochastic process is a probability process (Doob, 1942), while markov chain is stated as stochastic process which deals with markov property using discrete state space and a discrete or continuous parameter space which signifies the time (Balzter, 2000). This process is characterized by the property in which the value of the process in time t relies on its value in time t-1 and not on the series in time t-2, t-3, ..., t-n (Weng, 2002).

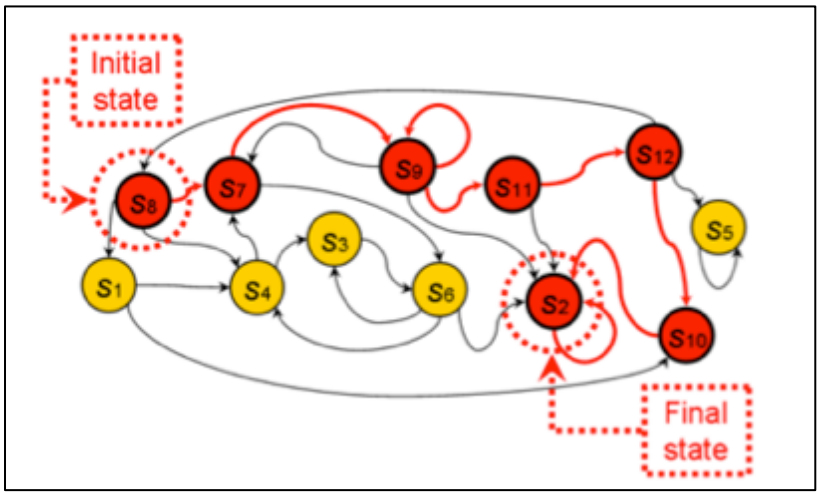


Figure 4.3 Example of Markov Chain

Figure above shows the schematic transition diagram of markov chain and describes one possible path. It assumes the initial state in time X_0 is s_8 and the final state is s_2. States are shown as red circle, while possible transition among the states is shown as red arrows. Time step can be denoted as Xn and state is denoted as S. For instance, in the figure above, $X_3 = S_9$ represents that the system shows state S_9 at time n=3. The system begins at time 0 (X_0) and move from one state to another with a

certain probability which is commonly called as transition probability. In this case, stochastic works based on present state (Izquierdo et al., 2009).

According to Balzter (2000), the conditional probabilities are denoted as $P\ (X_t = j|X_s = i) = \ p_{ij}(s,t)$ which indicate the transition probabilities from state i to j for indices $0 \leq s\ \leq t, with\ 1 \leq i,j\ \leq k$.

Then, the transition matrix P of k states is denoted as:

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{1k} \\ P_{21} & P_{22} & P_{2k} \\ \dots & \dots & \dots \\ P_{k1} & P_{k2} & P_{kk} \end{bmatrix} \dots (1)$$

For the initial state at time 0, it is denoted as $P\ (X_0 = i) = \ p_i(0)\ \forall i \in \{1, \dots, k\}$. Additionally, the probabilities of state at time t, which is $p_i(t)$ at time t, are calculated for consecutive number of the *k* states as follows:

$$\underline{p}\ (t) = (p_1(t), p_2(t), \dots, p_k(t)) \quad \dots (2)$$

Eventually, the equation to predict the future probabilities could be formulated using:

$$\underline{p}\ (t) = \underline{p}(t-1).P\ \dots (3)$$

As the function of time series, Markov chain will achieve a constant probability vector named limiting distribution.

$$\underline{p}\ (\infty) = \lim_{t \to \infty^-} \underline{p}\ (t)\ = \lim_{t \to \infty^-} \underline{p}\ (0)\ .P^t \dots (4)$$

4.1.1.2. Stochastic Markov Model

Stochastic markov model has been carried out using IDRISI Taiga software. Before applying stochastic module, markov chain analysis was performed between forest cover in 1996 and 2000 to predict forest cover in 2005. This process produced a transition probability matrix (table 4.1), a transition area matrix (table 4.2), and conditional probability images (figure 4.4). Transition probability matrix reports the probability of land cover to change to another category, whereas transition areas matrix reports the area as the number of pixels which are supposed to alter from one category to another category in respect to specified number of time. The last result of markov analysis which is probability images records the possible category that can be found. These images are projected from the two previous images. Having these

products from markov analysis, stochastic analysis was performed by using conditional probability images as an input.

Table 4.1 Transition Probability Matrix between Forest and Non Forest

		2000	
		Forest	**Non Forest**
1996	**Forest**	0.828	0.172
	Non Forest	0.159	0.841

Table 4.2 Transition Area Matrix between Forest and Non Forest

		2000	
		Forest	**Non Forest**
1996	**Forest**	142,553	29,686
	Non Forest	33,077	174,567

Transition area is represented as a number of pixels, which has the original unit of image. It is 30 x 30 meters. Figure 4.4 shows the conditional probability images; and consecutively, predictive image in 2005 as a result of Stochastic Markov Model is shown in figure 4.5. The results of this model shows salt and pepper appearance because there is a lack of spatial distribution knowledge among the categories. Hence, the weakness of stochastic markov is a lack of spatial dependency (Eastman, 2009).

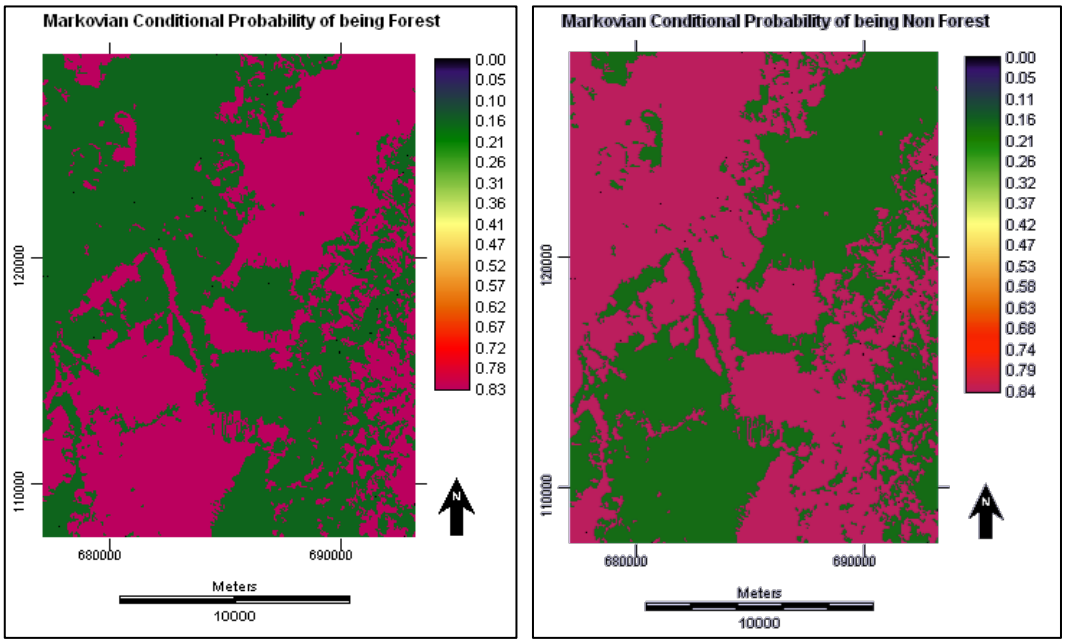


Figure 4.4 Conditional Probability Images

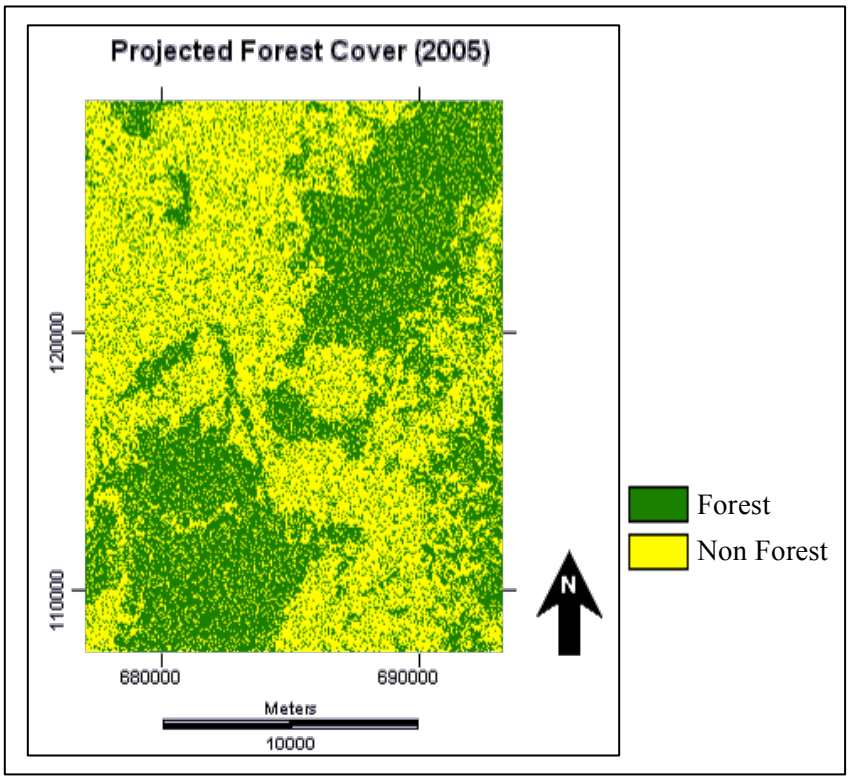


Figure 4.5 Projected Forest cover in 2005 using Stochastic Markov Model

4.1.2. Cellular Automata Markov (CA_Markov)

The second model, which is CA_Markov model, also occupies markov chain analysis. Additionally, it includes Cellular Automata concept in order to give sense of spatial dependency.

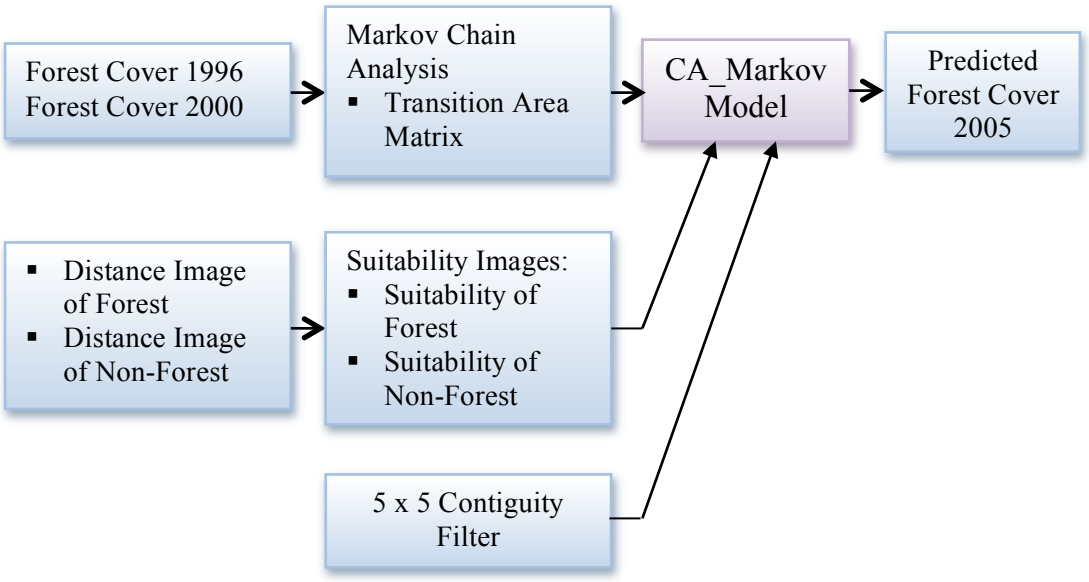


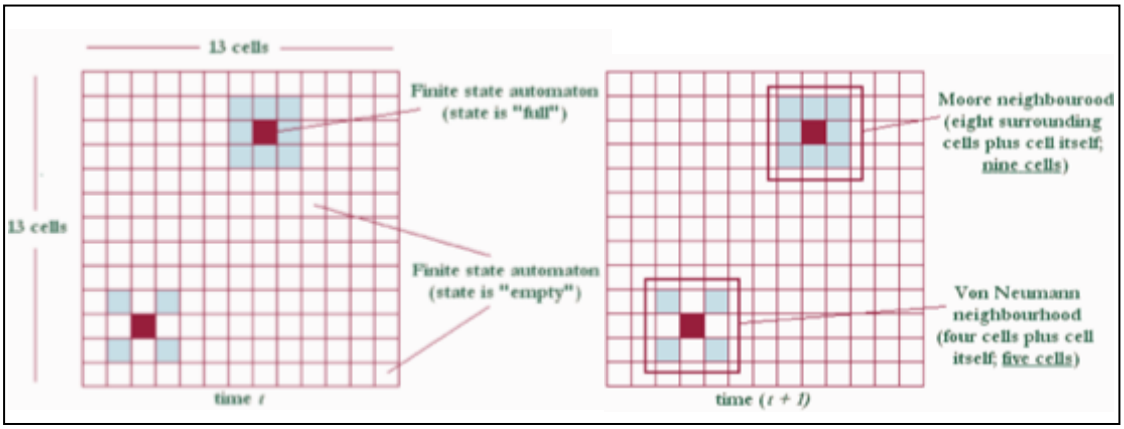
Figure 4.6 Flow chart of Cellular Automata Markov Model

4.1.2.1 Cellular Automata Markov (CA_Markov) Process

Cellular automata are defined as "a class of spatially and temporally discrete mathematical systems characterized by local interaction and synchronous dynamical evolution" (Ilhacinski, 2001). Another definition, "a cellular automaton is an agent or object that has the ability to change its state based upon the application of a rule that relates the new state to its previous state and those of its neighbors" (Eastman, 2009b). A mathematician named John von Neumann in 1950s introduced this concept initially for biological self-reproduction modeling (Ilhacinski, 2001). A cellular automaton involves several following components (Torrens, 2000):

1. a lattice (space where CA exists) of cells
2. a cell comprises a finite set of state
3. the neighborhood nearby the automaton
4. a set of transition rule that defines the performance of each automaton
5. the temporal space of automaton

Theoretically, the lattice of CA could be performed in one-dimension up to *n*-dimensional but for geographic schemes, it is considered in two-dimensions (Torrens, 2000). The most common two-dimension neighborhoods are Moore and Von Neumann as shown below.



Source: Torrens (2000)

Figure 4.7 Two-dimensional cellular automata

Mathematically, Cellular Automata is denoted as follows (Maerivoet & Moor, 2005):

$CA = (\mathcal{L}, \Sigma, \mathcal{N}, \delta) \ldots (1)$

Where,

$\mathcal{L}$ is a discrete lattice representing the physical environment

Σ is set of possible states in which every ith cell of the lattice at time t

the physical environment is represented by the discrete lattice $\mathcal{L}$ and the set of possible states denoted by Σ. Each $i\ th$ cell of the lattice has at time step t has a state $\sigma_i(t) \in \Sigma$.

$\mathcal{N}$ is the neighborhood of a certain cell which is denoted as $\mathcal{N}_i(t)$

δ is the local transition rule denoted as

$$\delta : \Sigma^{|\mathcal{N}|} \longrightarrow \Sigma : \bigcup_{j \in \mathcal{N}_i(t)} \sigma_j\ (t) \longmapsto \sigma_i\ (t+1)\ \ldots (2)$$

Based on the equation (2), the state of $i\ th$ cell at $t\ +\ 1$ is estimated using δ based upon the states of neighboring cells at time t.

Because the local transition rule is determined by a rule table, the output state will vary depending on the neighboring configuration. Assumed the amounts of $\Sigma\ and\ \mathcal{N}$, the possible rules equals $\left|\Sigma^{\Sigma^{\mathcal{N}}}\right| \ldots\ (3)$

Respecting to all the states for the whole cells at time t, a CAs global could b configured as

$$C(t) = \bigcup_{j \in \mathcal{L}} \sigma_j\ (t)\ \ldots (4)$$

in which $C\ (t)\ \in\ \Sigma^{\mathcal{L}}$

Occasionally, $C(t)$ is formulated using polynomial as

$$C(t) = \sum_{j=0}^{|\mathcal{L}|} \sigma_j\ (t)\ x^j\ \ldots (5)$$

By applying the local transition rule for the whole cells in the lattice, then CA could be denoted as

$$G : \Sigma^{\mathcal{L}} \longrightarrow \Sigma^{\mathcal{L}} : C(t) \longmapsto C(t+1)\ \ldots (6)$$

Because CA considers time and space, it will be iterated from initial configuration at $t = 0$ of $C(0)$ to the subsequent configurations as

$$C(0) \longrightarrow \mathrm{G}\,(\mathrm{C}(0)) \longrightarrow \mathrm{G}^2\,(\mathrm{C}(0)) \longrightarrow \mathrm{G}^3\,(\mathrm{C}(0)) \longrightarrow \cdots (7)$$

4.1.2.2 Cellular Automata Markov (CA_Markov) Model

Cellular Automata (CA) Markov used to model forest cover change employs markov chain to quantify the transition probability of land cover change in a certain time which is combined with CA to predict spatially explicit change between the land cover (Mondal & Southworth, 2010). Cellular Automata could simulate the transition from a number of categories to other categories, unlike Geomod which could only simulate for two categories for instance to simulate first category to the second category and not in the other way around (Pontius & Malanson, 2005).

CA_Markov predicts quantity of every category in time 2 based on extrapolation of gain and loss for every category from time 1 in which the extrapolation of quantity is a function of transition matrix generated by Markov chain between time 0 and time 1 (Pontius & Malanson, 2005). However, cellular automata has been improved not only depends on the former state but also neighboring state (Eastman, 2009). In this case, CA_Markov considers the spatial dependency. In fact, there are two considerations to predict location of land cover change, they are suitability image and contiguity rule (Pontius & Malanson, 2005).

Suitability image is used in CA_Markov to extrapolate the transition in defining the category; it chooses the particular category based on the largest suitability. Moreover, Pontius & Malanson (2005) stated that "CA_Markov uses a suitability image for each transition that it extrapolates", it means that suitability images incorporating the driving factors are built based on some categories contained in land cover map. In this research, land cover map used to predict forest cover comprises forest and non-forest. Consequently, suitability images were built from these two maps. These suitability images were only employed for CA_Markov and Geomod model since Stochastic Markov does not allow suitability images to model the land cover change. Before conducting suitability images for both factors, Boolean images for each factor have been performed. For instance, Boolean images of forest cover categorize forest as 1 and non-forest as 0.

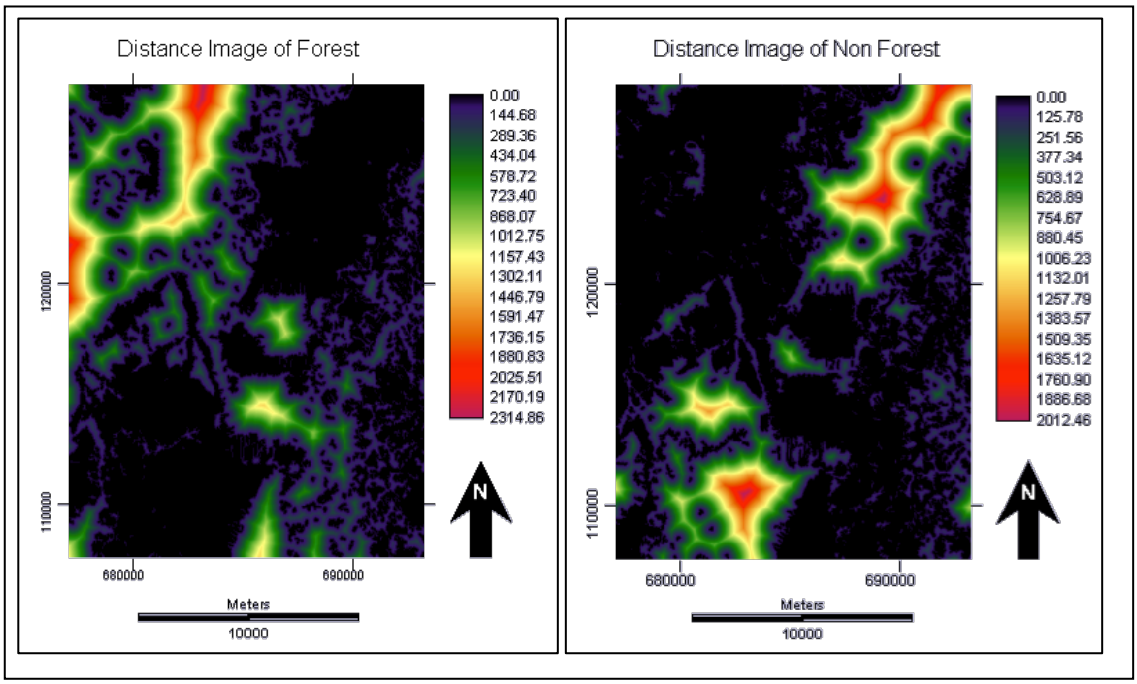


Figure 4.8 Distance Images of each land cover in 2000 (unit: meters)

After preparing 2 Boolean images, distance images were created for each of them which show the distance of each pixel to a certain category. Afterwards, suitability images were performed based on these distance images; it aims at standardizing the suitability images in a same unit measurement. Suitability images for all of them were adjusted in the same unit from 0 to 255. The highest value which is 255 shows that this particular area has the highest suitability; and on the other hand, 0 indicates that this area has the lowest suitability image. Furthermore, as input to predict the location of change, contiguity rule has also been performed. It is related to filter that is chosen as regard for neighboring assessment. When there are only few of neighboring pixels belong to certain category, then the suitability value is not applied; on the other hand, when there are many of neighboring pixels belong to certain category, the suitability value is applied (Pontius & Malanson, 2005). Eventually, the predictive model has been performed using CA_Markov module provided in IDRISI Taiga by inputting transition area matrix, suitability images, and defining the contiguity filter. In this study filter 5 by 5 is used as contiguity filter.

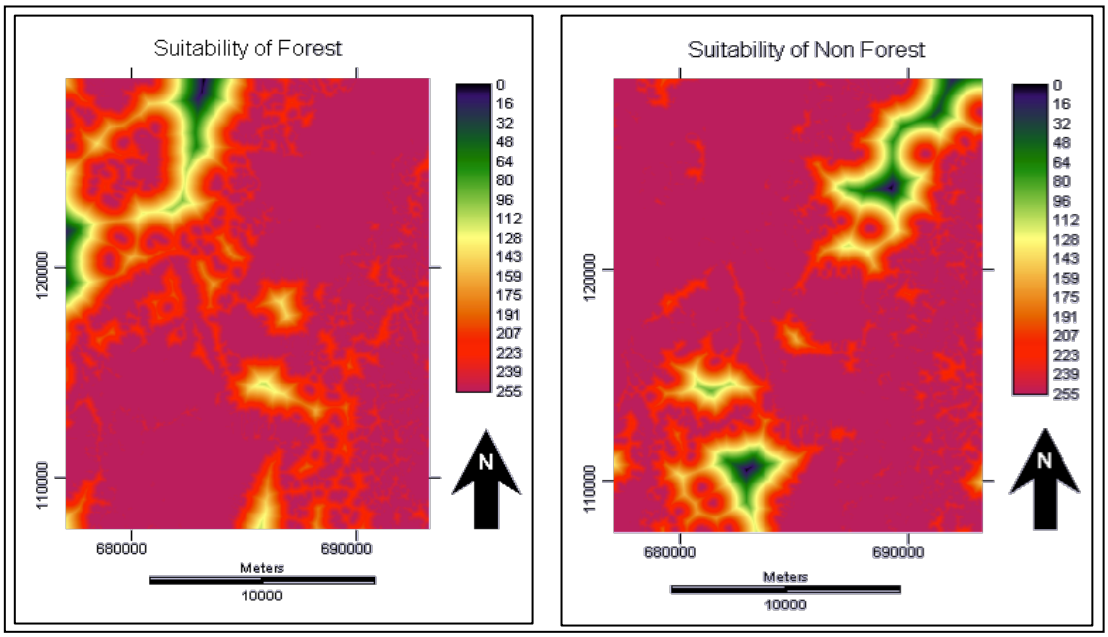


Figure 4.9 Suitability Images of each land cover in 2000

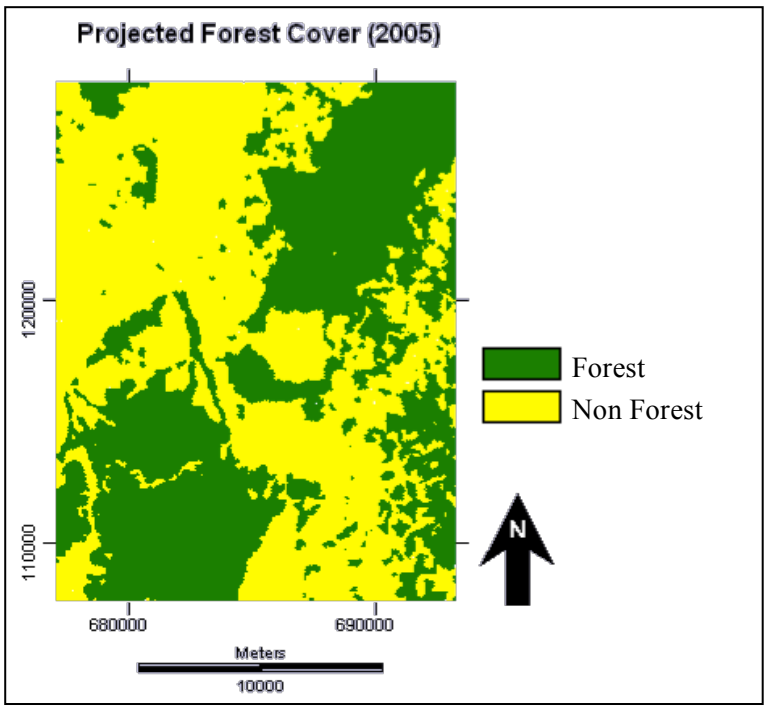


Figure 4.10 Projected Forest cover in 2005 using CA_Markov Model

4.1.3. GEOMOD

GEOMOD has been performed for several researches in order to determine the baseline scenario of deforestation for carbon budgeting as the concern of international agencies in climate change issue (Pontius & Chen, 2006). The difference of this model compare to two models before is this model does not incorporate markov chain analysis.

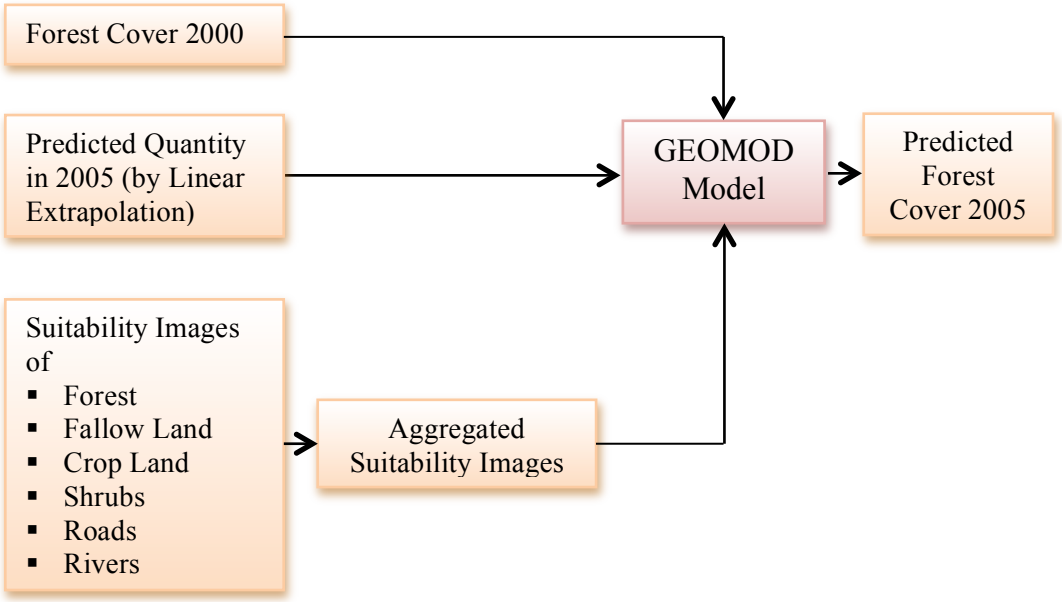


Figure 4.11 Flow chart of GEOMOD

In addition, GEOMOD only works to simulate either gain or loss from particular category to another category, it simplifies the complexity of CA_Markov in term of the capability to simulate the transition of several categories (Pontius & Malanson, 2005). To predict the quantity of change, Pontius & Malanson (2005) suggested linear extrapolation. Thus, to calculate the quantity in time 2, we need to specify the number of such category from time 0 and time 1. In this research the quantity of forest in 2005 were predicted using linear category based on the amount in 1996 and 2000 (see the detailed explanation in Appendix D).

Table 4.3 Factor Weights for the Suitability Images

Driving Factors	Factor Weight
Suitability Image of Forest	0.1667
Suitability Image of Fallow land	0.1667
Suitability Image of Cropland	0.1667
Suitability Image of Shrubs	0.1667
Suitability Image of Roads	0.1667
Suitability Image of Rivers	0.1667

Pontius & Chen (2006) explained that there are 4 fundamentals to predict the location of change. The first one is GEOMOD only allows one-way transition, from non-developed to developed or vice versa as an example. If user considers the developed one then GEOMOD will simulates from non-developed to developed and other cells will remain and there is no change from developed to non-developed. The

second one is stratification capability of GEOMOD; it can stratify the analysis based on strata. In this sense is political unit, quantity of change and calibration of suitability image can be perform separately for each stratum. The last two fundamentals are similar to CA_Markov with considering suitability image and contiguity rule. Suitability images can be created using Multi Criteria Evaluation or by using GEOMOD itself with inputting the driving factors. Basically, suitability images are created as same as the one that created to model CA_Markov. These suitability images are derived from several distance images. The last one, contiguity rule is assigned by defining the window for neighboring analysis.

In this research, GEOMOD has been performed by using forest cover map in 2000 as the basic map to predict forest cover map in 2005 and used several driving factor maps then aggregated them using Multi Criteria Evaluation (MCE). Initially, some distance images have been prepared before making suitability images. Distance image was applied for every single Boolean map. Boolean map represents the considered category to be simulated. For example, to create Boolean map of forest area, the category which is non-forest is assigned as 0 and the other hand, forest area is assigned as 1. Some driving factors have been selected regarding some references that have been done with similar research. The first factor is distance to forest because it indicates that the area closed to forest is more vulnerable to be changed into non forest. Since based on this research that forest tends to change into fallow land, shrubs and cropland so these land cover types were also taken into account due to the large amount of conversion forest into this types as also stated by Pérez-Vega (2011) who modeled land cover change in deciduous tropical forest that some human activities influence forest conversion. The other driving factors such as distance to roads and distance to rivers were also considered, as also used in similar research which analyzed forest loss between 1976 and 2020 in southern Chile carried out by Echeverria et al. (2009), these factors were considered because they make accessible log transportation.

Subsequently, all the distance images were used to produce suitability images with similar unit measurement from 0 to 255. Afterwards, these all suitability images were aggregated using MCE and same weight for six factors has been applied because all the factors were considered having same influence for this research, which is actually not, but it needs further investigation to define the weight for each factor.

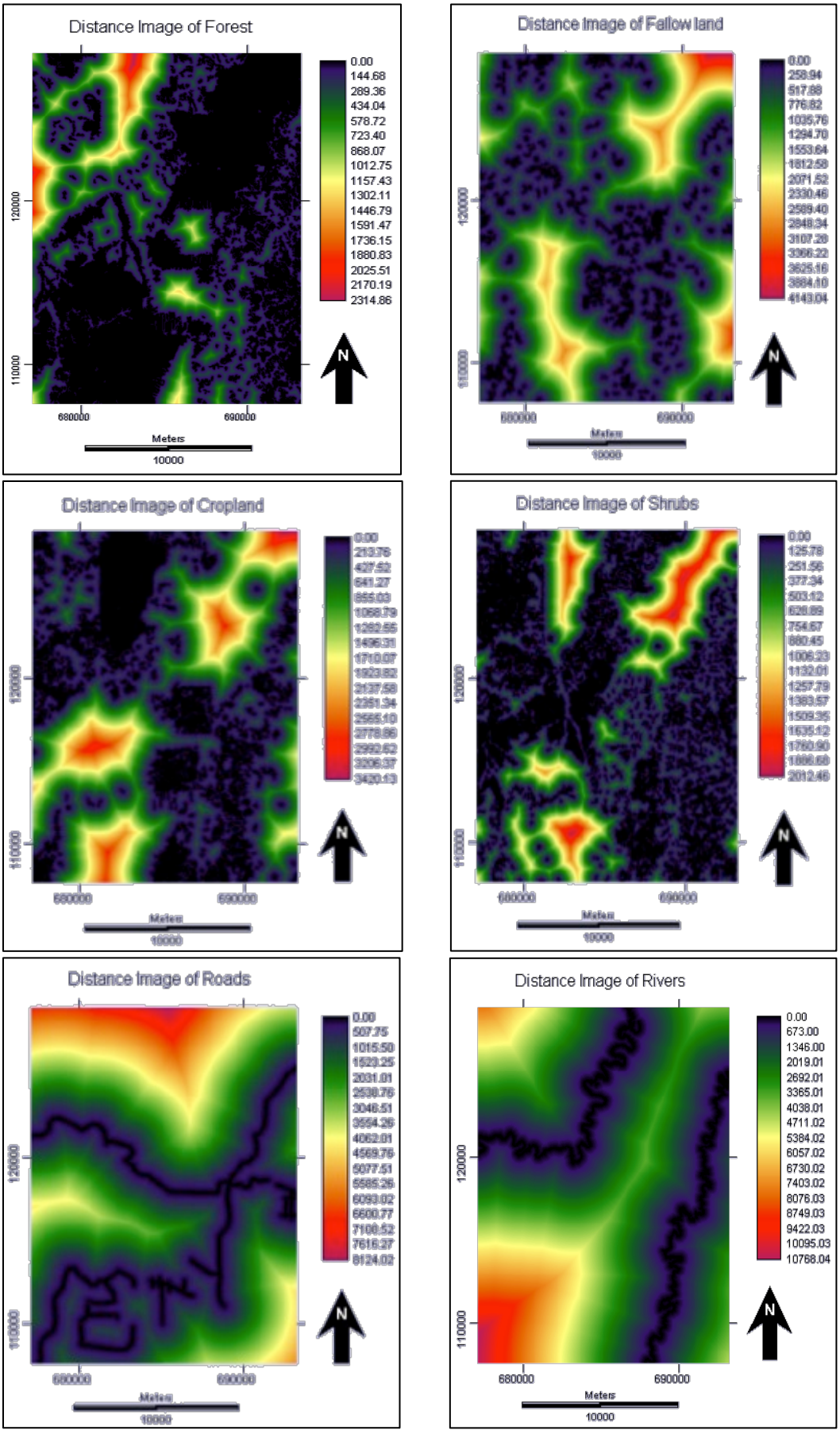


Figure 4.12 Distance Images of driving factors (2000) (unit: meters)

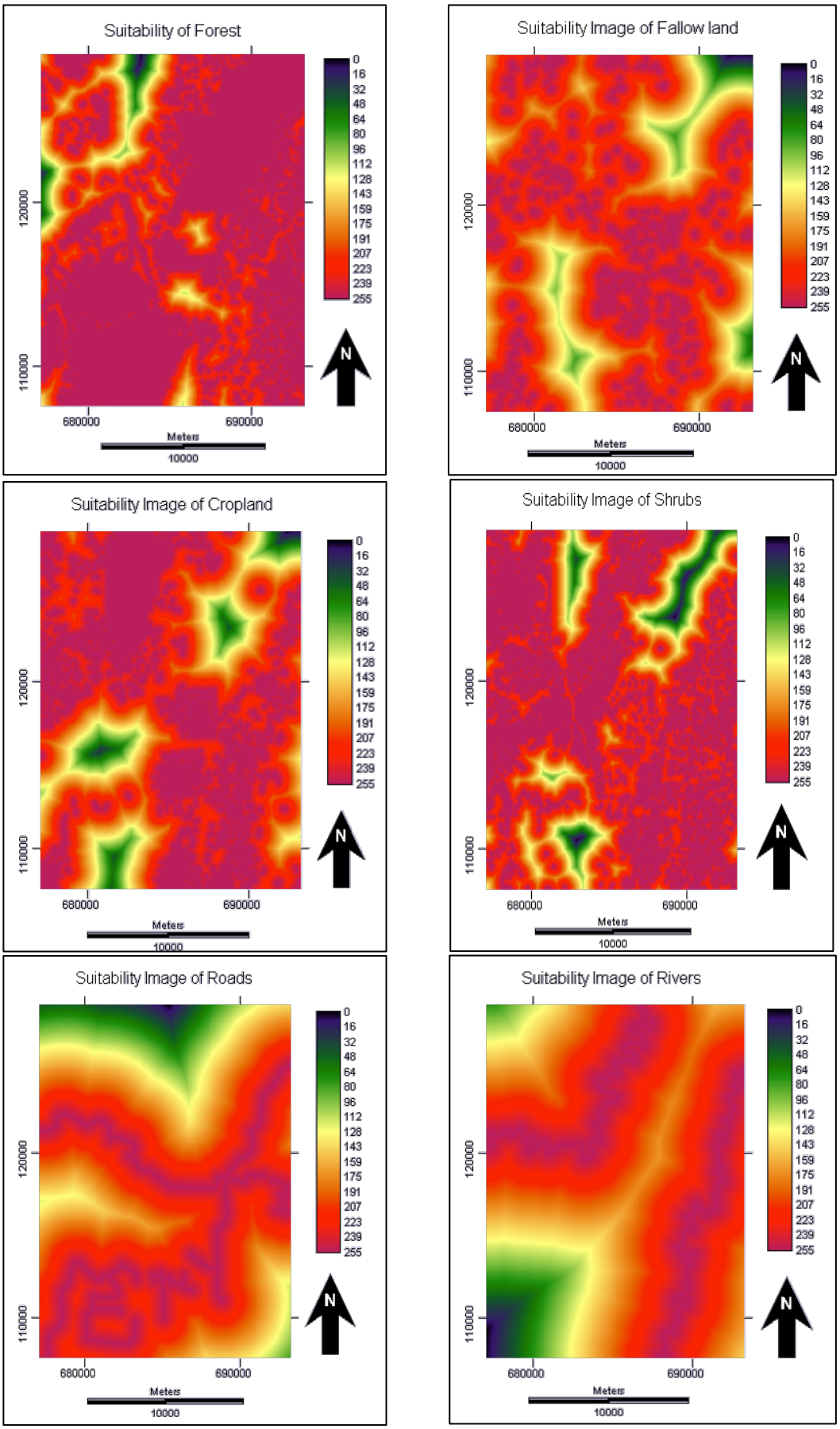


Figure 4.13 Suitability Images of driving factors (2000)

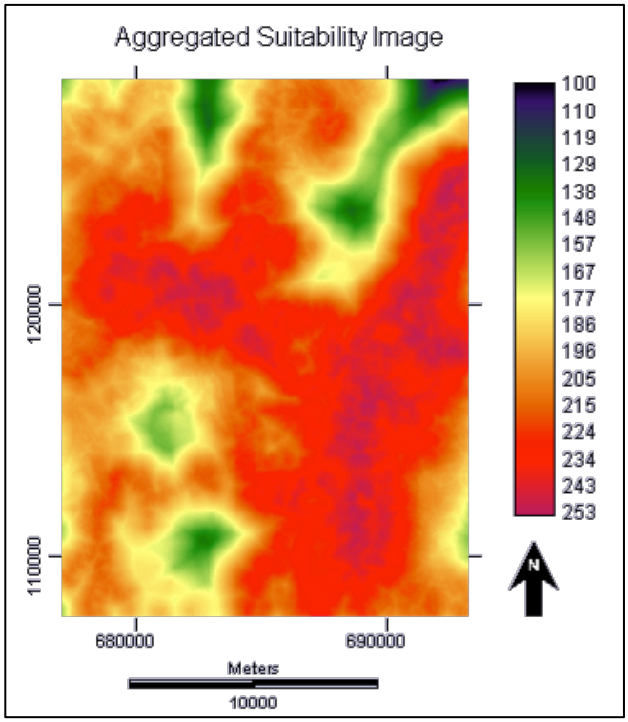


Figure 4.14 Aggregated Suitability Image

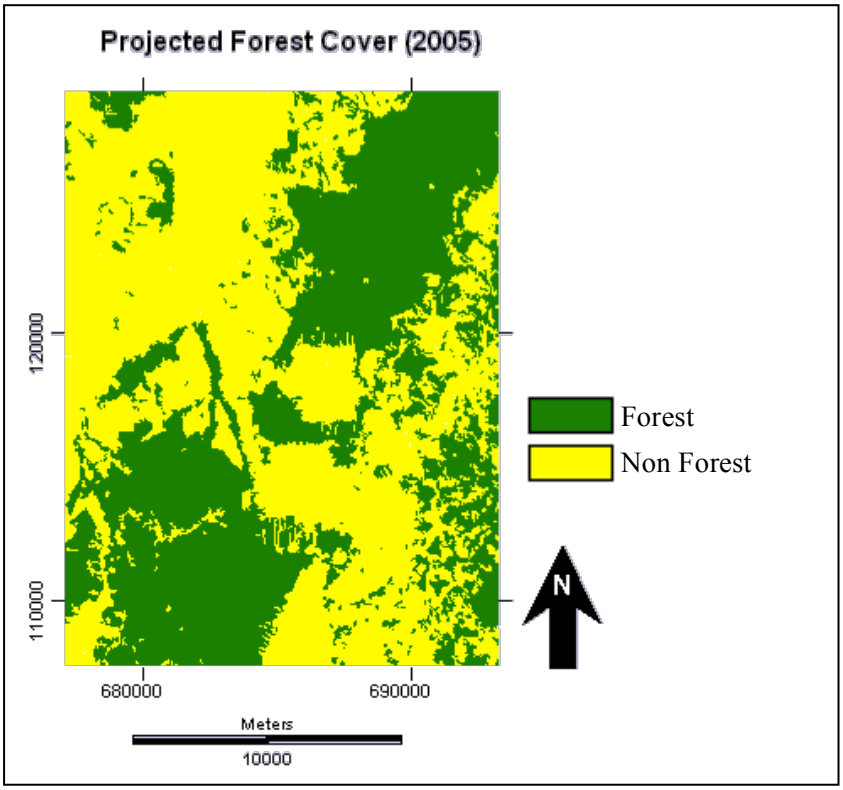


Figure 4.15 Projected Forest cover in 2005 using GEOMOD

4.2. Validation

In order to understand how accurate the model to predict land cover change, Kappa index of agreement has been chosen to measure the goodness-of-fit. Kappa could measure the similarity of predicted land cover and reference land cover for the same period. The formulation of Kappa is described in Appendix B.

Pontius (2000) portrayed that Kappa has following properties, Kappa = 1 if categorization is perfect, Kappa > 0 if observed proportion which is correct is greater than expected proportion because of chance, Kappa = 0 if observed proportion which is correct is equal to expected proportion because of chance, Kappa < 0 if observed proportion which is correct is less than expected proportion because of chance. In addition, Landis & Koch (1977) divided the range of kappa index in different category as follows:

Table 4.4 Categorization of Kappa Index

Kappa Statistic	Strength of Agreement
< 0	Poor
0.00 - 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 - 1	Almost Perfect

There are some variation of Kappa index including Kno, Klocation, and Kstandard. Kstandard describes the model's ability to obtain perfect prediction, whereas Klocation describes the accuracy to predict the location and Kno describes the quantity classified accurately related to the expected quantity classified accurately by a model with no ability to identify accurately quantity or location (Pontius, 2000).

Comparing the kappa index from the table below for three indexes either Kstandard, Kno, or Klocation, GEOMOD presents the highest kappa index. It means that GEOMOD gives better prediction than two other models. Besides that, by visually inspection to prediction error map in 2005 as the result of overlaying

predicted image and reference image, it is obvious that GEOMOD produces the least error even though CA_Markov also could give a proper prediction as the kappa index indicates a good result but conversely, Stochastic Model contributes a lot of error.

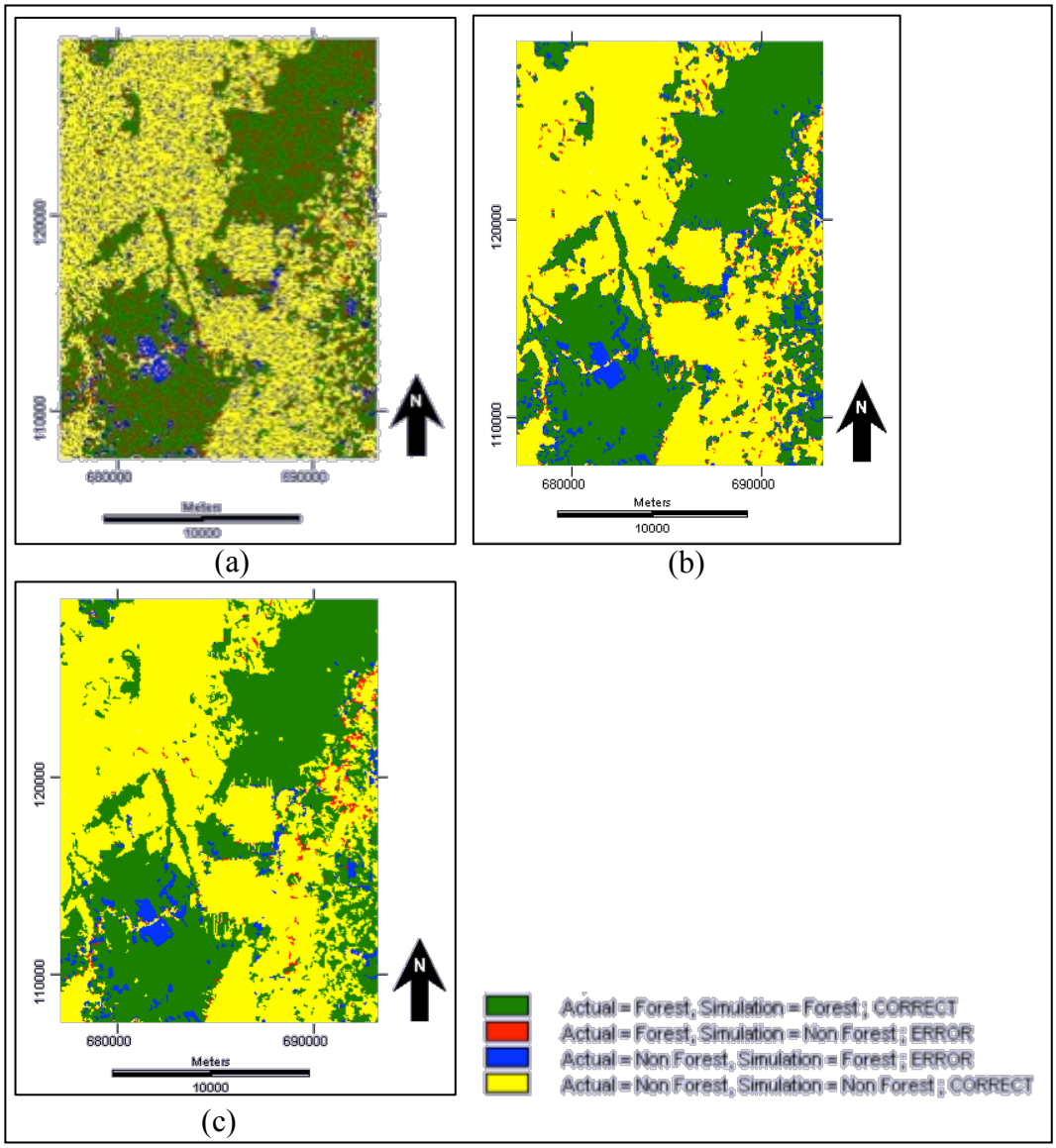


Figure 4.16 Prediction error map 2005 of (a) Stochastic Markov, (b) CA_Markov, (c) GEOMOD

Table 4.5 Kappa Index for each model

	Stochastic Markov	CA_Markov	GEOMOD
Kstandard	0.5620	0.8386	0.9131
Kno	0.6745	0.8801	0.9358
Klocation	0.6043	0.9030	0.9462

4.3. Discussions

Three models with their own characteristics have been completed in this research. The first model, which is stochastic markov model, does not allow applying contiguity rule. It means that this model assigns the new category without considering the neighborhood. As the result, the projected land cover produced by this model shows salt and pepper for all coverage. It implies that stochastic model lacks of spatial dependency. Apparently, knowledge of spatial dependence could be achieved by applying cellular automata in markov chain analysis because this model takes into account suitability image and contiguity rule to predict the location. Contiguity rule defines the filter window to consider the neighboring pixels in order to locate new category so it will assign new category for a certain pixel by considering the dominant category around that pixel otherwise it will remain as before and in the same time it assigns the category based on the suitability image, the more suitable of one pixel for particular category the more chance that pixel to be changed for that category. These procedures to predict location in order to give sense of geography are also working for GEOMOD. In addition, GEOMOD allows several driving factors to be considered as factors that could influence of one condition and these benefits would improve the model of prediction land cover change.

Even though GEOMOD contributes the highest Kappa, this model has a limitation in involving the number of category. GEOMOD only could apply for modeling two categories. Hence, it is recommended to apply CA_Markov for future modeling which involves more than two categories since CA_Markov also contributes high kappa. Moreover, another thing that we need to take into account when employing the model for prediction is the driving factor. As we can see in the prediction error map in figures above, the models were unsuccessful to predict in some parts though they present high kappa. It can be caused by the insufficient of driving factor that have been chosen or also could be caused by weight factors that have been applied. Thus, it is important for future work to consider a proper driving factor and weight factor for each driver. In order to do that, a local knowledge about particular case in certain area would be a great asset to get better model.

CHAPTER 5

FUTURE FOREST COVER MODELING

5.1. Modeling

GEOMOD has been performed to predict forest cover in 2015 since it contributes the highest Kappa to model forest in 2005 as explained in previous section. In order to predict the quantity of forest in 2015, the forest quantity in 2000 and 2005 have been taken into account. Thus, forest cover map in 2005 was selected as the beginning time to forecast forest cover in 2015. Estimation of quantity using linear regression has been applied (see Appendix D for further detailed calculation).

Prediction of location has been achieved by applying suitability maps and contiguity filter. Suitability maps were made by using similar driving factors as used in previous section, but they were derived from land cover map in 2005 since this map was used as the beginning time to predict forest cover in 2015. Nevertheless, the road map and river map used are similar to the one used to forecast forest cover in 2005 because there were no difference for these coverage. Contiguity rule using filter 5 x 5 has been employed. Further details to build GEOMOD can be seen below.

Table 5.1 Specification of GEOMOD parameters

Beginning land use image	Forest cover image of 2005
Mask or strata image	Unchecked
Neighborhood search mode	Constrained (5 x 5)
Beginning time	2005
Ending time	2015
Time step	1
Suitability Map	Use existing suitability image (created from MCE)
Environmental impact	Unchecked
Use Validation Image	Unchecked
Number of pixel Total	380024
Number of pixel State 1 BGN	163349
Number of pixel State 1 END	145572
Number of pixel State 2 BGN	216533
Number of pixel State 2 END	234310

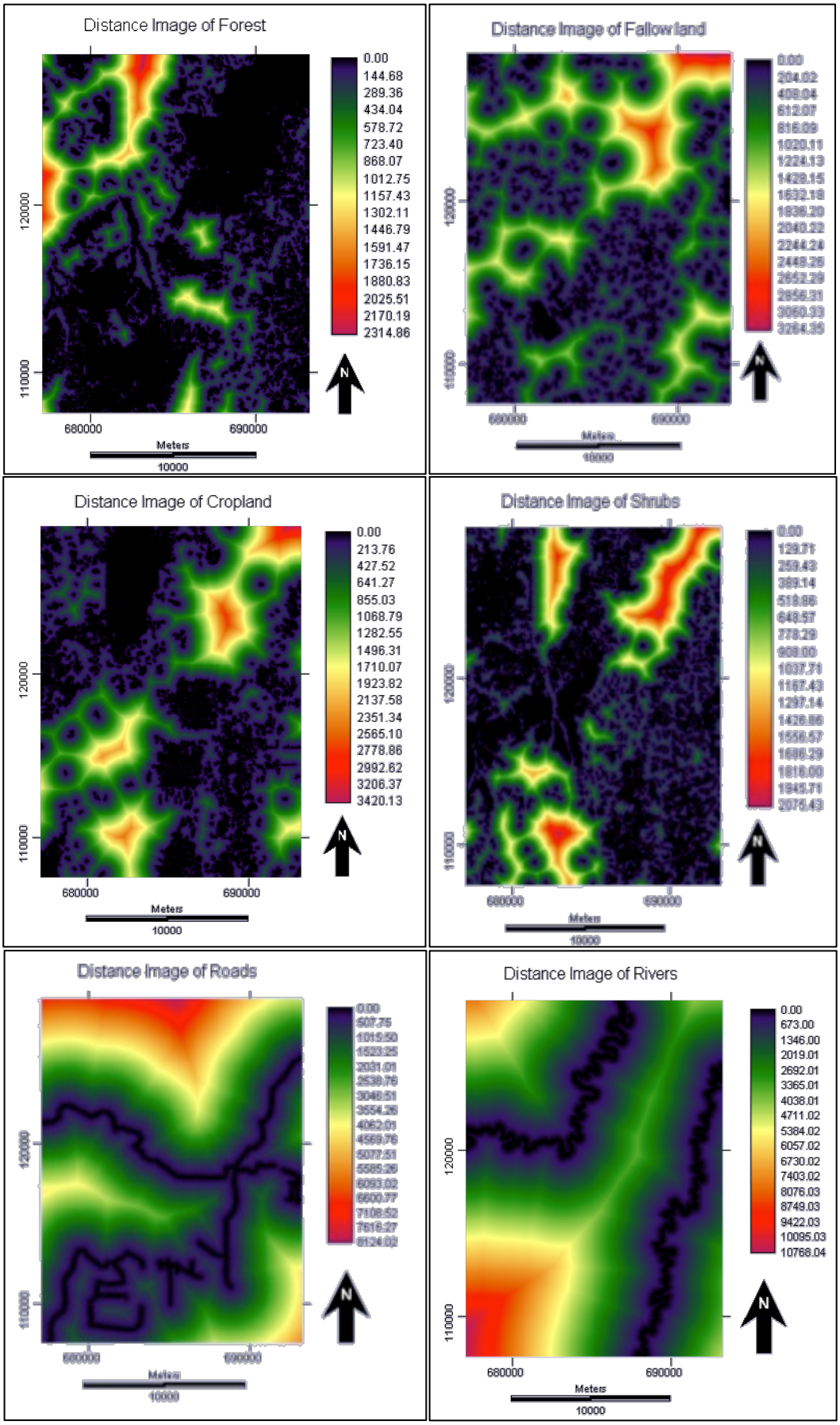


Figure 5.1 Distance Images of driving factors (2005) (unit: metres)

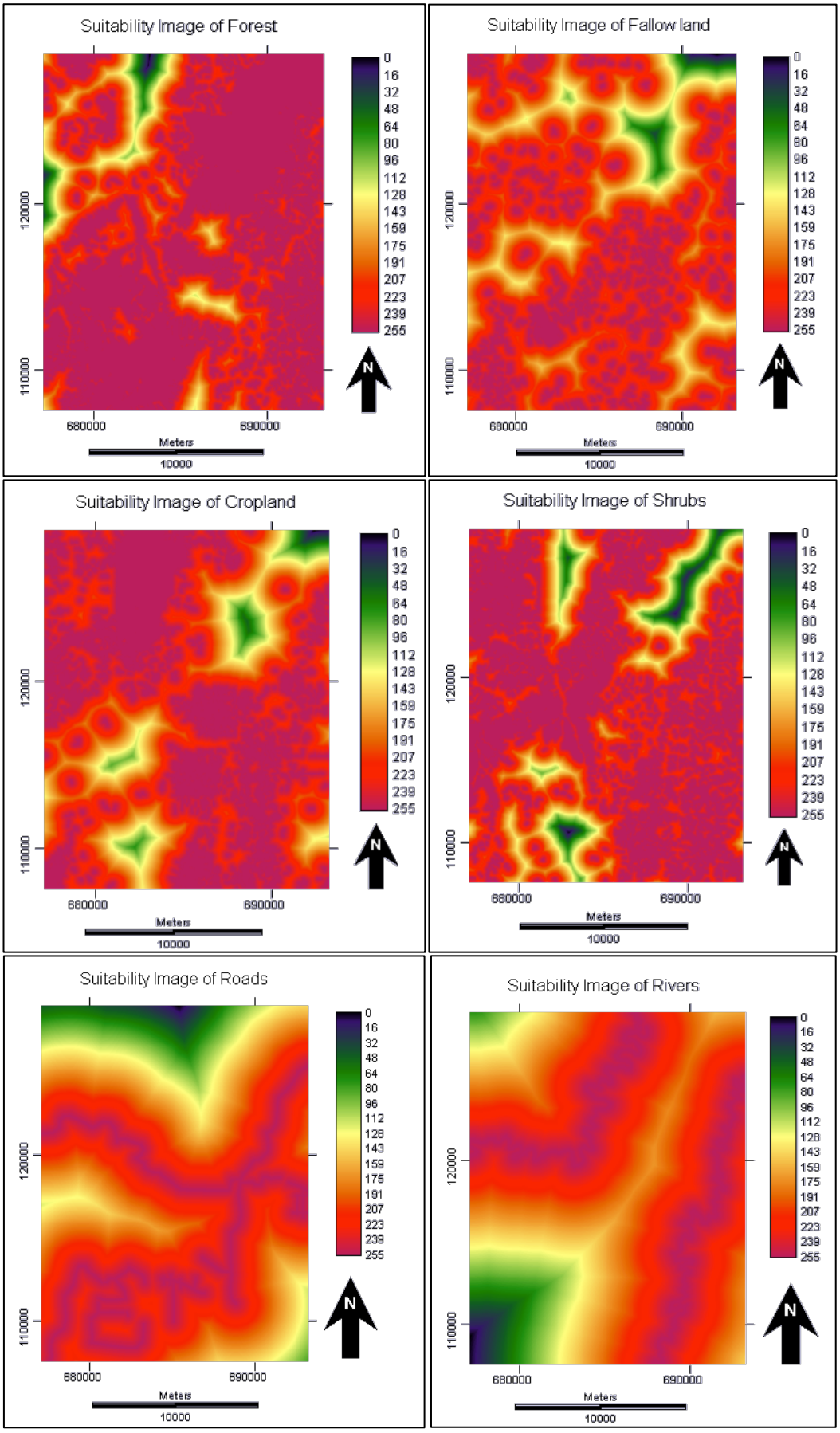


Figure 5.2 Suitability Images of driving factors (2005)

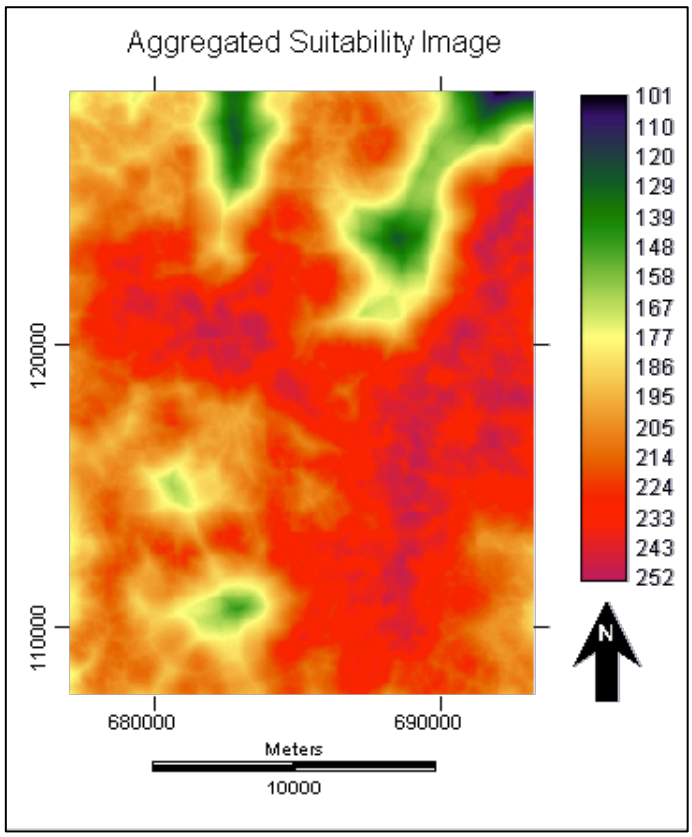


Figure 5.3 Aggregated Suitability Image (2005)

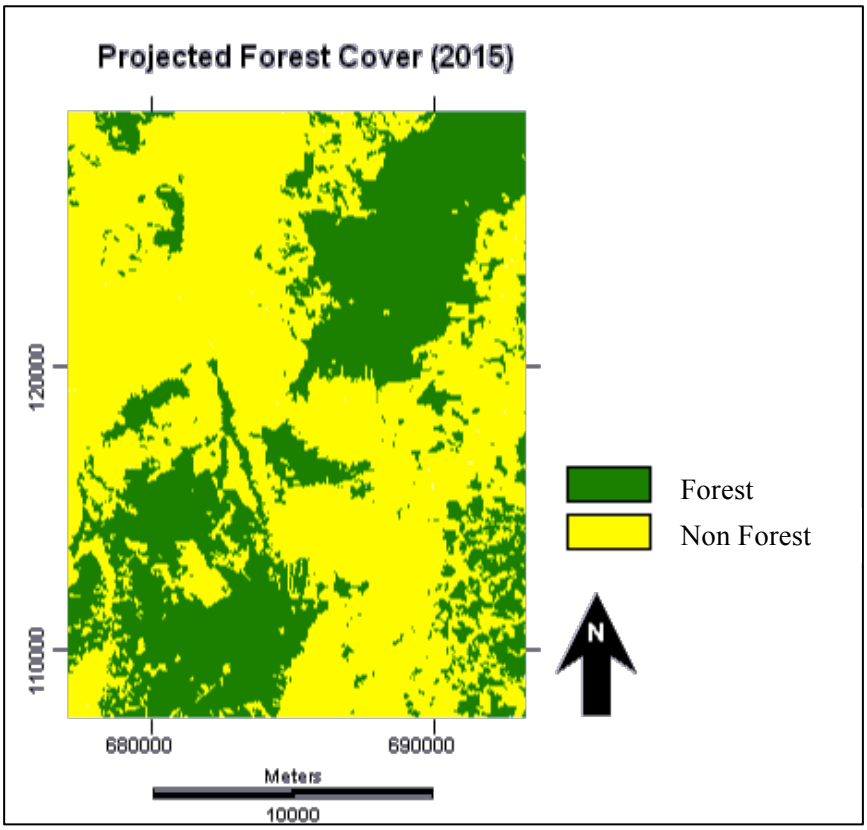


Figure 5.4 Projected Forest cover in 2015

5.2. Discussions

As the result, forest cover map in 2015 will decrease to 13,101 Ha and it shows that forest cover will be decreased by 12% during 2005-2015. It indicates a significant decrease. However, this prediction does not take into account information about protected area in which forest conversion will not allow to happen. For future research by using GEOMOD, it is recommended to consider protected area so that model could perform the reality closely. Indeed, information about future forest cover is very important as a consideration for further management by the stakeholders so that some decisions to prevent deforestation could be managed in a better way for forest sustainability.

As also mentioned in previous section that GEOMOD models a one-way transition, model that has been conducted in this research only considers the loss of forest instead of the gain. So, non-forested area in the initial time (2005) will remain non-forest in ending time (2015) but it is still acceptable because from the historical condition of forest between 1996 and 2005, there was no significant reforestation (gain of forest). In this case, GEOMOD works by choosing the location of cells to categorize as one of the two classes for the ending time. It is obvious that there is a net increase in non-forested category from initial time to ending time, then GEOMOD will search and select among the forested cells to be converted to non-forested area during this time interval.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

Conclusions have been drawn as follows:

1. Landsat images as remotely sensed data used in this research are sufficient to perform forest cover change analysis. Landsat image with its moderate resolution are able to support land cover type in the study area for further analysis. However, Landsat has a lack to avoid cloud cover because Landsat as an optical remote sensing works with short wavelength.
2. In this research, NDVI differencing combined with image classification has been performed to analyze the forest cover change and it proves the hypothesis that forest cover was decreasing in the study area.
3. Forest has decreased slightly between 1996 and 2000 by 1% and significantly decreased between 2000 and 2005 by 5%. Overall, forest has decreased by 6% between 1996 and 2005.
4. To forecast forest cover change in the study area, three kinds of models have been applied. These methods are Stochastic Markov Model, CA_Markov Model, and GEOMOD. As the result, GEOMOD has been used to predict forest cover in 2015.
5. Selection of the best model was attained by performing validation using Kappa index. Based on the simulation using three models, GEOMOD performed the highest Kappa so it was chosen to predict future forest cover in 2015 in the study area.
6. GEOMOD as the selected method to predict forest cover change contributed high Kappa index in this research, it can be stated that GEOMOD is qualified for forecasting forest cover change or other land cover change prediction.
7. Forest cover in 2015 will decrease to 13,101 Ha; it implies that forest cover will be decreased by 12% during 2005-2015.

6.2. Recommendations

The following recommendations could improve the similar research for future work:

1. Atmospheric disturbance in tropical area cannot be avoided and undoubtedly, it influences Landsat image with cloud coverage. It is recommended to have adequate reference data in order to adjust the coverage based on ground-truth data. Another way is by using another complementary image that has ability to penetrate cloud such as image with longer wavelength. This kind of image will inform better the existing land cover and it is also free of cloud, even though it would be very costly. Not only to prevent cloud cover, but also appropriate ground-truth data are very useful to measure how accurate our classification.
2. To obtain better model for prediction, it is not enough to be satisfied with high result of Kappa index. The complexity of land cover change will influence the simulation. In the case which has few complexity will result higher Kappa than the case with more complexity. So, it is advisable to engage appropriate driving factors. Hence, modeler should be able to have a good local knowledge of the study area in order to assign correct driving factors and improve the model to predict future cover. In addition, some constraints could be considered to improve the model such as protected area, which does not allow forest conversion.
3. Since GEOMOD works only a one-way transition either gain or loss between two categories, CA_Markov could be applied to simulate a prediction in which the gain and loss from several categories change dynamically as also CA_Markov still performs a sufficient model as found in this research.

REFERENCES

Anderson, J., Hardy, E., Roach, J., & Witmer, R. (1976). *A Land Use And Land Cover Classification System For Use With Remote Sensor Data.* U.S. Geological Survey. Washington: United States Government Printing Office.

Balzter, H. (2000). Markov chain models for vegetation dynamics. *Ecological Modelling, 126*, 139-154.

Biradar, C., Thenkabail, P., Gangodagamage, C., & Islam, A. (2003). *Landsat Enhanced Thematic Mapper (ETM+) Mosaic: Nominal 2000's Mosic for the Limpopo River Basin.* International Water management Institute (IWMI), Srilanka.

CCRS. (n.d.). *Natural Resources Canada.* Retrieved on November 25th, 2011, from http://www.nrcan.gc.ca/earth-sciences/geography-boundary/remote-sensing/fundamentals/1430

Centre Statistics Agency of Riau. (2010). *Riau in Fugures 2010.* Riau: Centre Statistics Agency of Indonesia - Riau Province.

Centre Statistics Agency of Rokan Hulu. (2009). *Rokan Hulu in Figures 2009.* Riau: Centre Statistics Agency of Indonesia - District of Rokan Hulu.

Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment, 37* (1), 35-46.

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B., et al. (2006). *Reducing Greenhouse Gas Emissions from Deforestation in Developing Countries: Considerations for Monitoring and Measuring* . Report of the Global Terrestrial Observing System (GTOS) number 46 , GOFC-GOLD report 26.

Denman, K. et al. (2007). *Couplings Between Changes in the Climate System and Biogeochemistry. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.

Doob, J. (1942). What is a Stochastic Process? *The American Mathematical Monthly , 49* (10), 648-653.

Eastman, J. (2009). *IDRISI Taiga Guide to GIS and Image Processing.* Massachusetts, USA: Clark University.

Eastman, J. (2009b). *IDRISI Taiga Tutorial.* Massachusetts, USA: Clark University.

Echeverria, C., Coomes, D., Hall, M., & Newton, A. (2009). Spatially explicit models to analyze forest loss and fragmentation between 1976 and 2020 in southern Chile. *Ecological Modelling , 212*, 439-449.

ERDAS. (2009). *Image Analysis for ArcGIS: Geographic Imaging by ERDAS.* ERDAS, Inc , Georgia.

ESRI. (2011). *ArcGIS Resources Center*. Retrieved on December 22nd, 2011, from http://help.arcgis.com

ESRI. (n.d.). *ESRI Developer Network*. Retrieved on December 01st, 2011, from http://edndoc.esri.com/arcobjects/9.2/net/shared/geoprocessing/spatial_analyst_tools/how_maximum_likelihood_classification_works.htm

FAO. (2007). *Brief of National Forest Inventory (NFI) of Indonesia.* Food and Agriculture Organization of the United Nations.

FAO. (2005). *FRA 2005-global table*. Retrieved on June 3rd, 2011, from http://www.fao.org/forestry/32033/en/

Fuller, D. O. (2006). Tropical forest monitoring and remote sensing: A new era of transparency in forest governance? *Singapore Journal of Tropical Geography, 27* (1), 15-29.

GOFC-GOLD. (2009). *A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation.* GOFC-GOLD Report version COP15-1, GOFC-GOLD Project Office, Natural Resources Canada, Alberta, Canada.

Gómez, O., & Milego, R. (2005). *CORINE Land Cover : How to analyze changes.* European Environment Agency.

Hayes, D., & Sader, S. (2001). Comparison of changedetection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. *Journal of Photogrammetric Engineering Remote Sensing, 67* (9), 1067-1075.

Holmgren, P. (2008). *Role of satellite remote sensing in REDD.* UN-REDD PROGRAMME. .

Ilhacinski, A. (2001). *Celullar Automata : A Discrete Universe.* Alexandria, Virginia, US: World Scientific Publishing, Singapore.

IPCC. (2003). Good Practice Guidelines on Land Use, Land Use Change and Forestry.

IPCC. (2000). *IPCC Special Report on Land Use, Land Use Change and Forestry.* The Hague: IPCC.

Izquierdo, L. R., Izquierdo, S. S., Galán, J. M., & Santos, J. I. (2009). Techniques to Understand Computer Simulations: Markov Chain Analysis. *Journal of Artificial Societies and Social Simulation, 12* (1).

Knight, J., & Lunetta, R. (2003). An Experimental Assessment of Minimum Mapping Unit Size. *IEEE Transactions on Geoscience and Remote Sensing*, 2132-2134.

Koomen, E., & Borsboom-van Beurden, J. (2011). *Land-Use Modelling in Planning Practice* (Vol. 101). Springer Netherlands.

Koomen, E., Stillwell, J., Bakema, A., & Scholten, H. (2007). *Modelling land-use change: Progress and applications* (Vol. 90). Springer Netherlands.

Lambin, E., Geist, H., & Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual review of environment and resources, 28* (1), 205-241.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Lillesand, T., & Kiefer, R. (2000). *Remote Sensing and Image Interpretation.* New York, USA: John Wiley and Sons, Inc.

Maerivoet, S., & Moor, B. D. (2005). Cellular automata models of road traffic. *Physics Reports , 419*, 1-64.

Metz, B., Davidson, O., Bosch, P., Dave, R., & Meyer, L. (2007). *Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

MoF. (2008). *Forestry Planning.* Ministry of Forestry of Indonesia, Jakarta.

Mondal, P., & Southworth, J. (2010). Evaluation of conservation interventions using a cellular automata-Markov model. *Forest Ecology and Management, 260*, 1716-1725.

NASA. (n.d.). *Earth Observatory*. Retrieved on December 20th, 2011, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php

NASA. (2008). *Landsat 7 Science Data Users Handbook.* Landsat Project Science Office.

Pachauri, R., & Reisinger, A. (2007). *Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC, Geneva.

Pekkarinen, A., Reithmaier, L., & Strobl, P. (2009). Pan-European Forest/Non-Forest Mapping with Landsat ETM+ and Corine Land Cover 2000 Data. *International Society for Photogrammetry and Remote Sensing (ISPRS) Journal of Photogrammetry and Remote Sensing* , 171-183.

Pérez-Vega, A., Mas, J., & Ligmann-Zielinska, A. (2011). Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environmental Modelling & Software, 29*, 11-23.

Pontius, R. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing, 66* (8), 1011-1016.

Pontius, R., Boersma, W., Castella, J., Clarke, K., de Nijs, T., Dietzel, C., et al. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science, 42*, 11-37.

Pontius, R., & Chen, H. (2006). *Land Change Modeling with GEOMOD.* Clark University.

Pontius, R., & Malanson, J. (2005). Comparison of the structure and accuracy of two land change models. *International Journal of Geographical Information Scienc , 19* (2), 243-265.

Pontius, R., & Schneider, L. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment, 85*, 239–248.

Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R., & Anderson, G. (2008). Using classification and NDVI differencing methods for monitoring sparse vegetation coverage: a case study of saltcedar in Nevada, USA. *International Journal of Remote Sensing, 29* (14), 3987-4011.

Quinn, J. (2001). Retrieved on November 15th, 2011, from http://web.pdx.edu/~emch/ip1/bandcombinations.html

Rashid, M., Lone, M., & Romshoo, S. (2011). Geospatial tools for assessing land degradation in Budgam district, Kashmir Himalaya, India. *Journal of Earth System Science* , 1-11.

Rashmi, M., & Lele, N. (2010). Spatial modeling and validation of forest cover change in Kanakapura region using GEOMOD. *38* (1), 45-54.

Research and Development Agency of Province of Riau. (2011). *Competitiveness Investment in Province of Riau.* Riau: Research and Development Agency of Riau Province.

Schoene, D. (2007). *Definitional issues related to reducing emissions from deforestation in developing countries.* Food and Agriculture Organization of The United Nations.

Torrens, P. (2000). *How cellular models of urban systems work (1. Theory).* London, UK: Centre for Advanced Spatial Analysis (UCL).

Uryu, Y. et al. (2008). *Deforestation, Forest Degradation, Biodiversity Loss and CO2 Emissions in Riau, Sumatra, Indonesia* . WWF Indonesia, Jakarta.

USGS. (2010). *U.S. Geological Survey.* Retrieved on December 14th, 2011, from http://landsat.usgs.gov/products_productinformation.php

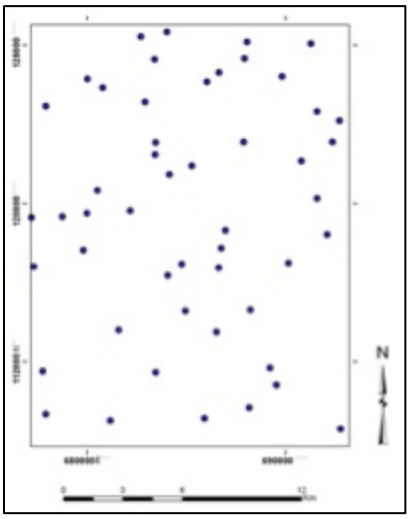
Weng, Q. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management , 64*, 273-284.

World Bank. (2006). *Sustaining Economic Growth, Rural Livelihoods, and Environmental Benefits: Strategic Options for Forest Assistance in Indonesia.* The World Bank.

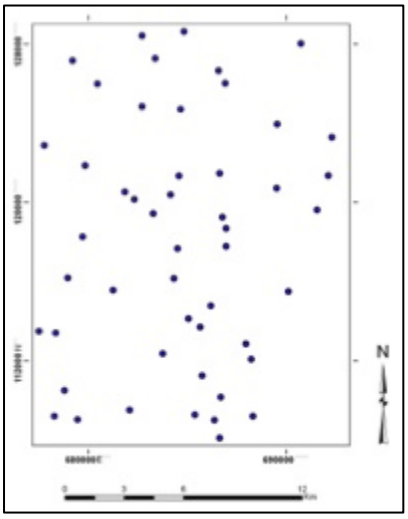
Appendix A
Parameters of Normalization

A.1. Training Samples

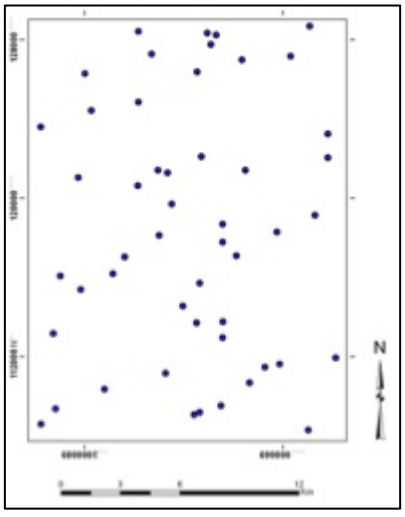
Training samples have been collected from each of NDVI image pair. In order to normalize one image to another image, 50 samples of unchanged land cover were subtracted as shown in figures below.



(a)



(b)

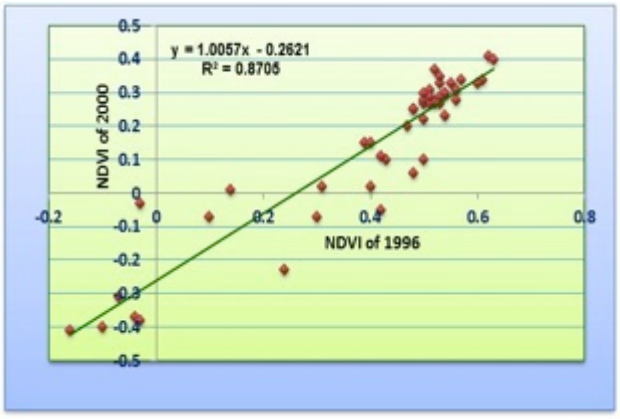


(c)

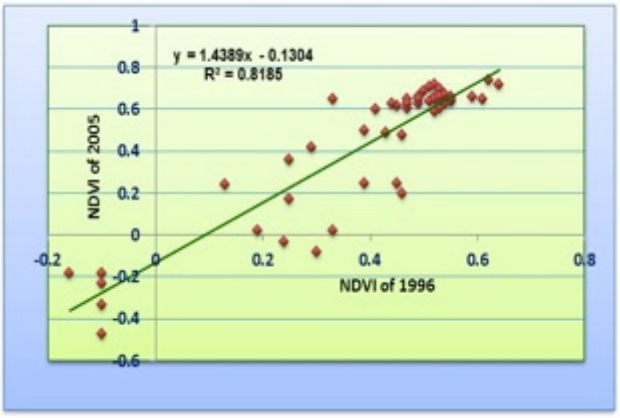
Figure A.1 Distribution of training samples for normalization between (a) 1996-2000, (b) 2000-2005 and (c) 1996-2005

A.2. Linear Function

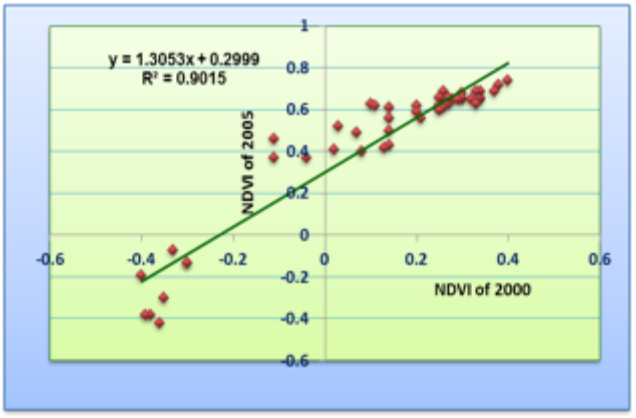
Linear function has been suggested to apply normalization. Thus, investigation of relationship between NDVI values subtracted from each pair has to be done to attain the linear function. Figures below show the relationship among NDVI image pair.



(a)



(b)



(c)

Figure A.2 Relationship between NDVI values of (a) 1996-2000, (b) 1996-2005 and (c) 2000-2005

Appendix B
Basic Terminologies of Accuracy Assessment

B.1. Overall Accuracy Assessment

Error matrix or confusion matrix illustrates a number of sample sets in row and column. By using these sample sets, one can assess the accuracy between reference data and classified data. An example below will explain further the accuracy assessment.

Table B.1 Confusion matrix between reference and classified map

Classified	Reference				Total
	A	B	C	D	
A	2	0	0	2	4
B	0	3	4	0	7
C	3	0	4	3	10
D	1	1	0	5	7
Total	6	4	8	10	28

Overall accuracy assessment defines the percentage of the total number of correct pixels (sum of diagonal) divided by total number of pixels. Additionally, error matrix also expresses the accuracy assessment of each single category involved in assessment. It is known as producer's accuracy (omission error) and user's accuracy (commission error). Producer's accuracy defines the percentage of correct number of particular category divided by total number of pixels for that category in reference, while user's accuracy defines the percentage of correct number divided by total number of pixels which have been classified to particular category.

Example:
Overall accuracy = (14/28) x 100% = 50%
Producer's accuracy:
A = (2/6) x 100% = 33.33%
B = (3/4) x 100% = 75%
C = (4/8) x 100% = 50%
D = (5/10) x 100% = 50%
User's accuracy:
A = (2/4) x 100% = 50%
B = (3/7) x 100% = 42.85%
C = (4/10) x 100% = 40%
D = (5/7) x 100% = 71.43%

B.2. Kappa

Kappa shows the agreement or similarity between observed or simulated map and the reality or reference map.

$$Kappa = \frac{P_o - P_c}{P_p - P_c}$$

Where,

P_o : The observed proportion correct
P_c : The expected proportion correct
P_p : The proportion correct

Then, the variation of Kappa index could be estimated by considering the contingency table of *J* categories (as shown in table B.2). The column indicates the reality or reference map, while the row indicates the classified or simulated map.

Table B.2 Contingency table of *J* Categories in which inputs are the proportions in study area

Simulation	**Reality**				
	1	**2**	**…**	***J***	**Total**
1	P_{11}	P_{12}	…	P_{1J}	$S_1 = \sum P_{1j}$
2	P_{21}	P_{22}	…	P_{2J}	$S_2 = \sum P_{2j}$
…	…	…	…	…	…
J	P_{J1}	P_{J2}	…	P_{JJ}	$S_J = \sum P_{Jj}$
Total	$R_1 = \sum P_{j1}$	$R_2 = \sum P_{j2}$	…	$R_J = \sum P_{jJ}$	1

Source: Pontius (2000)

This contingency table will be performed to define the proportion correct of simulated model in term of its ability to identify accurately both quantity and location.

Kappa for no ability or Kno performs the proportion correct classification in respect to the expected proportion of correct classification by a simulation in which there is no ability to identify quantity or location accurately.

$$Kno = \frac{P_o - NQNL}{1 - NQNL}$$

Table B.3 The Proportion correct classification regarding a simulation capability to identify correctly location and quantity

Ability to Specify Quantity	**Ability to Specify Location**		
	No. (NL)	**Medium (ML)**	**Perfect (PL)**
No. (NL)	$\frac{1}{J}$	$\left(\frac{1}{J}\right) + Klocation\left[(NQPL - \left(\frac{1}{J}\right)\right]$	$\sum_{j=1}^{J} min\left[\left(\frac{1}{J}\right), R_j\right]$
Medium (ML)	$\sum_{j=1}^{J} min\left(S_j, R_j\right)$	$Proportion\ correct\ observed, denoted\ P_o$	$\sum_{j=1}^{J} min\left(S_j, R_j\right)$
Perfect (PL)	$\sum_{j=1}^{J}\left(R_j^2\right)$	$PQNL + Klocation\ (1 - PNQL)$	1

Source: Pontius (2000)

Kappa for location or Klocation performs the ability of simulation to identify the location divided by the maximum possibility of simulation to identify location perfectly.

$$Klocation = \frac{P_o - MQNL}{MQNL - NQNL}$$

Standard Kappa or Kstandard performs the ability of simulation to achieve a perfect classification.

$$Kstandard = \frac{P_o - MQNL}{1 - MQNL}$$

Appendix C
Accuracy Assessment

C.1. Accuracy Assessment of Classified Image in 1996

Classified	Reference					Total
	Forest	Cropland	Shrub	Water bodies	Fallow land	
Forest	52	0	12	0	0	64
Cropland	0	26	1	0	4	31
Shrubs	1	3	28	0	3	35
Water bodies	0	0	0	1	1	2
Fallow land	0	0	0	0	8	8
Total	53	29	41	1	16	140

Overall Accuracy:	Producer's accuracy:	User's Accuracy:
$= \frac{52+26+28+1+8}{140}$ $= 82.14\%$	Forest $= \frac{52}{53} = 98.11\%$	Forest $= \frac{52}{64} = 81.25\%$
	Cropland $= \frac{26}{29} = 89.66\%$	Cropland $= \frac{26}{31} = 83.87\%$
	Shrubs $= \frac{28}{41} = 68.29\%$	Shrubs $= \frac{28}{35} = 80\%$
	Water Bodies $= \frac{1}{1} = 100\%$	Water Bodies $= \frac{1}{2} = 50\%$
	Fallow Land $= \frac{8}{16} = 50\%$	Fallow Land $= \frac{8}{8} = 100\%$

C.2. Accuracy Assessment of Classified Image in 2000

Classified	Reference					Total
	Forest	Cropland	Shrub	Water bodies	Fallow land	
Forest	46	0	14	0	0	60
Cropland	0	20	4	0	8	32
Shrubs	0	0	36	0	3	39
Water bodies	0	1	1	1	0	3
Fallow land	0	0	2	0	4	6
Total	46	21	57	1	15	140

Overall Accuracy:	Producer's accuracy:	User's Accuracy:
$= \frac{46+20+36+1+4}{140}$ $= 76.43\%$	Forest $= \frac{46}{46} = 100\%$	Forest $= \frac{46}{60} = 76.67\%$
	Cropland $= \frac{20}{21} = 95.24\%$	Cropland $= \frac{20}{32} = 62.50\%$
	Shrubs $= \frac{36}{57} = 63.16\%$	Shrubs $= \frac{36}{39} = 92.31\%$
	Water Bodies $= \frac{1}{1} = 100\%$	Water Bodies $= \frac{1}{3} = 33.33\%$
	Fallow Land $= \frac{4}{15} = 26.67\%$	Fallow Land $= \frac{4}{6} = 66.67\%$

C.3. Accuracy Assessment of Classified Image in 2005

Classified	Reference					Total
	Forest	Cropland	Shrub	Water bodies	Fallow land	
Forest	45	1	17	0	0	63
Cropland	0	21	9	0	3	33
Shrubs	0	4	30	0	1	35
Water bodies	0	0	1	1	1	3
Fallow land	0	0	0	0	6	6
Total	45	26	57	1	11	140

Overall Accuracy:	Producer's accuracy:	User's Accuracy:
$= \frac{45+21+30+1+6}{140}$ $= 73.57\%$	Forest $= \frac{45}{45} = 100\%$	Forest $= \frac{45}{63} = 71.43\%$
	Cropland $= \frac{21}{26} = 80.77\%$	Cropland $= \frac{21}{33} = 63.64\%$
	Shrubs $= \frac{30}{57} = 52.63\%$	Shrubs $= \frac{30}{35} = 85.71\%$
	Water Bodies $= \frac{1}{1} = 100\%$	Water Bodies $= \frac{1}{3} = 33.33\%$
	Fallow Land $= \frac{6}{11} = 54.54\%$	Fallow Land $= \frac{6}{6} = 100\%$

Appendix D
Quantity Prediction Using GEOMOD

GEOMOD does not have a certain method to extrapolate the change of quantity from one category to another category because the main aim of GEOMOD is to forecast the location of one-way transition (Pontius & Malanson, 2005). Nonetheless, GEOMOD has to specify the quantity of prediction so it is suggested to use linear extrapolation which is Y = ax + b. Y defines the area; x defines the time; a and b define the slope and intercept between 2 (two) parameters, respectively.

Modeler must supply the predicted quantities from one category; it could be the quantity of the first category or the second category in the ending time. Then GEOMOD will calculate the other one. In this research, the sequences of forest quantity in 1996 and 2000 have been chosen to predict the forest quantity in 2005. It was also applied to predict the forest quantity in 2015 based on forest quantity in 2000 and 2005.

D.1. Quantity Prediction of Forest Cover in 2005

Based on the classified image, it is known that forest area in 1996 and 2000 is 174,107 and 172,239 (pixel based), respectively.

By applying $Y = ax + b$

With 2 parameters,

$$\begin{array}{l} Y_2 = ax_2 + b \\ Y_1 = ax_1 + b \\ \hline Y_2 - Y_1 = a\,(x_2 - x_1) \end{array} -$$

$$x = \frac{Y_2 - Y_1}{x_2 - x_1} = \frac{172{,}239 - 174{,}107}{2005 - 1996}$$

$$x = -467$$

and by substitution,

$$b = Y - ax$$

$$= 174{,}107 - (-467 \text{ x } 1996)$$

$$= 1{,}106{,}239$$

$so, x = 2005$

$Y = (-467 \text{ x } 2005) + 1106239 = 169{,}904$

Then this result expresses the quantity of forest in 2005, which is based on pixel as the software works based on raster.

D.2. Quantity Prediction of Forest Cover in 2015

Based on the classified image, it is known that forest area in 2000 and 2005 is 172,239 and 163,350 (pixel based), respectively.

By applying $Y = ax + b$

With 2 parameters,

$$\begin{array}{l} Y_2 = ax_2 + b \\ Y_1 = ax_1 + b \\ \hline Y_2 - Y_1 = a(x_2 - x_1) \end{array} -$$

$$x = \frac{Y_2 - Y_1}{x_2 - x_1} = \frac{163{,}350 - 172{,}239}{2005 - 2000}$$

$x = -1777.8$

and by substitution,

$b = Y - ax$

$= 172{,}239 - (-1777.8 \text{ x } 2000)$

$= 3{,}727{,}839$

$so, x = 2015$

$Y = (-1777.8 \text{ x } 2015) + 3{,}727{,}839 = 145{,}572$

Then this result expresses the quantity of forest in 2015.

Masters Program in Geospatial Technologies

SPATIO-TEMPORAL DATA MODELING IN RESPONSE TO DEFORESTATION MONITORING

(A CASE STUDY OF SMALL REGION IN RIAU PROVINCE, INDONESIA)

DIAN NURAINI MELATI

Supported by:

Education and Culture

Erasmus Mundus