Laboratory of Geo-Information Science and Remote Sensing

Thesis Report GIRS-2015-21

Predicting Oil Palm Land Use Following Deforestation Using Available Spatial Parameters

Sukmo Pinuji

May 2015

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Predicting Oil Palm Land Use Following Deforestation Using Existing Spatial Parameter

Sukmo Pinuji

Registration number 820410653010

Supervisors:

Sarah Carter

dr. ir. Lammert Kooistra

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

at Wageningen University and Research Centre,

The Netherlands.

May 2015

Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2015-21
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

ABSTRACT

Understanding the characteristics of agricultural expansion, particularly oil palm, is important to study its impact on the world's land. Land use modelling is a tool that can be used to help to understand the key process of oil palm expansion, to assess the current state, the drivers, the processes and the impact of oil palm expansion. Using spatial datasets from different sources, this research models the process of land use change using IDRISI Land Change Modeller to understand the follow up oil palm land use after deforestation events in Indonesia, as well as to predict where the deforestation will likely to occur due to the process of oil palm expansions. Artificial Neural Network method was used to build sub-models during the observation period for the year of 2000 – 2006, while Markov Chain Method was used to predict future land use in 2009.

The results shows that the actual trend of deforestation and oil palm expansion during observation and prediction periods was change significantly, thus affecting the model accuracy in predicting changes, especially for the class that dominated the change after deforestation, such as other class in Riau and oil palm class in West Kalimantan and East Kalimantan. In Riau, other class were overestimated 47.63% more than the actual. In East Kalimantan and West Kalimantan was predicted by the model 32.86% and 54.57% less than the actual increases. Apart from those two classes, the following land use after deforestation events were predicted better by the model, and only gives small differences with the actual changes. We also found that distance from the existing oil palm plantations was the most significant variable in the process of deforestation and oil palm expansion. We conclude that Artificial Neural Network and Markov Chain Method are useful to model and to predict land use change following deforestation and agriculture expansion, but failed to account external factors of deforestation and agriculture expansion drivers that were not included during the modelling process.

Keywords: deforestation, oil palm, land use change modelling, Artificial Neural Network, Land Change Modeller, agriculture expansion.

Foreword and Acknowledgement

“Tetapi apakah kecerdasan pikiran itu sudah berarti segala-galanya? Bila orang hendak sungguh-sungguh memajukan peradaban, maka kecerdasan pikiran dan pertumbuhan budi harus sama-sama dimajukan”

"Is that, however, the intellectuality really meant everything? If men really want to develop civilization, then all the intellectuality, mental development and conscience have to be grown altogether”

(Kartini, 1900)

Environmental issues are always being my concern. The chance of working this thesis is both a fortune and challenge, as I learn a lot during the research, not only with the new technical knowledge but also in the process of learning itself, which give me a precious experience and amazing adventure.

First of all, I want to thank my supervisors Lammert Kooistra and Sarah Carter for guiding my research. They were always open for simulating discussions and encouraged me to improve and learn more. They also help me a lot in improving my scientific and analytical skill during this thesis.

I would also thank to all my fellow master students and all of friends for the supports and encouragements through a nice time and thousand cups of coffee that we spent in thesis room. A big thank also for my parents to be such a great parents as they are.

Lastly, Qoria and Widarmoko, without whom I shall never be confident to be in this path as far, thanks for being such a great team. You are my inspiration, my spirit, and place where my heart always found home.

Life is a great adventure and full of wonders, all we have to do is open our eyes and learn. Thanks God for this wonderful life you gave me.

Index

List of table

List of figure

ABBREVIATIONS

MoF	:	Ministry of Forestry, referring to the government authority of Indonesia
ANN	:	Artificial Neural Network
Modified GA	:	Modified Genetic Algorithm
CA	:	Cellular Automata
WDPA	:	World Database Protected Area
TPM	:	Transitional Probability Matrices
LCM	:	Land Change Modeller
MCK	:	Map Comparison Kit
BPS	:	Badan Pusat Statistik or National Statistical Bureau of Indonesia
FAO	:	Food and Agricultural Organization
DEM	:	Digital Elevation Model

I. INTRODUCTION

Deforestation has become a global concern, and many issues related to forest loss have been studied extensively in various scales and aspects. Research and discussion about the drivers and causes of deforestation has been conducted, and many anthropogenic factors have been identified as the main causes of forest loss and forest degradation (Hosonuma et al. 2012, Kissinger et al. 2012). Furthermore, agricultural activities are mentioned as one of the main causes of deforestation and forest degradation (Geist and Lambin 2001, Etter et al. 2006, Morton et al. 2006, Ramankutty et al. 2008), and oil palm is referred to as one of agricultural commodities responsible for the decreasing of forests area, especially in the equatorial area (Sandker et al. 2007, Fitzherbert et al. 2008).

To understand the characteristics of oil palm expansion in changing global forest covers is crucial to account its extent and the impacts on the environment, as well as to predict the future expansions. Spatial based land use modelling can be a powerful tool to understand the processes and stages of agricultural expansions in changing the world's cover, and provides the flexibility to work with various input and methods that are widely available now (Sklar and Costanza 1991, Lambin 1997, Veldkamp and Lambin 2001). Furthermore, the availability of open source spatial datasets improves the possibilities of modelling land use change in various levels of details and complexities and provides wider opportunities to explore the process of agricultural expansions and its impacts on the environment.

I.1. Background information

Agricultural activities provide crucial services for human life through food demands, and also in economic development related with agricultural industries (Johnston and Mellor 1961, Hassan et al. 2005, Etter, McAlpine et al. 2006). In line with population growth, the demand of food production is increasing, and requires agricultural industries to intensify its production, as well as to expand the production area. As a trade-off, environmental problems such as tropical deforestation, biodiversity loss, the degradation of soil quality and greenhouse gas emissions are rising (Foley et al. 2005, Etter et al. 2006, Ramankutty et al. 2008).

Many studies have been conducted to investigate and quantify the environmental impact of agricultural activities, and spatial analysis through land use change modelling has become a powerful tool (Lambin 1997, Lambin et al. 2000). Modelling land use change can provides a better understanding of the proximate and underlying drivers of land use change, to detect the trends and dynamics, to assess the impacts and to predict the future changes (Lambin 1997, Veldkamp and Verburg 2004). Many methods of spatial modelling of land use change have been developed, and often, the choice of method and level of detail in land use modelling are highly depending on the data availability and data quality (Veldkamp and Verburg 2004). Furthermore, with the advance development in GIS and remote sensing, spatial data are now extensively available in diverse formats and scales. The technology also allows non-spatial information such as crop yield, population density and other census-based data to be mapped and be presented spatially, and be provided in the form of thematic maps. Many GIS platforms also provide datasets that can be easily accessed. The easiness of access and the broad range of available data on various scales put open source data as a promising sources for GIS analysis, including the issues of land use change and agriculture expansion modelling.

I.2. Problem definition

While open source data are widely available, the challenges arise when the data have a different source and format, different processing technique and different spatial and temporal resolution, which often cause inconsistencies when the users try to compile several data for some purposes (Verburg et al. 2011). To compile and to select suitable data for land use change modelling are sometimes challenging. Often, global open source data with thematic information are available in coarse resolution. Moreover, global census-based maps, such as crop maps and yield maps usually lack of validation, where no independent remote sensing survey can be carried out to validate the data at time where they compiled, thus the level of accuracy of the data cannot be verified (Strahler et al. 2006, Kongsager and Reenberg 2012). This condition often raises the questions of for what can the data be used and are the specifications of the data available to meet the need of analysis. We study the possibility to associate the usage of available open source datasets to get a better understanding on the trend and dynamic of oil palm expansion in relationship with deforestation. Several open source data contain information that are considered as variables of land use change related with agricultural expansion and deforestation such as slope, elevation, protected area, and crop map were used (Boyd 1996, Lambin 1997, Overmars et al. 2003). Together with available data at national level provided by the government that are available in finer resolution, these datasets were incorporated to build a model of land use change for predicting oil palm expansion following deforestation. We choose oil palm as a representative of agricultural commodities, as this crop is mentioned as the fastest growing commodity in our study area (Sunderlin 1996, Sandker et al. 2007, Susanti and Burgers 2012). Since we incorporated some variables that we consider have complex relationship with the process of deforestation, we use Artificial Neural Network (ANN) as the method of land use modelling. This method is referred capable to model a nonlinear relationship within input variables and has been proved give better result compare to other methods, especially when the relationship between variables are unknown (Hill et al. 1994). Furthermore, we use Markov Chains method to build predictive land use change in our desirable year (Eastman and Toledano 2000, Gilks 2005). These methods were available in Land Change Modeller (LCM) extension in IDRISI SELVA software developed by Clark Labs, Clark University, California, which was used in this research.

I.3. Research objectives and questions

The overall aim of this study is to assess the possibility of using available open data as spatial parameters to predict oil palm land use following deforestation in Indonesia. To ascertain the main objective, we break down three sub-objectives as below:

1. Describe land use dynamics in the study area.
2. Design and test a model to predict deforestation event using available crop map dataset and indicators of suitability for conversion to agriculture.
3. Assess the accuracy and suitability of using crop maps to predict follow up land use following a deforestation event.

To answer our objectives, several research questions are defined as listed below.

1. What is the trend and dynamic of land use change including deforestation events?
2. Are there any relationship between existing land use and follow up land use following deforestation?
3. Is the model can be used to predict land use change following deforestation?
4. What is the accuracy of the model?
5. Can we use the crop map dataset to predict follow up land use following deforestation?

I.4. Thesis Content

The first chapter of this report presents the introduction, background information, problem definition and research objective and questions, to introduce on why it is important to model the process of agriculture expansion, what are the available data that can be used and why we choose our selected method in modelling land use.

In the second chapter, we summarized theoretical background related to our study. We described the condition of deforestation and oil palm plantation in Indonesia, what variables that are considered as the drivers and constraints, and some description about Artificial Neural Network and Markov Chain methods that we used in this research.

In the third chapter, we described the methods that we used to answer the research questions and objective, while the results and the discussions of our findings were summarized in the fourth and sixth chapter respectively. Lastly, we summarized the conclusion in the seventh chapter.

II. THEORETICAL BACKGROUND

II. 1 Deforestation and Oil Palm in Indonesia

Forests in Indonesia account for 11% of the world's remaining tropical forest area, and are important for biodiversity, home for numerous endemic and rare species, and play an important role as a carbon sink which can help to prevent global warming and climate change (Sunderlin 1996, Iremonger et al. 1997, Fearnside 2000, Lee et al. 2007).In 2013, the Ministry of Forestry (MoF) reported that Indonesian forests cover an area of 98,072.7 million ha, or approximately 52.5% of the total land area (Kehutanan 2014). However, Indonesia is also experiencing rapid deforestation, as Sunderlin (1996) found that deforestation rates in Indonesia have been increasing, from 300,000 ha/ year in 1970 to 600,000 ha/ year in 1980 and 1 million ha/ year in 1990 (Sunderlin 1996). Forest Watch Indonesia found that the rate of deforestation increased by 1996-2000 by approximately 2 million ha/ year (Indonesian Forest Watch, 2002). Many debates arise on the causes of deforestation in Indonesia, from shifting of cultivation or traditional agriculture (Davis and Ackermann 1990, Mertz et al. 2009), logging activities (Barbier et al. 1995, Palmer 2001) population pressures (Fraser 1998, Sunderlin and Resosudarmo 1999), transmigration and resettlement program (Fearnside 1997), and oil palm expansion (Sandker et al. 2007, Budidarsono et al. 2013**)**. Oil palm is one of the fastest growing agricultural commodity around equatorial area, including Indonesia (Fitzherbert et al. 2008, Koh and Wilcove 2008, USDA 2010, Wicke et al. 2011, Susanti and Burgers 2012). This commodity is mainly used as a source of edible vegetable oil and oleo-chemical industries. Palm oil has become a favourable commodity compare to other plant-based oil regarding the relatively low of cost production that make this commodity available at competitive prices. Further development also puts palm oil as a promising alternative for bio fuel and renewable source of fuels, as the demand to reduce the usage of fossil fuels is increasing regarding to environmental issues (USDA 2010, Susanti and Burgers 2012, Locke and Henley 2013).

In Indonesia, palm oil industry plays important roles for national economic development, as this commodity contributes substantially to national incomes, and in some regions becomes the major driver of economic development (USDA 2010, Susanti and Burgers 2012). However, palm oil is also denounced as a major cause of deforestation and environmental degradation, even though this statement is still debated among the scientists (Fitzherbert et al. 2008, Koh and Wilcove 2008, Budidarsono et al. 2013).

Currently, the Indonesian Government states that approximately 24.5 million ha of lands are potentially suitable for oil palm cultivation, with around 7.65 million ha has already been developed and another 6.5 – 7 million ha of undeveloped land have had a permit/ concession for oil palm plantation (USDA 2010). Two primary land types have been targeted to be converted into oil palm plantation by the MoF, which are production forest and dry land agriculture (USDA 2010). Nonetheless, Hunts (2010) mentioned that some land use such as grassland are not favourable for oil palm investment, even though this land type is easier to be converted into oil palm plantation prior to the land preparations such as forest clearing process. The reason why forest area is more favourable for the new expansion is the obtainments of additional profit from by-product such as timber harvesting from forest clearing (Hunt 2010).

Forest conversion into oil palm plantation is regulated through the Decree of Ministry of Forestry and Agriculture Nr. 376/Kpts-ii/1998 (Manurung 2000). The decree mentioned that type of forest that can be converted into oil palm plantation should has maximum slope of 25%, locates on the altitude of 0 – 300 m above sea level, has rain falls of 1750-4000 mm/year with the average of dry month per year 0 – 3 months, has the effective soil depth for mineral soil more than 100 cm and for peat area less than 200 cm, and has average temperature per year 24° - 29° C (Decree of Ministry of Forestry and Agriculture Nr: 376/Kpts-II/1998, 1998).

Considered as high values commodities, oil palm plantations in Indonesia are not solely dominated with the large plantations. Smallholder plantations, typically owned by the local farmers with the area of cultivations for less than 2 ha, also characterized the plantations in Indonesia. The rising of smallholder plantations was supported by the establishment of State Regulation Nr.13/ 1995 about the Industrial Business Licence, which allowed local farmers to access technology and market through nucleus estate and partnership scheme with large companies, as a way from the Government to boost up palm oil industries. In consequences, many local farmers shifted their land into oil palm, and some of them start to open new plantations by clearing the forests (Budidarsono et al. 2013). Even though single smallholder cultivates only approximately 2 ha of plantation per household, they play a significant part in the palm oil industry in Indonesia. Smallholder plantations account for 35%-45% of the total area of planted oil palm, and contribute up to 33% of the total production (Vermeulen and Goad 2006, Rist et al. 2010). The existence of smallholder plantations, however, influences the pattern of oil palm expansion within regions. To avoid deterioration, palm oil fresh fruit bunches must be milled within 24 hours of harvest time. Hence, access to infrastructures, nearby plantation or location within delivery distance to the nearest mill is necessary (Vermeulen and Goad 2006). In contrary, large plantations are usually established in scarcely populated area, or even in empty frontier area, disregard the existing infrastructures or settlement as consideration and in most of the case will stimulate new growth centre within the area (Budidarsono et al. 2013).

II. 2 Explanatory variables of deforestation drivers

Several variables are considered to correlate with the process of deforestation events and characterize the process of land use changes regarding oil palm expansion. These variables are usually specific within the location. Recognizing and quantifying them are useful in spatial modelling as these can be used to explain the process of agriculture expansion and explain its proximate and underlying causes.

Several variables have been considered as the drivers of deforestation and agriculture expansion. Besides anthropogenic factors, topographical conditions also characterized the process. Elevation, slope, proximity to residential area, proximity to road, proximity to forest edge and population density are the most common variables included in land use modelling (Boyd 1996, Lambin 1997, Schneider and Pontius 2001, Overmars et al. 2003, Rajan 2010). Moreover, some variables also considered as constraints, such as protected area or certain land use type, which hinder the changes of land use due to restriction and limitation(Verburg et al. 2006, Rajan 2010). These variables can be mapped, and many modelling tools such as CA-MARKOV, CLUE-S and LCM provide the possibilities to quantify them and explain the contribution of each variable in land use change process statistically (Verburg et al. 2002, Pontius Jr et al. 2008, Camacho Olmedo et al. 2013). Nonetheless, regardless the importance of the variables and its significance in influencing land use change, the choices of including certain variables in land use modelling usually depend on the availability of the data.

II.3 Modelling Land Use Change

Land use modelling has been widely used to explain the trends and dynamics of shifting of land use over the time. Model of land use should have the capabilities to present the complexity of the land use system in the past and present, and to project the alternative pathways into the future. Land use model should also provide the possibilities to understand the key process of land use change, which is important to understand the current state, the drivers, the processes and the impacts of changes (Sklar and Costanza 1991, Veldkamp and Lambin 2001). Land use modelling has been widely used for many purposes, from spatial planning (Zander and Kächele 1999, Ligtenberg, et al. 2001, Waddell 2002) to monitoring and evaluation (Johnes 1996, Lambin 1997, Jat, et al. 2008), and from local scale (Matthews 1983, Blennow and Persson 1998, Morgan, et al. 1998, Verburg, et al. 2002, Waddell 2002) to global scale (Prentice, et al. 1993, Monfreda, et al. 2008, Ramankutty, et al. 2008).

The principle of modelling land use change is how to capture the complexities of land use dynamics over the time and describe it in quantitative terms (Lambin, et al. 2001). Several instruments have been developed for land use change modelling purpose, and provide various methods for its quantifications, either deterministic or stochastic models. In general, land use modelling methods can be categorized regarding of some assumptions concerning the number of possible categories, types of category transitions, spatial dependency, feedbacks, cross-scale linkages and many more (Pontius and Malanson 2005). Some of these issues are concerning with highly complex behaviour of the model such as interactions of the parameters, and some are concerning with more basic issues such as the format of required data (Pontius and Malanson 2005). Moreover, the selection of modelling method is highly depends on several considerations such as the purpose of the modelling, data availability and the scale or extent of the model (Verburg et al 1999, Pontius and Malanson, 2005).

Some land use modelling techniques have been implemented in Indonesia for many purposes. Romijn et al. (2013) used standard Ordinary Least Square (OLS) multivariate regression to predict future changes at district level in order to quantify the effects of different forest definitions and their impact on developing REDD+ reference emission levels in Indonesia. Warlina (2007) used CLUE-S model as a method for sustainable spatial planning in the district scale, and Suprayogi (2003) implemented ANN method to predict the availability of water supply at district level. From several methods that had implemented in Indonesia, the results show that the applied methods were proven can be used to simulate the land use change process in Indonesia and addressing the questions related to them. (Romijn, et al. 2013)(Suprayogi 2003, Warlina 2007)

II.4 Artificial Neural Network (ANN) Multi Layer Perception

ANN is widely used in many fields to quantify and to model complex behaviours and patterns of different parameters in the process of land use change. Its capabilities to recognize nonlinear patterns and the relationships between variables make this method is being widely used in various sectors, from medicine (Dybowski et al. 1996, Khan et al. 2001), hydrology (Thirumalaiah and Deo 1998, Tokar and Johnson 1999, Nagy et al. 2002), to image classification and remote sensing (Hepner et al. 1990, Kanellopoulos et al. 1992, Civco 1993). This method has a non-linear mapping structure based on the function of the human brain, and has many advantages such as the ability to handle non-linear functions, to perform model-free function estimation, to learn from data relationships that are not otherwise known and to generalize unseen situation (Mas et al. 2004).

ANN has also been widely used for modelling land use (Pijanowski et al. 2002, Mas et al. 2004). Compared to other statistical methods of land use change such as logistic regression, for example, ANN has the advantages in incorporating the correlation within complex spatial variables , that failed to be accommodated by logistic regression method (Li and Yeh 2002). A brief example of the advantages of using ANN was explained by Mas et al. (2004) when illustrating the work of Sader and Joyce (1988), who studied the forest area change process and its association with the slope and road network in Costa Rica. They found that in the first phase, forest clearing had a strong relationship with proximity to road, and that the deforestation on shallow slope before 1961 was low. This condition occurred because the 0-5% slopes in the study area were inaccessible due to the absence of the road network. By the time, deforestation on the shallow slope that previously was less accessible increased by the construction of a highway (Sader and Joyce 1988, Mas et al. 2004). Thus, Mas (2004) illustrated that forest clearing process is not the result of the sum of the effects of each factor in an independent form, but rather the combination of them. This aspects, according to Mas (2004), were failed to be included in traditional statistical modelling of land use, in which ANN can accommodates them (Mas et al. 2004)

The architecture of ANN was adapted from the model of biological neural network, which capable to recognize and memorize patterns from their interactions with the environment. Multi-layer perception is one of the most widely used architectural networks of ANN (Rumelhart et al. 1985). The simplest ANN construction consists of three layers, although additional layers can be added. The layers are one input layer, one output layer and at least one hidden layer located between the input and the output layer. Next, a set of observation X consist of n different variables is used as the input data, and connected directly to the input layer. Each layer in the network is interconnected with the preceding and following layer through the nodes or neurons, which act as processing elements of the layers (Murata et al. 1994, Hsu et al. 1995, Li and Yeh 2002, Lee et al. 2004). In the input layer, each node will receive a single value which corresponds to an element in X, and each node generates output that will be used as input in the next layer. Therefore, nodes within a layer are not interconnected as well as nodes in non-adjacent layer, and each layer will only connect to the adjacent layer. The interaction between nodes is denoted with the weights, that are used to address the strengths of the network interconnection between associated neurons (Hsu et al. 1995, Eastman and Toledano 2000, Lee et al. 2004).The algorithm will calculates the weights for each node in the input layer, hidden layers and output layer. The signals generated from node to node are modified according to the weight associated with each connection, and the later node will sum all the weighted signals from previous connected nodes, and will be treated in the same way for the consequent nodes as “activation function”. The data moves forward from node to node with multiple weighted summation process before ending at the output layer (Pijanowski et al.

2002). Next, the learning procedure in ANN architecture is based on the concept that if the network gives the wrong answer, then the weight is corrected so the error is lessened and the next responses from the layers are expected to be correct. This concept, however, raises a problem in ANN method. As the network is designed to minimize the error of training set, this condition cause of over fitting on the model, and the more layers with more complex function is likely to be over fitted (Mas et al. 2004). Figure II.1 illustrates the architecture of the ANN with the example of three hidden layers.

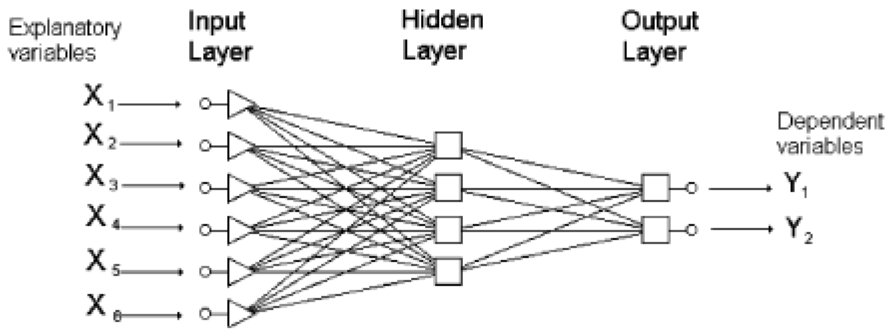


Figure II.1 Schematic illustration of the three layers in ANN. In land use change modelling, explanatory variables X represent the variables that are considered as drivers and constraints of land use change, while dependent variables Y represent the output of land use changes, for example area change and non-change (Mas et al. 2004)

II.5 Markov Chains Method

Markov Chains method has been widely used to model land use change. This method is commonly used to predict geographical characteristics, and has become an important prediction method in geographical research (Jokar Arsanjani et al. 2013). In general, Markov Chains is a technique for estimating using a probability matrix that record the probability of each transition. This method is commonly used to project future change of land use in urban or non-urban area at various spatial scales (Bell 1974, Weng 2002, Wu et al. 2006).

The core principle of Markov Chains method is a continuation of historical development. This method uses matrices to represent changes between category (Koomen and Borsboom-van Beurden 2011). Markovian analysis work based on the assumption that the next stage depends only on the current stage and the previous stages does not influence the outcome, and the model only influenced by the amount of training set received from the current state (Aaviksoo 1995, Myint and Wang 2006). This analysis uses matrices to represent changes between categories. The matrix cell values are derived directly from the input maps and indicate as proportions of land use classes. Under the assumption that the sample is a representative of the whole study area, the proportional changes become the probabilities of land use change and for the transition matrices. Markov chains method also work under an assumption of stationary condition, which mean that the model assume that temporal rate of change and amount of change stay the same(Muller and Middleton 1994, Koomen and Borsboom-van Beurden 2011). Markov chains method has proven to give better results compare to the simple linear extrapolation from the side of the possibility in including all transitions between the types and in allowing prediction in long periods (Aaviksoo 1995). Another advantages of the Markov Chains method in predicting land use are the simplicity of the logical operation and efficiency in computation, even though the assumption of the stationary of the transition matrix failed to consider the spatial dependency information in the model (Myint and Wang 2006).

In land use modelling, Markov Chains method is popular to be combined with other statistical analysis method (Koomen and Borsboom-van Beurden 2011). For example, CA-MARKOV model integrates Cellular Automata (CA) method with Markov Chains to model land use change and to predict future change (Li and Reynolds 1997, Pontius and Malanson 2005, Yang et al. 2012). Tang et al. (2007) also combines traditional Markov Chains model with Modified Genetic Algorithm (Modified GA) to model the spatio-temporal urban change in Daqing City, China. Furthermore, Markov Chains was also used with ANN modelling for prediction purposes (Bohling and Dubois 2003, Razavi 2014). Most of this integration using statistical methods (such as Modified GA, CA and ANN) to calculate Transitional Probability Matrices (TPM), which are required to describe the Markov Chain behaviour, while Markov Chains was used to predict future condition. (Tang, Wang et al. 2007)

III. MATERIALS AND METHODS

III. 1 Study Area

We choose three provinces in Indonesia for our study area, Riau, East Kalimantan and West Kalimantan. These provinces are located in the two main islands in Indonesia. Riau is part of Sumatera, while East Kalimantan and West Kalimantan are in Kalimantan. Oil palm has been cultivated commercially in Indonesia since 1911 under the Dutch Administration, which established plantations along the east coast of Sumatera Island (Budidarsono et al. 2013).During the 1970's, palm oil became one of the fastest growing agriculture commodities around the globe, including Indonesia (Koh and Wilcove 2008, USDA 2010, Wicke et al. 2011, Susanti and Burgers 2012). Since the 1970's the Indonesian government extensively expanded oil palm plantations in Sumatera and then into Kalimantan and Papua by 1990 (USDA 2010). Riau province is the largest palm oil producer in Indonesia, and has the largest oil palm plantation area within the country; therefore we choose this province as our study area. We choose East Kalimantan and West Kalimantan, as after Riau had been developed for oil palm, expansion into these provinces occurred. The choice to compare West Kalimantan and East Kalimantan was mainly based on the topographical differences between these provinces. West Kalimantan has relatively flat topography, while East Kalimantan has a coarser surface dominated with highlands and mountainous areas. These historical and topographical differences were expected to represent variation in proximate and underlying drivers of deforestation related to agricultural expansions. Figure III.1 below shows the location of the study area, and a general overview of the three provinces is given below.

Figure III.1 The study area in Riau, West Kalimantan and East Kalimantan. Riau is located in Sumatera island, and the largest oil palm producer in Indonesia. West Kalimantan and East Kalimantan are located in Kalimantan Island, and oil palm has been expanded to these provinces after Riau.

III.1.1. Riau

Lying between 01˚05' South Latitude and 02˚25' North Latitude, and between 100˚00' – 105˚05' East Longitude, Riau covers an area of 8,915,016 ha. Riau has a wet tropic climate with the average of rainfalls between 1700-4000 mm per year, varied between dry season and wet season. Riau has a relatively flat area with more than 47% of the area has slope less than 2%. Forest covers 35% of the total area, and categorized as conversion forest, protected forest and production forest. Tesso Nilo National Park and Bukit Tigapuluh National Park, two of the protected areas in Indonesia, also located in this province. In the economic sector, agriculture play important role in contributing regional revenue, and mostly dominated by oil palm as the main commodity. This commodity occupied for almost 91% of the total plantation in Riau (BPS 2014), and mentioned as responsible for the 7% for the total forest loss in Indonesia (Gelling 2007, Nawir and Rumboko 2007).

III.1.2. East Kalimantan

East Kalimantan located in 2˚33' North Latitude to 2˚25' South Latitude and 113˚44'-119˚00' East Longitude, and has the area of 127,267.52 km^2. East Kalimantan has rough topography and dominated by upland area, especially in the western part of the island, where more than 36% of the area is located at altitude for more than 500 m above sea level and around 73% of the area lies in the slope more than 15%. East Kalimantan also has large forest covers, which cover almost 61% of the total area. From this forests area, in 2013, around 1,698,171 ha are designated as protected areas and 468,503 ha as conservation forests. Kayan Mentarang National Park, one of the last remaining primary forests in South East Asia, is located in this province.

In agriculture sector, palm oil is the largest commodity, followed by rubber and coconut. From these plantations, almost 78% of them are large scale, the rest 22% are smallholder plantations. Oil palm has also become the third largest to contribute in regional income after mining and tax sectors.

III.1.3. West Kalimantan

Covering an area of 146,807 km^2, West Kalimantan lying along 2˚08' and 3˚02' South Latitude and 108˚30' – 114˚10' East Longitude. Located along the equator, this province has a tropical climate with high temperature and humidity. Topography in West Kalimantan is dominated by lowland area and some highland that lies along the Kapuas stream from the western to the eastern part of the region. There are also peat lands, swamp bogs and mangrove forests in some part of the low lowland. Most of the area is dominated by yellow-red podsolete or degraded lateritic soil type, which covers an area of 9.4 million ha, or almost 64% of the total area. Along the coastal line, the soil type is dominated by alluvial soil for almost 23% of the total area. Forests cover for almost 46% of the total area, and agriculture occupies for around 12%, while settlements occupy less than 1% of the total area. Several protected areas, such as Betung Kerihun, Bukit Baka, Danau Sentarum and Gunung Palung, which are designated as National Parks, also located in this province. This province has a population of 4,550 million, with a population density of 32 inhabitants/ km^2. The population are spread unevenly, mostly concentrated in coastal area. In agricultural sector, the main commodity of this province is palm oil, and usually being cultivated in the form of large plantations or smallholder plantations. The report from the National Statistical Bureau (BPS) also mentioned that during 2012, agricultural sector has the highest realization of capital investment, and plays important roles in increasing the regional revenue (BPS 2014).

III.2. Materials

We used several datasets in this research. Datasets were obtained from different sources with different format and also different spatial and temporal resolution. Some datasets are free open source data, and some datasets were available at National level provided by the Government. The description below explains the datasets used in this research.

Land cover dataset was obtained from the Ministry of Forestry (MoF) of Indonesia, and derived from the classification of mosaic Landsat TM/ETM+ data with 30 by 30 meters of resolution. This land cover map was classified into 22 land cover classes. This dataset is available in vector format; depict the polygons of land use for the year of 2000, 2003, 2006 and 2009. The entire classification process was undertaken by the forest management unit of Indonesia (BPKH) at sub-national levels, and then compiled by the MoF. (Romijn et al. 2013).

Oil palm dataset was downloaded from http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html. This website provides the global extent of various crops based on FAO classification and can be freely downloaded. The dataset is available in raster format and represent the areas of individual crop type for the year of 2000 at 5 min by 5 min spatial resolution in latitude by longitude, or approximately 10 km by 10 km. The data set is the combination of a new gridded map of global croplands for the year of 2000 with agricultural statistics representing individual administrative units throughout the world (Monfreda et al. 2008). Agricultural data were collected through census and survey information for desirable crop type from the smallest political units, which were mostly available at sub national level (state/ province and country/ district). For Indonesia, the data were collected from the Food and Agriculture Organization's (FAO) statistical database. The dataset represent single value of proportion area of certain crop type, which was calculated from the average of individual census from the year of 1990 – 2003, to get a single representative value for circa the year 2000. Each single crop map represents the proportion of certain crop area in each grid cell, which was calculated as:

$$fcrop_{xy} = fcropland_{xy}\left(\frac{crop\ pu}{cropland\ pu}\right)$$

Where $fcrop_{xy}$ is pixel value represent the proportion of certain of crop type, $fcropland_{xy}$ is the total cropland in certain pixel, $crop_{pu}$ is the number of certain cropland area in the inventory data and $cropland_{pu}$ is the total of cropland in the inventory data (the smallest political unit where the census data were available).

Slope map is available in grid raster format, and was generated from mosaic DEM (Digital Elevation Model) which was downloaded from ftp://xftp.jrc.it/pub/srtmV4/tiff/. The original slope map has spatial resolution of 90 by 90 meters. The slope is represented in percentage value, and projected in World Molleweide spatial reference (Carter and Stuiver 2014). We include this variable as input in this research as we consider that slope influences the process of land use change, regarding to the accessibility factor and suitability for certain plantation type.

SRTM map is an elevation map downloaded from http://srtm.csi.cgiar.org, and available in arc grid format with decimal degree projection system using datum WGS 84. The data was derived from the USGS/ NASA SRTM data, and being processed to provide seamless continuous topography surfaces (Jarvis et al. 2008). We include elevation in our study as elevation might determine the suitability of certain crop type to grow, which potentially influence the process of land use change regarding the agricultural expansion.

World Protected Area Map was downloaded from www.protectedplanet.net, which provides online access to the World Database for Protected Area (WDPA). This portal is a joint project of IUCN (International Union for the Conservation Nature) and UNEP (United Nations Environment Programme), and provide world protected areas maps over the worlds (IUCN and UNEP-MCMC 2014). The material downloaded for this research is 2014 version. Considering that protected area is designed for conservation purpose, we include this variable as a constraint of deforestation, with the assumption that there will be no change occurred in the protected areas.

Proximity distance to existing oil palm plantation

It is considered that the area located near existing oil palm plantations has high potential to change into oil palm plantation, following its neighbourhood land use. Therefore, we incorporated proximity distance from existing oil palm plantation as driver parameter of agricultural expansions. Proximity distance from existing oil palm plantation is provided as a distance map. We built distance map using Euclidean Distance tool based on the existing oil palm polygons generated from the land use map from the MoF for the year of 2000. The map provides information about relative distance from the existing oil palm polygons, and was formatted in 32 bit floating pixel type.

Proximity distance to existing settlements

We involve distance from settlements in our study, as we expect that deforestation and the opening of new oil palm plantation has correlation with the existing residential area, regarding to the accessibility and economic activities factors. Distance map was produced using Euclidean Distance. Settlement polygons were obtained from the land use map from the MoF for the year of 2000. Table III.1 below shows the resume of the datasets we used in the research.

Table III.1 List of the datasets, the abbreviation used in the report, format and source of the data

Name of data	Abbreviation	Source	Resolution	Source type
Crop map of oil palm dataset	OP dataset	http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html	5 min by 5 min or approximately 1km x 1km	Open data
Land cover/ land use map of Indonesia	LU Map	The Ministry of Forestry Republic of Indonesia (MoF)	30m x 30m	Data confidential
Global elevation dataset	Elevation Map	http://srtm.csi.cgiar.org.		Open data
Global slope dataset	Slope Map	Carter, S., Stuiver J, 2014, Slope derived from SRTM data, v 1, Wageningen University & CIFOR.	90 m x 90 m	Open data
Global Forest Protected area	PA map	The World Database of Protected Areas	Vector dataset	Open data
Distance from existing oil palm map		Derived from land cover/ land use map from the Ministry of Forestry	30m x 30m	Derived from confidential data
Distance from existing settlement area		Derived from land cover/ Land use map from the Ministry of Forestry	30m x 30m	Derived from confidential data

III.3. Methods

An overview of overall methods is shown in Figure III.2 below, while the following sub-chapters described the software and methods.

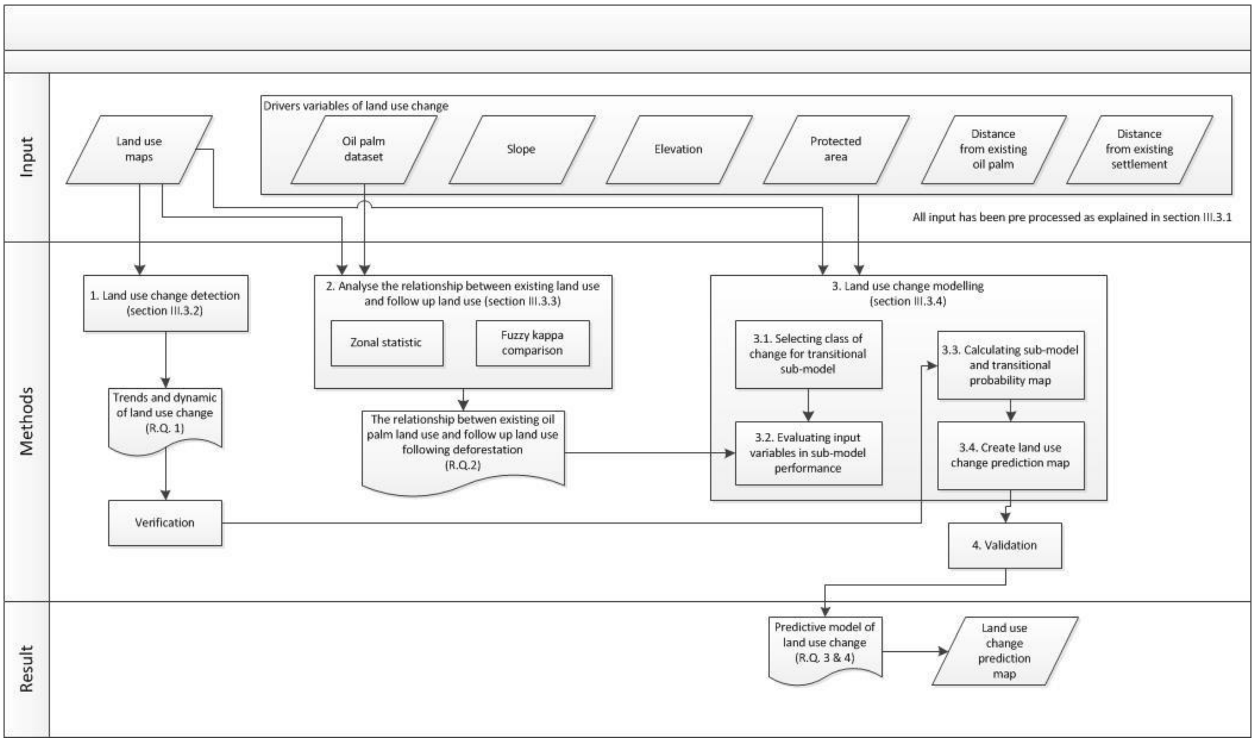


Figure III.2 Research flow chart

III.3.1. The software

We used three software in this research. We used Arc Map version 10.2.2 with spatial analysis extension for most part of the change detection and pre-processing step before modelling section. To build land change model and prediction, we used Idrisi Selva software with Land Change Modeller (LCM) extension, developed by Clark Lab University (http://www.clarklabs.org/products/Land-Change-Modeler-Overview.cfm). We also used Map Comparison Kit (MCK) software version 3.2 to compare oil palm dataset (OP map) with land use map (LU map) to see the degree of relationship between the two maps, as part to answer research question 2. This software was developed by the Netherlands Environmental Assessment Agency (PBL), in cooperation with the National Institute for Coastal and Marine Management (RIKZ) and the European Commission, and has been widely used for map comparison and validation in land use modelling (Visser and de Nijs 2006, BV 2010).

III.3.2. Datasets preparation

The purpose of datasets preparation was to provide data with required format in the next process. We cropped all maps for the area of interest and projected it with the same projected coordinate system. We also resample all raster datasets into 30x30 m pixel size; correspond to the Landsat ETM pixel size in which land use maps (LU maps) from the MoF were derived. We converted all vector based maps into raster format with 8 bit unsigned pixel type for nominal and categorical maps such as LU maps and protected area map (PA map), and 32 bit floating type such as slope and oil palm datasets. We also set all maps that will be used in IDRISI modelling so that they have the same spatial dimensions, including pixel size, projection system, spatial extent, spatial resolution and number of rows and columns. All datasets were exported into IDRISI software in ASCII format; therefore we converted all maps into ASCII format.

To simplify the analysis of land use change and dynamics in the study area, we aggregated data attributes from the MoF into five classes, as shown in Table III.2 below. These aggregated maps were then being used for the next analysis.

Table III.2 Aggregated data attributes from the MoF dataset

Aggregated category	Criteria	MoF data Classes
Forest	FAO forest definition –include degraded areas/ secondary forest, forest plantations and forest in the swamp or peat land area.	Primary upland forest, Secondary upland forest, Primary mangrove forest, secondary mangrove forest, crop forest, bushes/shrubland, primary swamp forest, Secondary swamp forest, swampy bush
Agriculture	Farming/grazing activities including matrix rural settlement/agriculture land uses	Upland farming, upland farming mixed with bush, transmigration, rice field
Oil Palm		Plantation/garden
Other	Natural or anthropogenic landscapes which are not forest or agriculture	Savannah, cultured fisheries/fish pond, settlement/developed land, transmigration, open land, mining/mines, water body, swamp, airport/harbour
Cloud	No data	Cloud

III.3.3. Land use change and trend detection

We mainly used Arc Map 10.2.2 software to identify the actual change and dynamic of land use. We identified the trend and dynamic of land use change including deforestation events by identify forest loss polygons for the period of 2000-2003, 2003-2006 and 2006-2009 based on LU Map from the MoF. We identified deforestation events and its follow up land use by overlaid the LU maps for two consecutive years, and presented the result in form of land use change maps and graphs. As an addition, change detection was also carried out in IDRISI Land Change Modeller (LCM) using change analysis menu, as this part was required for modelling step. The methods of change detection in IDRISI will be described in section III.3.5.

III.3.4. Analyzed the relationship between existing oil palm map (OP map) and follow up oil palm land use after deforestation

The main process of this step was to identify the relationship between OP map and actual LU map from the MoF. OP map represents the density of oil palm area in each pixel, while LU map was derived from remote sensing data and has been validated. We assumed that the new plantations were likely to be opened near the existing oil palm area, thus we can use OP map as a reference to detect the next expansion of oil palm. Therefore, we compared OP map with 2000 LU map from the MoF as the starting point. We also compared the change into oil palm that was detected from the LU map with the distribution of oil palm based on OP map, to see whether the pattern of change was interrelated with the pattern of oil palm density.

We implemented three steps to answer this question. First, as a starting point, we measured the fitness of OP map with the existing oil palm area from the LU map in the year of 2000 by overlaid the two maps in Arc Map. For each overlaid polygon, we calculated the mean value of the pixels from OP map using zonal statistic tool. Considering that small polygons will contribute less to the total mean value, we normalized the mean value by weighting it. Each mean value resulted from the zonal statistic calculation was multiplied by the area of correspond polygon and divided by total area. The equation is:

$$X = \frac{\sum_{i=1}^{n} xi.wi}{\sum_{i=1}^{n} wa}$$

Where X is the weighted mean value for all polygons in the same class, *xi* is the mean value of each polygon, *wi* is the area of the polygon and *wa* is the total area of polygon in the study area.

We calculated weighted mean of the OP map for all land use classes from the MoF, and compared the results.

Secondly, we observed the similarity between OP map with the 2000 LU map using the fuzzy kappa comparison method, as an additional step to measure the degree of similarity between the two maps. This method is useful to assess the similarities and dissimilarities of two maps, and give the possibilities to assess maps with different legends (BV 2010). Rather than comparing pixels as true and false when they are assigned as the same or different class, this method calculates the grade of similarity between pairs of cells in two maps (Hagen 2003, Hagen-Zanker et al. 2005, Visser and de Nijs 2006, Ahmed et al. 2013). We compared OP map with 2000 LU map to see the grade of similarity between them. We used MCK (Map Comparison Kit) software version 3.2 to calculate the fuzzy kappa value. First, we classify OP map into three classes, which are low, medium and high. Then, we convert the reclassified OP map and 2000 LU map in ASCII file to be imported to MCK software. We used fuzzy kappa algorithm method as this algorithm can be used to compare two maps with different legends. We used the default algorithm setting from MCK, with a radius of neighbourhood as 4 and exponential decay halving distance as 2. A similarity matrix that depicts the proximity closeness of each class should be determined for the calculation, therefore we used the estimated value obtained from the previous step when we compared the similarity of each class in LU map with the pixel values in OP map. We determined that high class in OP map will correspond to oil palm class in LU map, while medium class will correspond with other and agriculture classes from land use map and low class related with forest class. We used classification of strength of agreement from Sousa (2002) for kappa index as reference to justify the similarity of the maps as low, medium and high (Sousa et al. 2002). Table III.3 below shows the similarity matrix that we used in Fuzzy Kappa algorithm setting in MCK.

Table III.3 Similarity matrix in fuzzy kappa comparison. The values represent the degree of similarity for pair of classes. The value of 0 means that the classes are completely not similar, while the value of 1 means that the classes are similar.

Map 1/ map 2	Forest	Agriculture	Oil palm	Other
Low	1	0	0	0
Medium	0.5	1	1	1
High	0	0.5	1	0

In the third step, we identified the possibility of using OP map to estimates the next expansion of oil palm plantation based on its distribution pattern of pixel values in the changes related to deforestation and oil palm. We identified pixel values for the polygons that experienced change related to deforestation and oil palm expansion. We observed the possibilities of certain pattern of land use change based on the pixel values from OP map. We considered changes related to the deforestation like forest to agriculture, forest to oil palm and forest to other, and also changes related to the expansion of oil palm like other to oil palm and other to agriculture for the year of 2000, 2003, 2006 and 2009. To see the possibilities of the relationship between the extents of change with pixel value in OP map, we also measured the coefficient correlation of each change polygon. The correlation coefficient was calculated with the equation below, where r is the correlation coefficient, x is the pixel value, x is the mean pixel of all polygons, y is the area width and y is the mean area of all polygons.

$$r(x,y) = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sqrt{\sum(x-\overline{x})(y-\overline{y})}}$$

III.3.5. Modelling land use change

Modelling land use change was done in IDRISI Selva LCM. This part is mainly answer research questions 3, 4 and 5. Some steps in modelling part can be used as complementary explanations to answer question number 1 and 2. In this research, we divided modelling process into three parts, which are observation phase, prediction phase and validation. In observation phase, sub-models were built, which model the change of land use during this period. We used 2000 and 2006 LU maps derived from the MoF dataset as the input. This step produced Transitional Probability Maps for each change; depict the probability of each pixel to change into certain land use. We used Multi-Layer Perception Artificial Neural Network (ANN) method to build the Transitional Probability Maps. In prediction phase, a prediction model was generated for the year of 2009 based on the Transitional Probability Maps produced in the previous step, and a predictive model for the year of 2009 was produced. The prediction model was generated using Markov Chains method (Michalski et al. 2008, Rajan 2010, Ahmed and Ahmed 2012). In validation phase, we used 2009 LU map from the MoF to validate our prediction map. We implemented per category validation method and three maps comparison method as our validation techniques (Pijanowski et al. 2002, Mas et al. 2004). The description below explains the steps in land use modelling, while validation step was described in the next sub-chapter.

1. Importing land use maps and drivers variables
    a. LU maps for the year of 2000, 2006 and 2009
    LU maps from the MoF were used as the basic maps in the modelling part. We used 2000 and 2006 LU maps as the input in observation phase, while 2009 LU map was used during validation process. All LU maps were converted into ASCII format before being imported to IDRISI format (.rst). Once the maps have been imported, new legends should be assigned so that all land use maps will have the same legends with the same category for each nominal value.
    b. OP map
    As the map depicts a certain amount of oil palm area in each pixel, this dataset was treated as density map in land use modelling. Therefore, we calculated the class frequency of the map as density metric. Because this metric will be used to update the dynamic variable at a particular interval or stage, a macro file should be built and be assigned as a basic operation in transitional sub-model structure.
    c. Elevation map
    Elevation map was imported in real integer file format.
    d. Slope map
    Slope was imported in real binary format.
    e. Proximity distance to existing oil palm plantation map
    Proximity distance to existing oil palm plantation map was imported as real binary format.
    f. Proximity distance to settlements map
    Proximity distance to existing settlements map was imported as real binary format.
    g. Protected area map (PA map)
    The protected area was assigned as a constraint in land use change events, and we assumed that this area will not change over time. Existing PA map is available in polygon format, therefore we transformed it into raster format. We reclassified protected areas as 0 and non-protected as 1, since the pixel value will act as multiplier in the transitional sub model calculation. LCM also required categorical data to be converted into a set of Boolean (dummies) variables, or transformed using the Evidence Likelihood transformation method. Therefore, we performed evidence likelihood transformation for PA map, and this process was done in Variable Transformation Utility tools in IDRISI LCM (Eastman 2006). Considering the changes that might occur in protected areas for some reasons, this step also recalculated the probability of each pixel to change based on the real changes that already occurred, and assign new value on it (Camacho Olmedo et al. 2013).

2. Assigning land use change analysis
    Change analysis was implemented before the sub-model was created. In this step, we calculated changes during the observation period. This step produced the quantification of gain and losses for each land use and contribution of each class in overall land use changes. As an additional step, we also implemented trend detection. Trend detection provides an overview of generalized

pattern of specific change based on its spatial distribution. The pattern was derived by giving a value of 1 for change area and 0 for unchanged area, and treating them as if they were quantitative values. The numeric value produced did not show the real trend value, since it only shows the generalization about the pattern of change (Eastman 2006). We implemented 3rd order of polynomial to calculated the trends, and the result was a range of integers shows the degree of changes within the area, with the higher value means the higher changed occurred in that direction.

3. Determine class of changes that will be included in sub-model structure
   All changes were presented in variable transformation utility panel. Nonetheless, not all changes were included in sub-model structure, since we only interested in changes related to deforestation events and oil palm expansion. We choose Forest – Agriculture, Forest – Oil Palm, Forest – Other, Other – Agriculture and Other – Oil Palm as our sub models. It is important to note that some changes did not occurred in certain area, so the number of changes were different in each study area. Even though it is available to regroup changes into one class if it is considered to have the same drivers, we did not regroup the classes because we presumed that each sub-model has different driver's parameters.

4. Determine the drivers and constraints variables in sub-model structure
   Before the sub-models were calculated, a transitional sub-model structure should be designed, to determine the driver parameters that will be used for each sub-model and its role and operation during the sub-model calculation. All maps that were imported to IDRISI except LU maps were assigned as the variables of land use change. We also determined the role of each variable as dynamic or static variables. Static variables express the aspects of basic suitability for the transition under consideration, and are unchanging over time, while dynamic variables are time-dependent drivers that will be recalculated over time during the course of a prediction (Eastman 2006). For dynamic variables, basic layer type and operation should be specified. Table III.4 below summarize the explanation of driver variables specification.

Table III.4 Drivers and constraint variables and its role in the sub-model

Name of variables	Role	Basic layer type	Operation
Distance from existing oil palm plantation	Dynamic variable	Land use	Distance
Distance from existing settlement	Dynamic variable	Land use	Distance
Oil palm map	Dynamic variable	Land use	Density with frequency class pattern
Slope	Static variable	-	-
Elevation	Static variable	-	-
Protected area	Static variable	-	-

5. Built the sub-models of land use change for the year 2000 – 2006 (observation phase)

   Sub models were needed prior to predictive model, and were done for all classes of change which have been determined in step 3. The sub-model calculation generates the probability of each pixel to be converted into certain land use. We used Multi-Layer Perception Artificial Neural Network (ANN) method to calculate the sub-models. 2000 and 2006 LU maps were used as model input. Several independent variables which were already constructed in sub-model structure were used as the variables, and the performance of each variable was evaluated during sub-model calculation.

   Some parameters were set to generate the best model that represents the actual situation. The most significant parameters were the number of hidden nodes and learning rate. Dynamic learning rate was set as automatic. The initial set up for each sub-model were three hidden layer nodes with the sigmoid constant of 1.0, the momentum factor of 0.5, starting learning rate as 0.01 and end learning rate as 0.0001. We set stopping criteria for running the model as 0.0001 for RMS error, 10.000 times iteration or accuracy of 100% as the first setting. However, if at the end of the iteration the error monitoring graph still shows the variation or giving low accuracy value, we increased the iteration times for a half than the previous until a certain amount and the curve on the graph shows no variation or the accuracy rate did not improve significantly. If at the end of the iteration there was no improvement of accuracy rate, we add the number of hidden nodes until it reaches the level where the accuracy increased with insignificant number. In general, the accuracy rate around 80% is acceptable (Eastman 2006). Figure III.3 illustrated on when the calculation was repeated with the adjustment criterion and when the result of calculation can be used to build the transitional probability maps.

a b

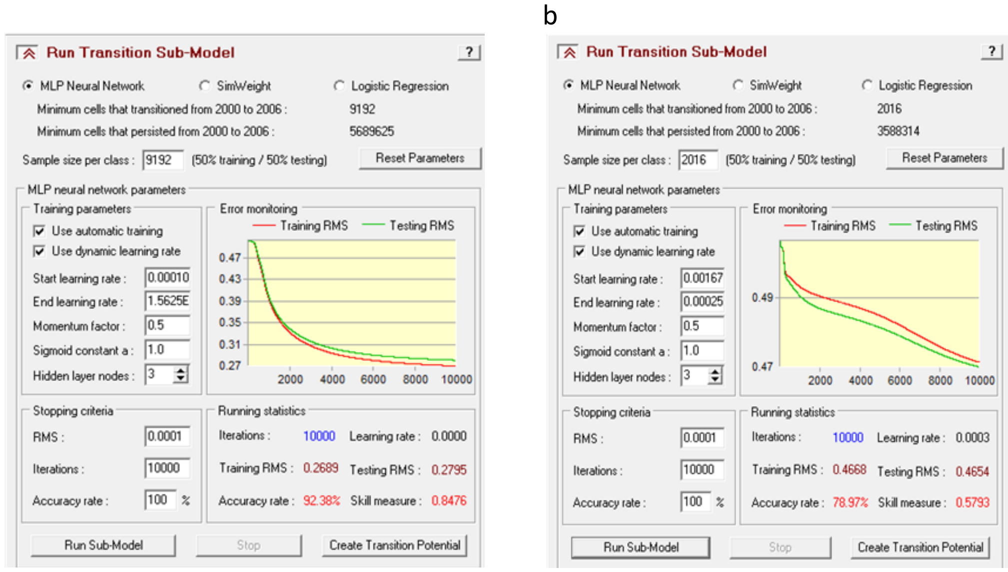


Figure III.3 The examples of the sub-model calculation result. a) The error monitoring graph shows a smooth graph and the accuracy rate was sufficient. Therefore, a transitional potential map can be calculated. b) Even though the accuracy rate was close to 80%, the error monitoring graph still shows a significant decrease. Therefore, we need to adjust the parameters and sub-model was calculated.

One of the result of this process was a document contains statistical explanations of the learning process. These explanations were useful to analyze the explanatory power of the independent variables. The document consists of three parts of the explanations: *general model information, weight information of neurons across layers and sensitivity of the model to forcing independent variables to be constant*. The first and second parts contain the general setting of the sub-model and the algorithm information of the sub-model. The last part, sensitivity of the model to forcing independent variables to be constant, was useful to judge the explanatory power of the variables which explain the performance of each variable in each sub-model. The relative power of each variable was tested by holding selected variables to be constant at their mean value, which will effectively remove variability caused by the variable concerned. Three different sensitivity analyses were carried out. First analysis was done by forcing one variable to be constant. The second was done by holding all variables constant except one, and the third was tested by developed the model using all variables, and then holds constant every variable in turn to determine which configuration of the variables has the least effect on model skill. If the result shows that there is no significant decrease of accuracy rate when a variable was held constant, it was considered that the concerned variable was less significant in the sub-model calculation. Therefore, this variable can be considered to be removed to reduce over fitting and to improve model skill. After the variables were tested and the selection was made on which variables that will be included in sub-model, the map of transitional potential or Transitional Probability Map was created. This process was repeated for all sub-models prior to prediction modelling.

6. Change prediction modelling (prediction phase)

   We used Markov Chains method to model 2009 predictive land use change. Transitional potential maps generated from the previous step were used to calculate transition probability matrix, which depict the probability of each land use to be converted into different class.

   There are two types of prediction produced by IDRISI LCM; hard prediction and soft prediction. Hard prediction calculates the prediction based on a specific scenario, and each pixel in the predicted map will be assigned as a certain class with the same categories as the input. In contrast, soft prediction calculates the vulnerability the pixels to change without specifying into what class it will change, but rather the degree to which areas have the right conditions to advance changes or the probability of each pixel to change (Eastman 2006, Camacho Olmedo et al. 2013). In this research, we only interested in hard prediction model, since our focus is to observe how the model predict certain changes, which was easier to be quantified rather than the soft prediction model.

III.3.6. Validation of the predictive model

We validated the prediction model using per category method by compared the two maps cell by cell based on a single class (Ahmed et al. 2013). We compared 2009 predictive map with 2009 LU map as an initial map. To assess the performance of the predictive model, we subtracting the

2009 predictive map with 2006 LU map to see how the predictive model predict the changes, and compared the result with the actual change at the same period.

To give more insight on the source of error in predicted model, we implemented three maps comparison technique. This technique allocates the accuracy as two classes: accuracy due to the persistence and accuracy due to the changes. This map comparison method allow the users to observe the origin of the error and depict the accuracy based on the change and persistence of land use (Pontius et al. 2004, Ahmed et al. 2013).Principally, three maps comparison technique is done by overlaying previous initial map (t_0), current initial land use map (t_1) and predicted land use map at current time (t_{1p}). Comparison between t_0 and $t_1$will characterize the observed change in the maps to reflect the actual dynamics of the landscape. Overlaid t_0 and t_{1p} will characterize the model's predicted change to reflect the behaviour of the model, while comparison of t_1 and t_{1p} will characterize the accuracy of the prediction. We choose this method as this technique is considered able to explain the misinterpretation of accuracy assessment due to pure persistence and changes (Pontius and Malanson 2005, Pontius Jr et al. 2008, Ahmed et al. 2013). In this research, t_0 was represented by 2006 LU map, t_1 was represented by 2009 LU map and t_{10} was represented by 2009 predicted map. We overlaid 2006 LU map, 2009 LU map and 2009 predicted map. The result generated 5 types of accuracy classes; consist of two components of agreements and three components of disagreement. The two components of agreements are persistence simulated correctly and changes simulated correctly, while the three components of disagreements are change simulated as persistence, persistence simulated as change and change simulated as change to the wrong category (Pontius and Malanson 2005, Pontius Jr et al. 2008).

IV. RESULTS

IV. 1. The trend and dynamic of land use change including deforestation event

We identify the trend and dynamic of forest change by investigating the change of forest polygon over the periods. We found that forest loss patterns in the study area were different within periods. In general, deforestation pattern in East Kalimantan and West Kalimantan increased during all the monitoring years, while in Riau, deforestation increased in the first and second period, but decreased in the third period. During 2000 until 2006, Riau has the highest clear cutting area amongst others, but in 2009 deforestation rate decreased significantly, from 262,064 ha into 190,778 ha, lower than East Kalimantan and West Kalimantan. In observation years, the pattern in East Kalimantan and West Kalimantan were steadily increasing, but during 2006-2009, West Kalimantan experienced intensive forest loss compared to the previous period. In this period, the area of deforestation in West Kalimantan increased 52% from earlier, put this province as the highest forest clearing area amongst others. Figure IV.1 below describe the trend of deforestation within periods.

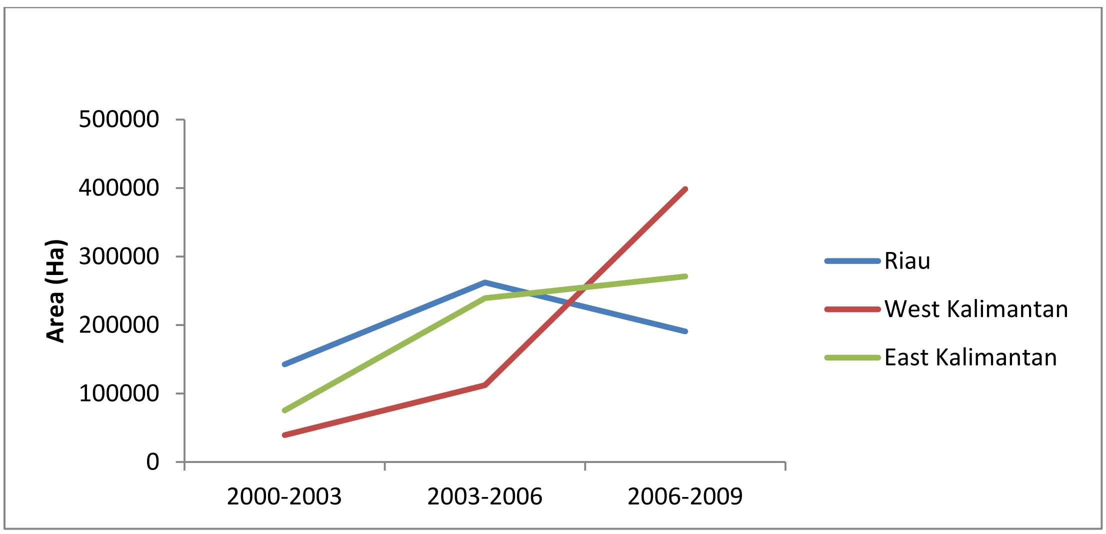


Figure IV.1 Annual deforestation area during 2000 - 2009 for the selected study areas in Indonesia. Data source:MoF

To get more insight on the trend of deforestation in the study area, we also observed followed land use after deforestation. We found that the trend was change during the observation and prediction periods. In Riau, deforestation was mostly followed by other class, while in East Kalimantan and West Kalimantan, this trend only occurred in the first period. In the first period, approximately 67% of the forest losses in East Kalimantan converted into other, while in the second and the third period, it was mainly turned into oil palm (42% in the second period and 72% in the third period). In West Kalimantan, in the first period, for about 44% of deforested area shifted into other, while in the second and the third period it were dominated by oil palm as for 65% in the second period and 47% in the third period. A significant increasing of agriculture class also showed in West Kalimantan during 2006-2009, with almost

41% of forest loss were converted into agriculture class. Figure IV.2 below illustrates the trend of shifting land use following deforestation, while Figure IV.3 shows the maps of forest change during 2000-2009.

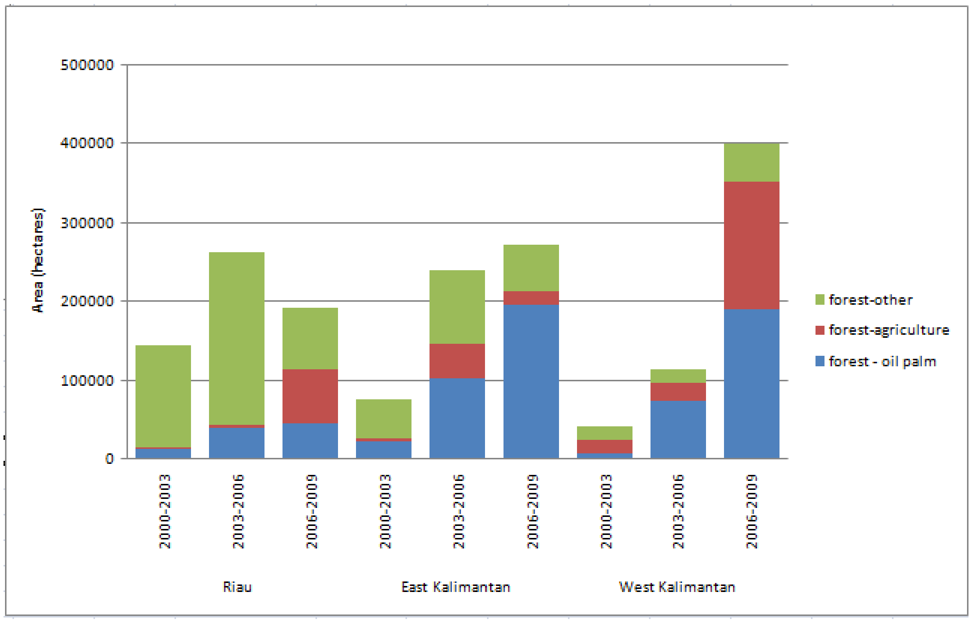


Figure IV.2 Trend of land use following deforestation. Data source: MoF

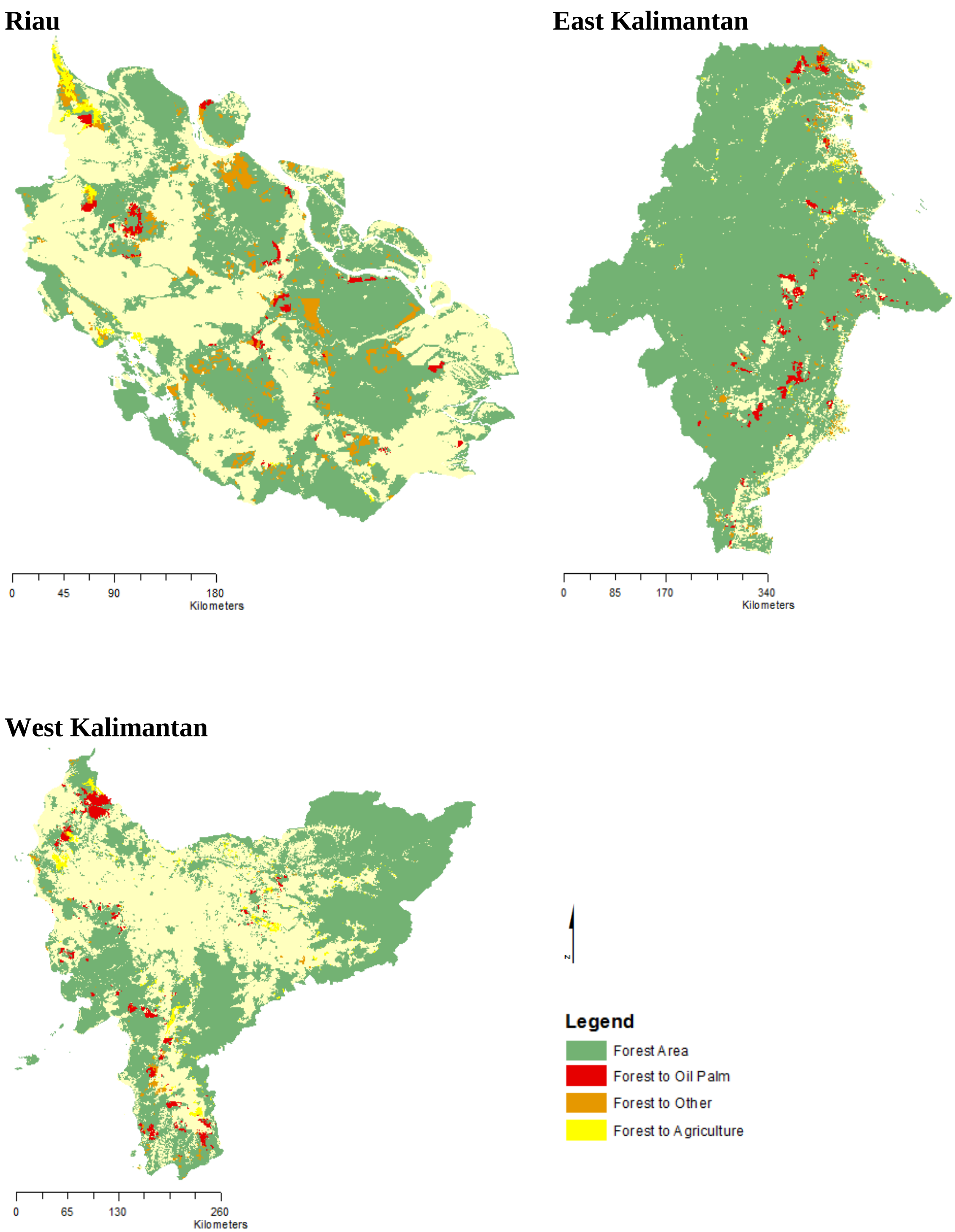


Figure IV.3 Area of deforestation during 2000 - 2009. Source: MoF

As other class was derived from the aggregation of several land use classes as shown in Table III.2, we consider that other class was sensitively change into different land use over the periods, thus we called it as intermediate changes that might be also important to explain the process of oil palm expansion. Therefore, we pay attention for deforestation in 2000 followed by other class in 2003, particularly in Riau, by disaggregated this class into the original classification from the MoF. We focus on Riau as during the monitoring years, forests loss were dominantly changed into other class, with 91% of the total change in the first period, 84% in the second period and 41% in the third period. We traced back the process of forest-other conversion from 2000 until 2009 in the same polygons. We found that forest-other polygons in the first period were all belong to open land class. In 2003, the same polygons shifted into oil palm (14%), bushes/ shrubland (14.4%), agriculture (0.9%), crop forest (19.55%) and the remaining as open land (50.7%). In 2009, approximately 90% of the remaining open land still remains as open land, while the rest 10% were converted into crop forest, oil palm, swampy bushes and agriculture land use. This condition implies that forest conversion can be followed by intermediate change before being converted into certain land use, including oil palm plantation. Figure IV.4 below illustrate the intermediate change of other class in Riau.

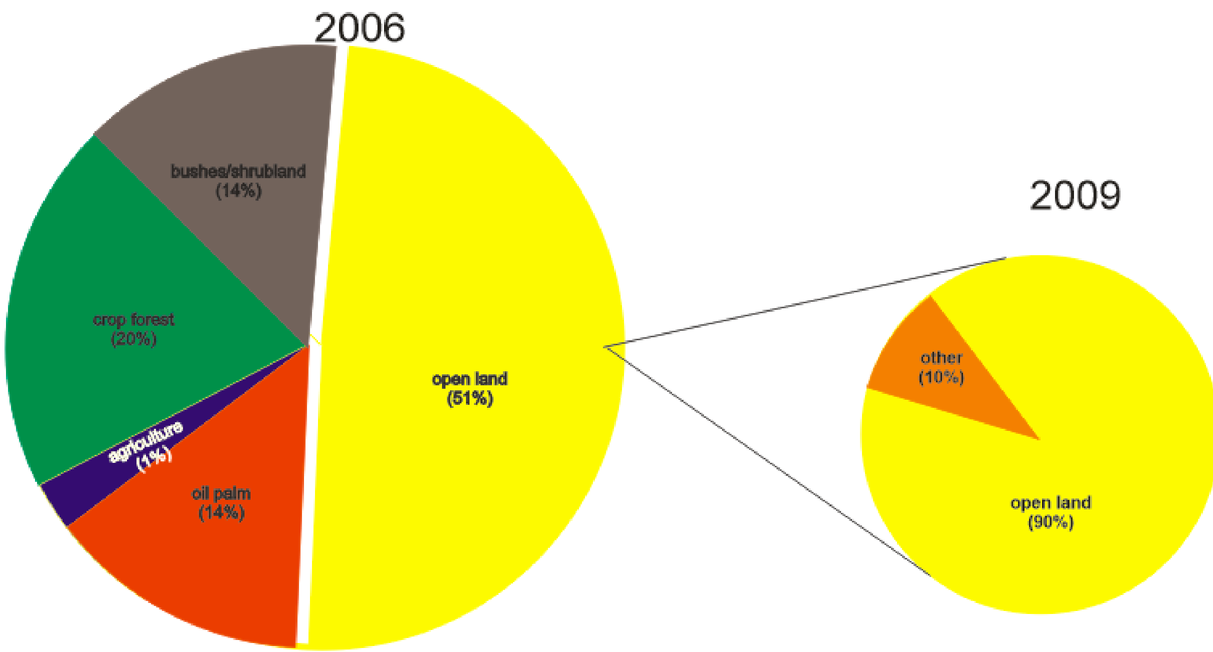


Figure IV.4 Conversion of other class/ open land during 2006 - 2009. This graph was derived from the deforested polygons that changed into other class/ open land in 2003. Data source: MoF

We also implemented change analysis in IDRISI land change modeller. This step was a part of land use modelling process. The result of this step was the quantification of gain and losses area of land use change, and show which class that has the most contribution in certain change. The result can be used to identify the contribution of each class in increasing or decreasing the other classes, even though the history of a single polygon cannot be traced individually. Figure IV.5 show the result of gain and loss quantification from IDRISI change analysis panel. The value is in hectare. A negative value means area loss, while positive values area obtained.

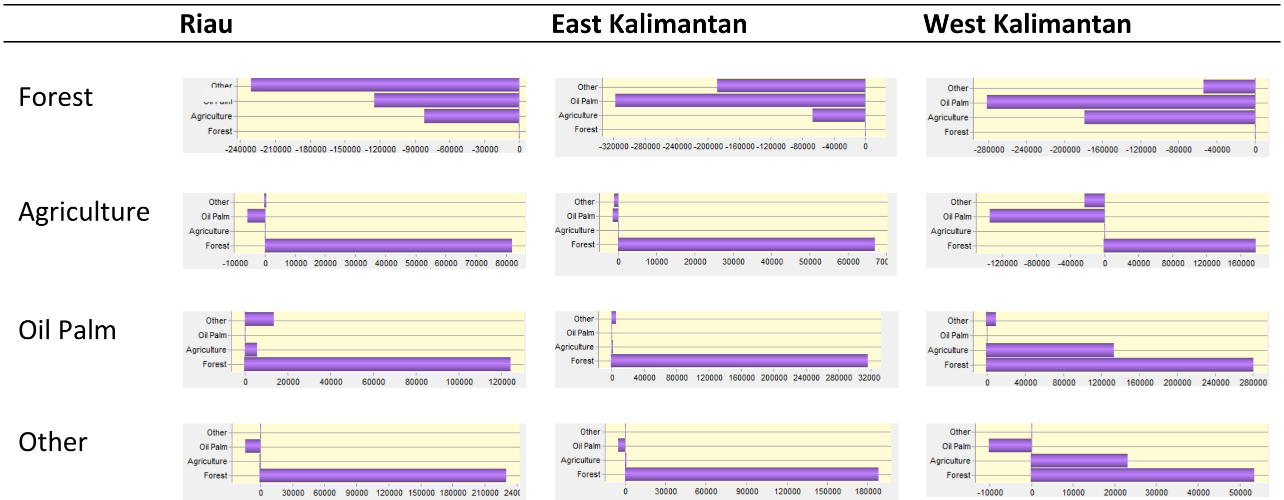


Figure IV.5 Gain and losses area of land use during 2000 – 2009. The result was derived from the LU maps produced from the MoF dataset, and was calculated using IDRIS LCM.

Figure IV.5 shows that in all provinces, forest contributes the most of increasing area of oil palm, agriculture and other. In Riau, forest conversions were mostly converted into other, while in East Kalimantan and West Kalimantan it was mainly converted into oil palm. The graphs also show that in the study areas, some of oil palm area were derived from other and agriculture class. Extensive conversion from agriculture into oil palm was found in West Kalimantan, covering an area of approximately 140,000 ha, and contributes to almost half of the total oil palm plantation. This trend of shifting cultivation also occurred in East Kalimantan and Riau, even though the number was not as extensive as in West Kalimantan.

We also produced trend maps of forest conversion into oil palm, which provide an overview of generalized pattern of changes based on its spatial distribution as shown in Figure IV.6. The red colour represents the high concentration of oil palm in that area, yellow represents medium concentration and the green colour represents low concentration of oil palm. The result shows that during 2000 – 2009, oil palm expansion in Riau were mainly concentrated in the middle part of the province, and spread throughout the area and covering along the west and east side of the province. In East Kalimantan, the trend tends to spread along the coastal area in the western part of the province and continuing from north to south. The trend tends to decreased in western part of the province, which was dominated by high land and steep slope area. Meanwhile, in West Kalimantan, trend of oil palm expansion starts to spread from the coastal line in the western part of the province, and was reduced with the increasing distance from the coastal area.

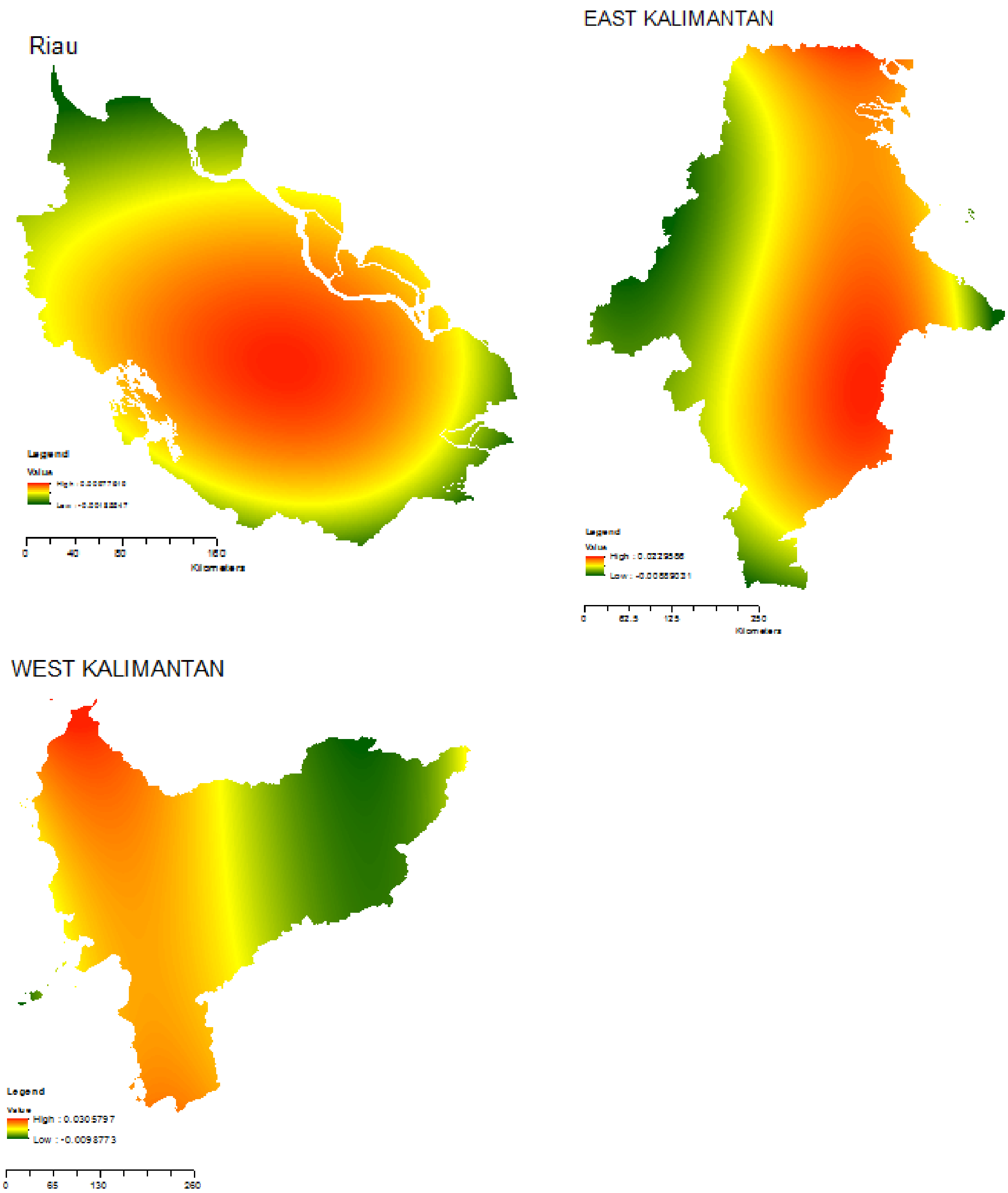


Figure IV.6 Spatial distribution trend of deforestation into oil palm during 2000 – 2009. The maps were derived from the LU maps produced from the MoF dataset, and was calculated using IDRISI LCM. The values in the legends did not represent the degree of change, but rather to explain the distribution of oil palm throughout the area.

IV.2. The relationship between existing oil palm map and follow up oil palm land use after deforestation

Table IV.1 shows the value of normalized mean pixels as the results of overlaid maps between OP map and LU map from the MoF for the year of 2000. The values indicate the percentage of oil palm based on the OP map in the correlated polygons from LU map. Figure IV.7 illustrates the example of overlaid of oil palm polygons from the LU map with the OP map. Lighter pixels represent high value of oil palm density, while darker pixels represent low value of oil palm density. It is important to note that because OP map was produced based on crop yield for every available administrative level, each study area has different maximum of pixel value of OP map, as represented in the right column in Table IV.1. The highest maximum value is in Riau, while East Kalimantan has the lowest maximum value.

Table IV.1 shows that in general, oil palm class has the highest value of mean pixels compare to agriculture, forest and other. We also noticed that the value of agriculture class only has small differences with oil palm class. Forest class has the lowest mean pixel value compare to other, and have significant differences with the other classes, especially in Riau.

Table IV.1 Normalized mean pixel value of OP map based on the aggregated LU map from the MoF dataset. The values were in percentage; depict the density of oil palm in certain pixel. The result was derived from the overlaid OP map with the LU map from the MoF for the year of 2000.

	Oil Palm	Agriculture	Forest	Other	Maximum value
Riau	6.93	6.69	2.72	5.41	12.4
East Kalimantan	0.71	0.48	0.32	0.51	6.7
West Kalimantan	2.32	2.31	1.00	1.70	10.7

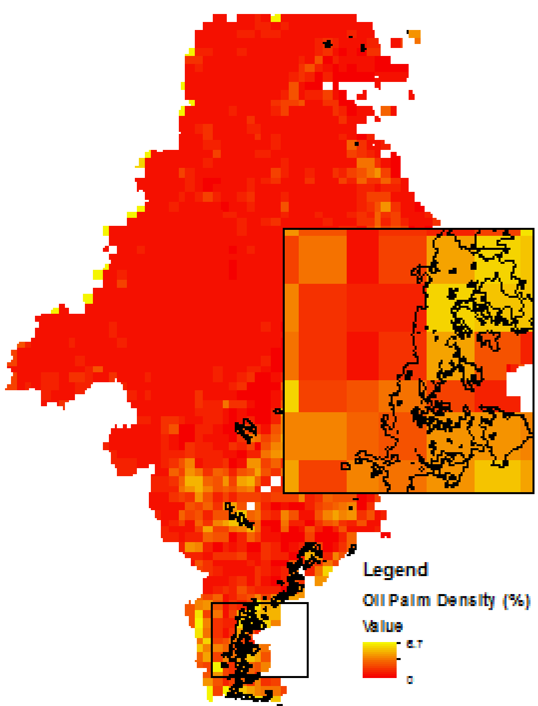


Figure IV.7 An example of overlaid LU map from the MoF with OP map.This map is an example of overlaid between oil palm class from the MoF in the year of 2000 with OP map for East Kalimantan. The legend indicates the density of oil palm. Darker colour represents lover pixel value, indicates that those area has low percentage of oil palm density based on OP map, while lighter colour indicates that those area has high density of oil palm.

Next, the result from the fuzzy kappa comparison also shows that all areas have low fuzzy kappa values. Riau has the value of 0.06 with the fraction correct of 0.497, while East Kalimantan has total fuzzy kappa value of 0.198 with the fraction correct 0.860. West Kalimantan has total fuzzy kappa value of 0.039 with fraction correct of 0.705. These values are considered low, indicated that the OP map has low agreement with the LU map for the year of 2000, even though the fraction correct in East Kalimantan and West Kalimantan are considered high. Nonetheless, the comparison map shows that the pattern between two maps has a higher degree of similarity in some area, as shown in Figure IV.8. The high similarity of the pixels, showed by high value of fuzzy kappa (green pixels) is usually found in forest areas.

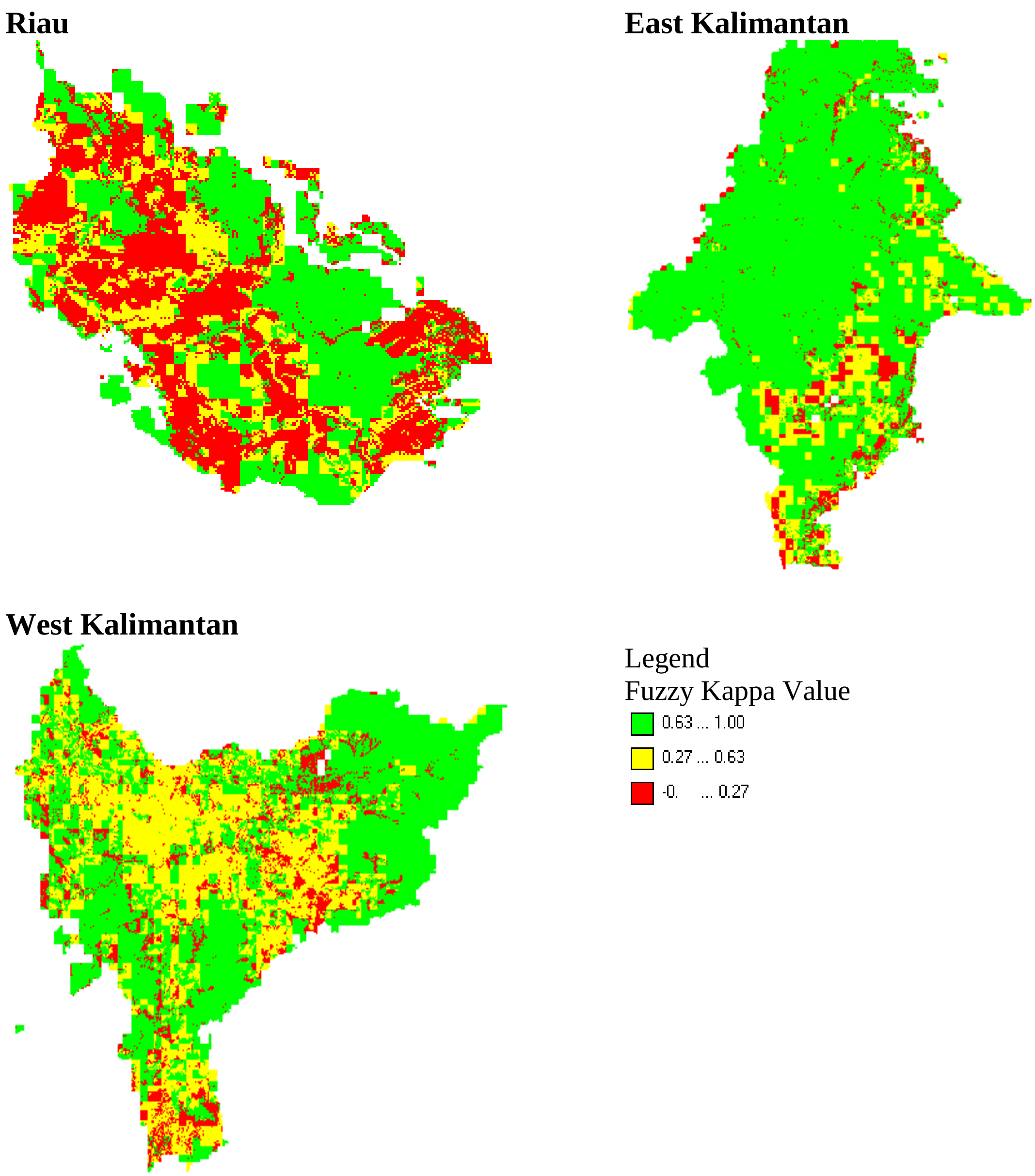


Figure IV.8 Fuzzy kappa comparison maps. These maps were produced using MCK software, by comparing OP map with the LU map from the MoF for the year of 2000 as an input. The green pixels show the high value of fuzzy kappa, which indicates that the two maps have the strong agreement in these pixels. The medium agreement is shown in yellow pixels, while low agreement is shown in red pixels.

Next, to investigate the possibility of predicting deforestation events using OP map, we calculated the normalized mean of pixel value for each deforested polygon. Table IV.2 below shows the value of the normalized mean pixel for deforested polygons that were converted into certain land use. Because we identified deforested areas, all changes were occurred in relatively low pixel value, for less than 4%, since this area were derived from forests which originally has the lowest pixel value in OP map. From Table IV.2, we found that forest change did not follow a specific pattern in the OP map. In all provinces, the highest value occurred in forest-agriculture and not in forest-oil palm as we expected.

Table IV.2 Normalized mean pixel value of deforested polygon (in percentage). The values were obtained by overlaying LU maps from the MoF and OP map.

	Riau			West Kalimantan			East Kalimantan		
	2000-2003	2003-2006	2006-2009	2000-2003	2003-2006	2006-2009	2000-2003	2003-2006	2006-2009
Forest - Oil Palm	2.63	2.69	2.33	2.10	0.67	1.43	0.28	0.44	0.42
Forest - Agriculture	3.81	3.69	2.28	1.79	1.27	2.09	0.38	0.52	0.44
Forest -Other	2.95	0.03	1.85	1.57	1.37	1.41	0.38	0.39	0.38

We also calculated statistical correlations between normalised mean value with the area of change to see the possibilities whether the pixel values has the correlation with the extent of change, but we found that only Riau that has a high negative correlation coefficient (-0.8). This indicated that in Riau, the wider deforested polygon, the lower pixel values in OP map. However, this correlation was not seen in other regions. West Kalimantan and East Kalimantan have correlation coefficient -0.3 and 0.2 respectively that were considered low.

IV.3 Modelling land use change

Two major products were generated in modelling land use change, which are transitional probability maps of each sub-model and prediction map for desirable year. Transitional probability maps depict the probability of all pixels to change, while prediction map was generated for the year of 2009. An intermediate process to test the power of explanatory variables was also being implemented. The description in sub chapter below explains the results of modelling part. An additional step to calculate the power of explanatory variables in sub-models was explained in appendix 2.

IV.3.1 Transitional probability maps from sub-model in observation period (2000-2006)

Sub-models were needed prior to change prediction. Sub models were built for the observation period, and the results were maps of transitional probability for each change. The maps show the probability of each pixel to change into a certain class. The probability is shown as value from 0 until 1, with 0 as have no probability to change, and higher value means the higher chance to change. Figure IV.9 illustrates the example result of transitional probability maps from forest into oil palm in Riau, East Kalimantan and West Kalimantan, while probability maps for other changes were presented in Appendix 3.

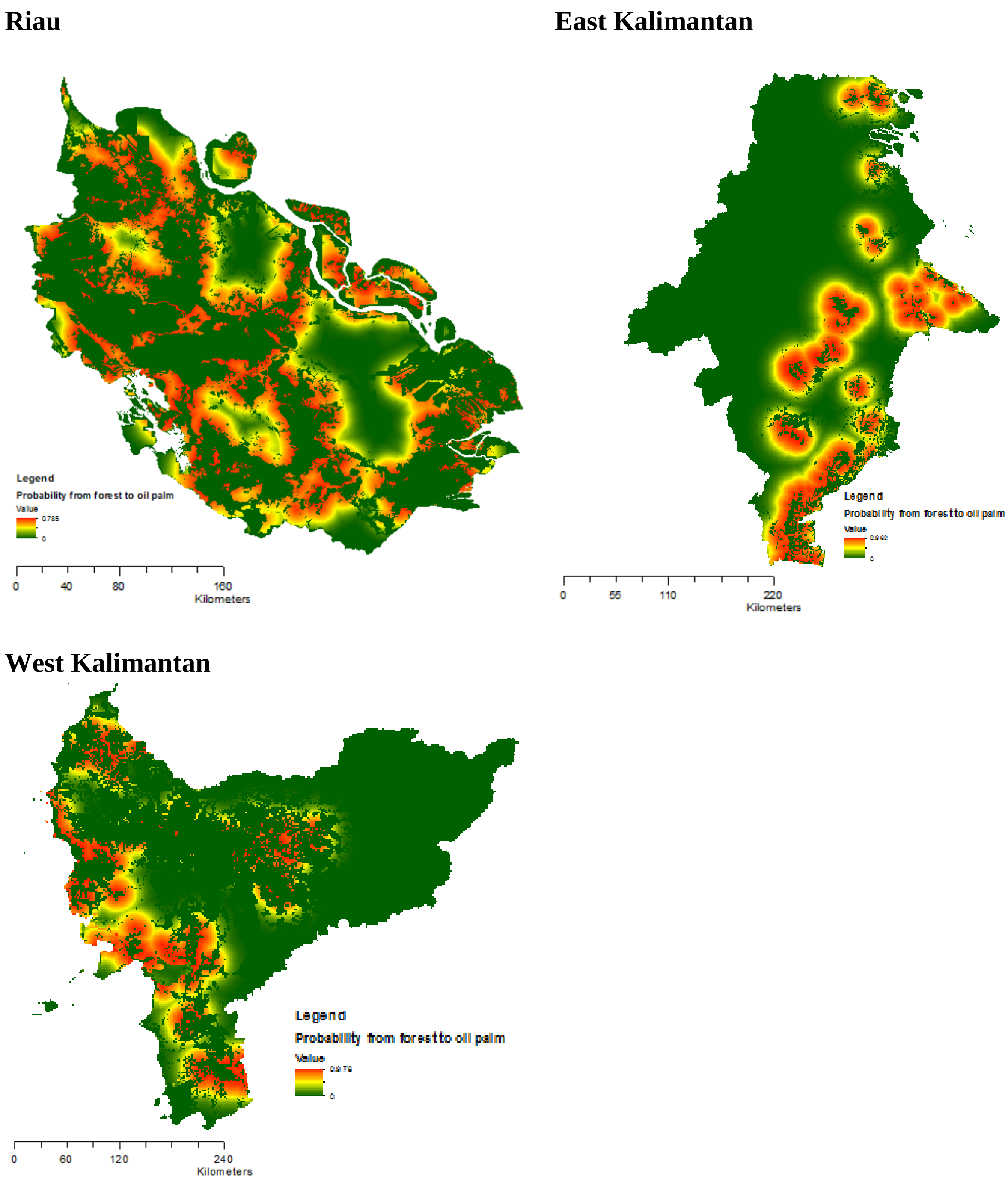


Figure IV.9 Transitional probability maps of forest-oil palm sub-model, derived from observation period in 2000-2006. The transitional probability was calculated in sub-model calculation in IDRISI LCM. The value in legend represent the probability of each pixel to be converted into certain class (in these maps were oil palm class)

IV.3.2 Prediction result for the year of 2009

The predictive model shows that from the three provinces, forest class was predicted to experience the biggest changes. Most of the forests area in Riau was predicted to change into other, while in East Kalimantan and West Kalimantan, it was predicted to change mostly into oil palm.

In Riau, the model predicted that approximately 10.83% of the area will change in 2009, or covering the area of 963, 590 ha. Forest was predicted to decrease 10.79% compared to the area in 2006, or 481,795 ha. Other class was predicted to increase 47.97% from its original area. Oil palm was predicted will increase to 3.97% from the previous, or 91,394 ha, while agriculture class was predicted will only increase of 0.65% or 8674 ha.

In East Kalimantan, the model predicted that approximately 1.4% of the total area in this province will change in 2009. Forest was predicted will only loss 0.8% of its area, or 139,910 ha in total. Oil palm was predicted will increase 12% from its initial, or 54,425 ha, while other class was predicted will increase 6.58%, or 63,754 ha.

In West Kalimantan, the model predicted that oil palm will increase 8% from the initial, or 42,838 ha. This change contributes 30.2% of total change during 2009. The model also predicted that forest will lose its area about 0.87% or 70,978 ha. Furthermore, agriculture class was also predicted to increase for 0.35% or 18, 742 ha, while other class will increase of 1.36% or 9, 249 ha.

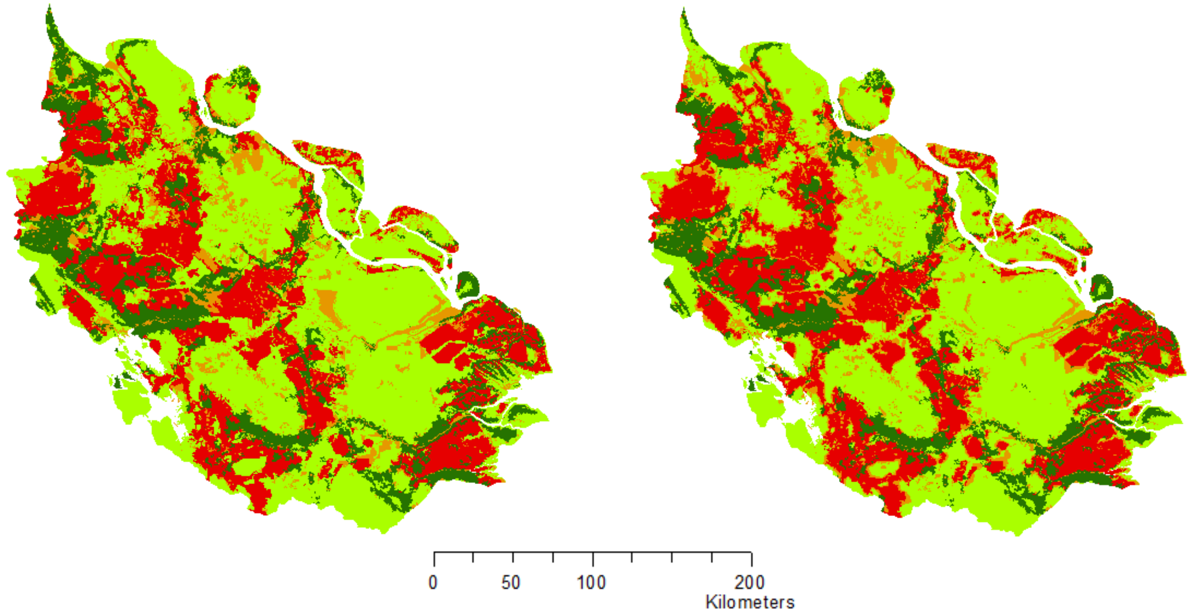
Table IV.3 below summarize the predictive area generated by the model and its comparison with the initial land use in the year of 2006 and 2009, presented in 1000 ha, while Figure IV.10 shows the comparison between 2009 initial land use and 2009 predicted land use.

Table IV.3 Comparison between initial land use in 2006 and 2009 with the predicted (pred) land use in 2009 (in 1000 ha)

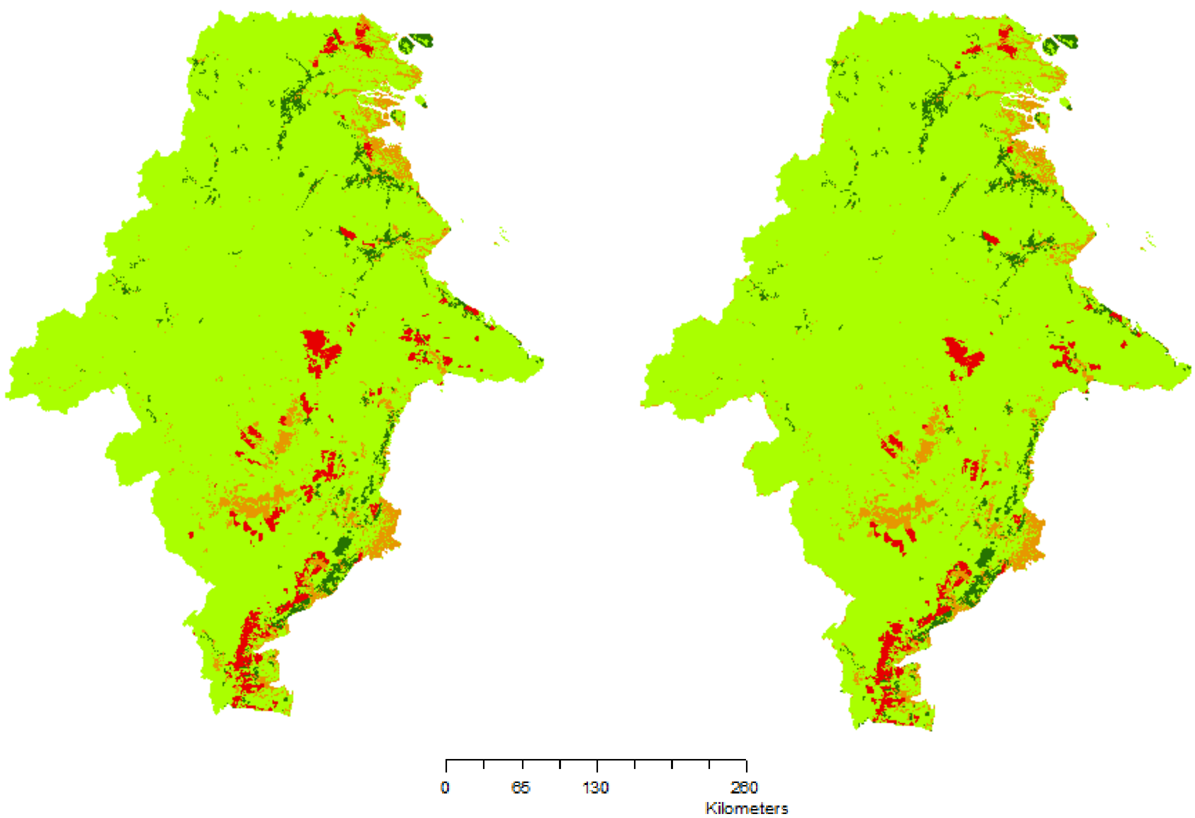
	Riau			**East Kalimantan**			**West Kalimantan**		
	2006 initial	2009 initial	2009 pred	2006 initial	2009 initial	2009 pred	2006 initial	2009 initial	2009 pred
Forest	4465.20	4322.60	3983.41	17792.48	17531.86	17652.57	8160.79	7789.88	8089.81
Agriculture	1328.19	1402.04	1336.86	572.54	589.92	594.27	5352.87	5345.79	5371.61
Oil Palm	2300.10	2372.60	2391.49	447.26	648.65	501.68	534.95	869.71	577.78
Other	795.72	798.24	1177.45	969.17	1011.01	1032.93	678.49	721.71	687.74

Initial land use map 2009 **Predicted land use map 2009**

Riau



East Kalimantan



West Kalimantan

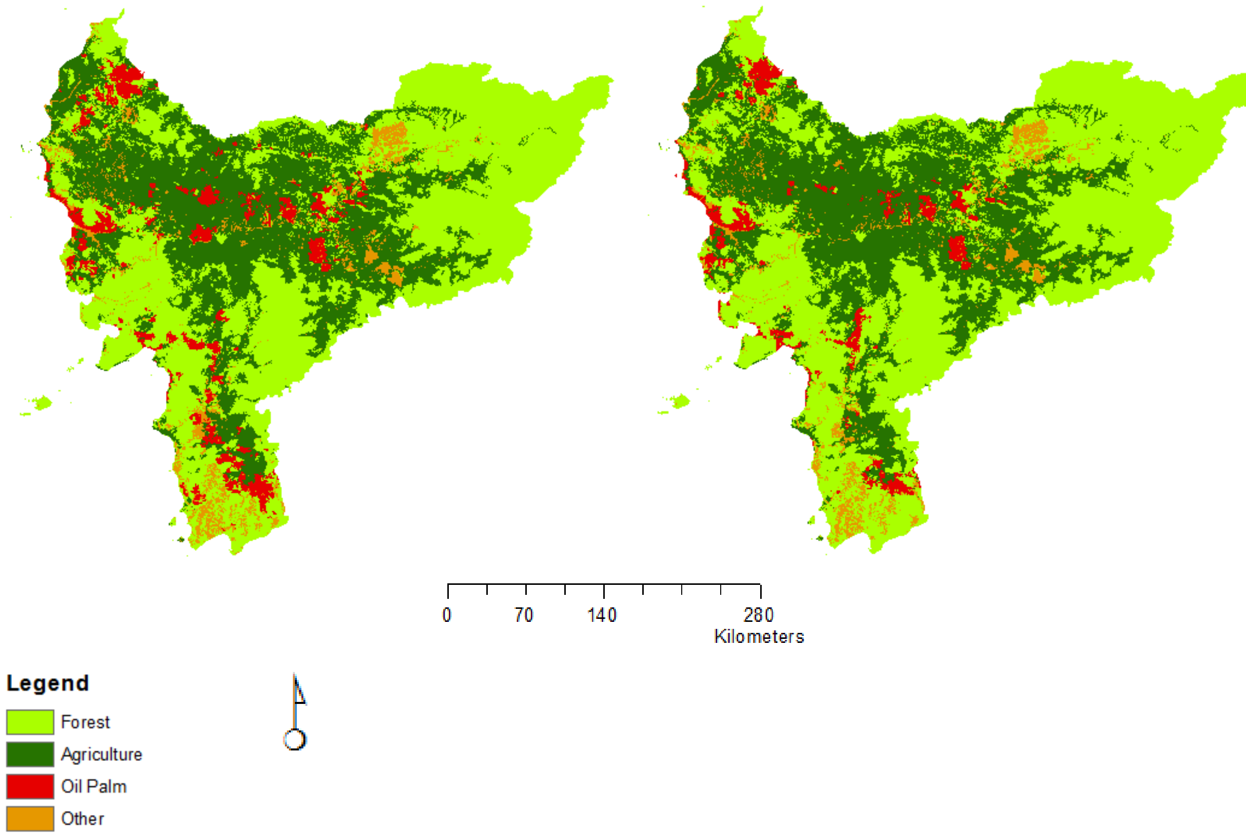


Figure IV.10 Comparison of the original land use from the MoF in 2009 with the prediction model result in the same year

IV.4 Calculating accuracy of the model (validation step)

Table IV.4 below shows the comparison of 2009 original change and 2009 predicted change in percentage, as a result of per category map comparison. The changes were calculated by subtracting 2009 initial and predicted maps with the initial land use map in 2006. Negative values indicate that the change occurred due to area loss, while positive value indicates that the change related to the increasing of the area.

Table IV.4 The comparison of percentage change from initial map and predicted map in 2009. The percentage was calculated by subtracting each class with the original land use in 2006 from the MoF. Negative values indicate area loss, while positive values indicate area gain.

	Riau		**East Kalimantan**		**West Kalimantan**	
Class	2009 initial change	2009 predicted change	2009 initial change	2009 predicted change	2009 initial change	2009 predicted change
Forest	-3.19	-10.79	-1.46	-0.79	-4.55	-0.87
Agriculture	5.56	0.65	3.04	3.80	-0.13	0.35
Oil Palm	3.15	3.97	45.03	12.17	62.58	8.01
Other	0.32	47.97	4.32	6.58	6.37	1.36

In Riau, forest and other were predicted overestimated than the actual change. Forest was predicted to loss its area of 7.60% more than the actual loss, where other class was predicted increased 47.63% more than the actual increase. This difference gives misclassification of 339,192 ha in forest class, and 379,213 ha in other class. The model predicted better in oil palm class, which only overestimated for 0.82% from its actual change, or 18, 896 ha. Meanwhile, agriculture class was underestimated by the model, with the differences of 4.91% or 18,896 ha.

Compared to Riau, in general, the model performed better in East Kalimantan. The model predicted forest loss only 0.68% less than the original change, which means that approximately 120,711 ha of the forests area were failed to be predicted as change. The model also predicted agriculture and other classes close to the real change, as the model only overestimated changes for less than 3% for these classes. Nonetheless, oil palm was predicted 32.86% less accurate compare to the real change, as the model predicted that this class will only increase 12.17%, but in reality, oil palm increased to more than 45% in 2009.

In West Kalimantan, the model performed better in predicting the changes in forest and agriculture classes. The model only gives a difference of 3.68% less than its actual in predicting forest loss. In agriculture class, the model predicted that the area will increase 0.35% but in actual it decreased to 0.13% from the initial area. Furthermore, oil palm was significantly underestimated by the model, where the model only predicted to increase for 8.01%, but in fact, this class increased to 62.58%. The distribution of error between 2009 initial map and 2009 predicted map is presented in Appendix 4

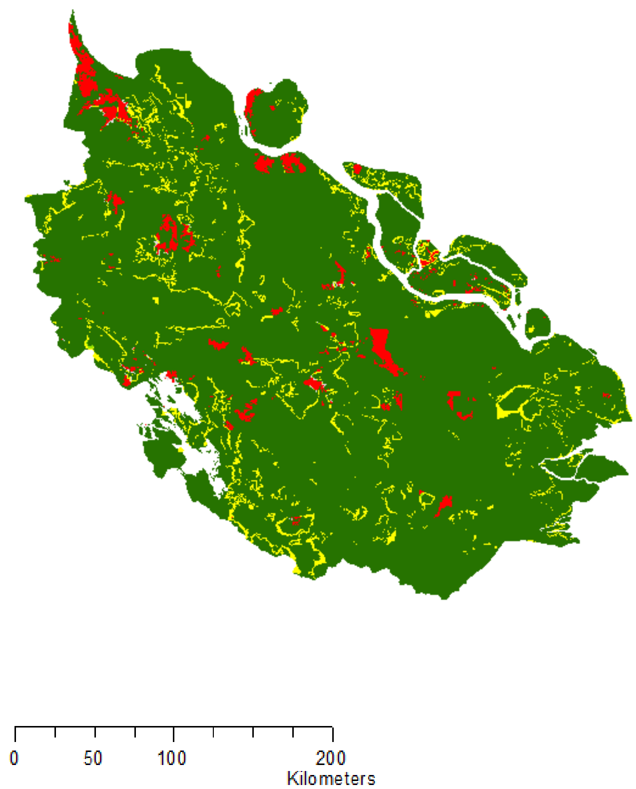
To give more insight of the sources of error in model prediction, we also implemented the three maps comparison as a complementary method for model validation. This method allocates the accuracy as two classes: accuracy due to persistence and accuracy due to changes. We compared 2006 initial map, 2009 initial map and 2009 predicted map. The three maps comparison method will allow the user to observe the origin of error, and depict the accuracy based on change and persistence of land use.

Table IV.5 show the total result of the three maps comparison accuracy assessment in percentage. The percentage was calculated by dividing the number of pixels in each class of accuracy with the total pixels. Of the three areas, the total agreements were mostly obtained from the persistence correct (the area that was not changes). East Kalimantan has the highest persistence correct with 97.87%, followed by West Kalimantan and East Kalimantan as 94.97% and 91.30%. Nonetheless, total agreements to predict the changes are low. The model only predicted changes correctly for 0.06% in Riau, 0.06% in East Kalimantan and 0.02% in West Kalimantan. Furthermore, total of disagreement in Riau is the highest compared to other study area, with 8.64% of disagreement error, followed by West Kalimantan with 5% and East Kalimantan for 2.08%. In Riau, total disagreements were mainly came from the error due to persistence in the initial map that was predicted as a change in the model, which contribute 5.41% from total disagreement error, whilst in East Kalimantan and West Kalimantan, these errors were mainly came from the error of change that was assigned as persistence by the predictive model. The maps presented in Figure IV.11 below shows the distribution of agreement and disagreement pixels from the three maps comparison.

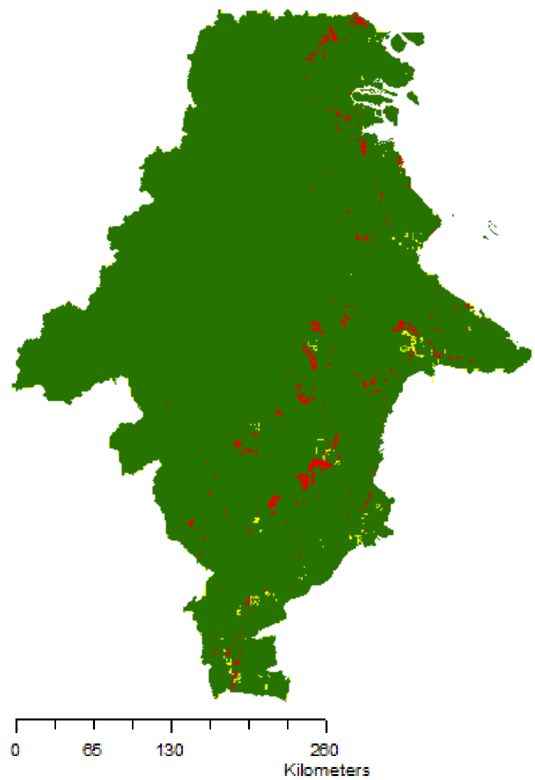
Table IV.5 Three map comparison result, as the result of overlaid 2006 initial LU map, 2009 initial LU map and 2009 predicted map. The values are in percentage.

	Riau	East Kalimantan	West Kalimantan
Persistence correct	91.30	97.87	94.97
Change correct	0.06	0.06	0.02
Total agreement	**91.36**	**97.92**	**95.00**
Error due to observed as change and predicted as persistence	3.10	1.42	4.46
Error due to observed as persistence and predicted as change	5.41	0.65	0.51
Error due to change but misclassified	0.12	0.01	0.04
Total disagreement	**8.64**	**2.08**	**5.00**

Riau



East Kalimantan



West Kalimantan

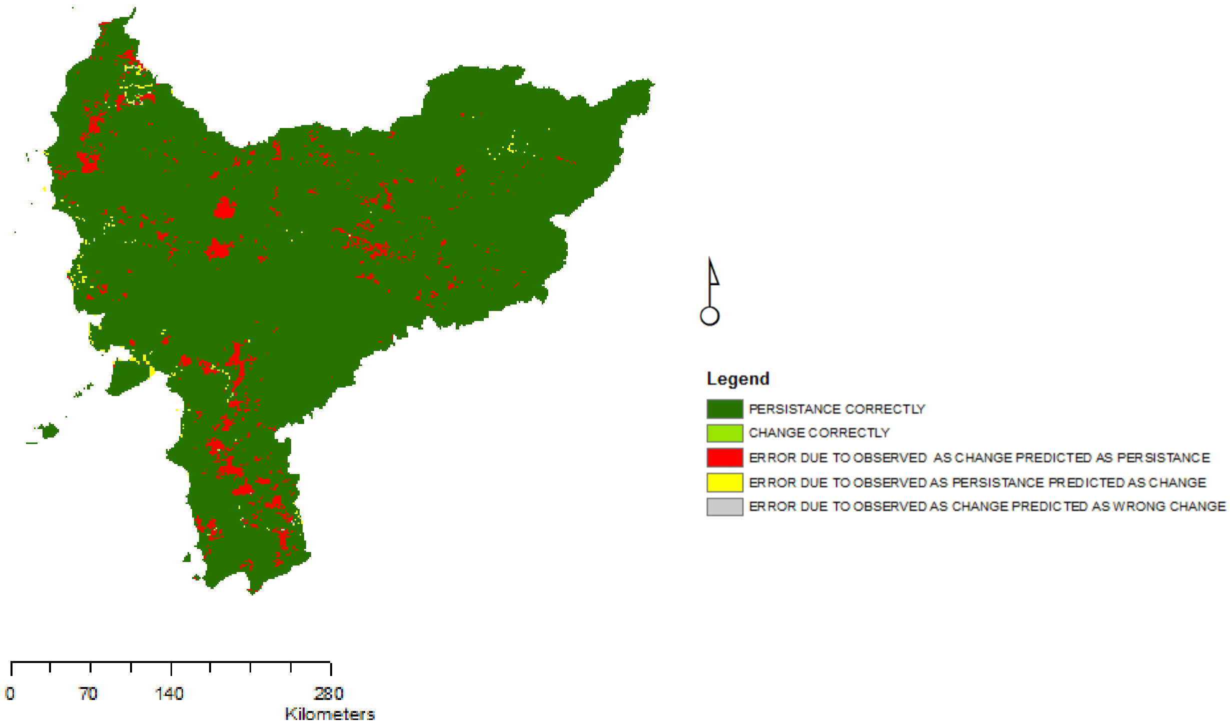


Figure IV.11 The class of accuracy assessment using the three maps comparison method. The maps were produced by overlaid 2006 initial LU map, 2009 initial LU map and 2009 predictive map.

V. DISCUSSION

V.1. Land use modelling and prediction, between reality and modelling

In this research, we use several assumptions regarding to modelling land use change. Firstly, we assumed that the patterns of predictor variables were constant over the time. Secondly, we assumed that there were no policy changes at global, national or regional level that will affect the pattern of land use change during the observation and prediction periods, and the economic drivers were assumed to be stable. Thirdly, we assumed that the variables related to the government policies were implemented ideally. Nonetheless, the real process of land use change sometimes far from the assumptions, as confirmed by earlier researches (Waddell 2002). In reality, anthropogenic and topographic variables sometimes change with unpredictable patterns by some reasons such as natural disasters, economic growth, and the policy change or the policy that has not been implemented effectively.

In this research, the prediction result shows that the model gives low accuracy in the most dominant changes during observation period, which are other in Riau and oil palm in East Kalimantan and West Kalimantan. In Riau, other class was overestimated 47.63% more than the actual change, while in East Kalimantan and West Kalimantan, oil palm were predicted 32.86% and 54.57% less than the actual increased, as indicated in Table IV.3 and Table IV.4. However, the class that has only small changes within the periods such as agriculture was predicted better by the model. Looking back to the actual trend from the reference map as presented in Figure IV.1 and Figure IV.2, we found that the actual trend changed significantly during the observation period and prediction period. In Riau, in general, the deforestation rate decreased during the prediction period and the intensity of change of other class also decreased significantly compare to the previous periods. This condition also occurred in East Kalimantan and West Kalimantan, when the actual trend shows that oil palm experienced a significant increasing during prediction period compare to the observation periods.

Several possible explanations could explain why these striking differences occurred. In 2009, the world crude oil palm price increased up to 780$/ ton, has resulted in a large profit in oil palm business, and encourage the producers to expand their operation (Wicke et al. 2011). During this period, Indonesia intensively increased oil palm production by extending the plantation area, and after Sumatera that has been intensively expanded, future development was directed to Kalimantan and Papua (Susanti and Burgers 2012, Budidarsono et al. 2013). Furthermore, the establishment of the Decree of Agricultural Ministry No. 26/ 2007 allowing the large, capital-intensive companies to expand their production to invest in labour-intensive oil palm projects with smallholder farmers through various partnership schemes, such as contract farming or *Perkebunan Inti Rakyat (PIR)*, also contributed in accelerating the expansion of oil palm plantation in Indonesia. The increasing of global oil palm demand, and the scheme of policy that allowed smallholders to access technology and market through partnership scheme, encourages the emergence of smallholder plantations in Indonesia (Sumargo et al. 2009, Susanti and Burgers 2012).The census shows that during 2006-2009, the establishment of large oil palm plantations in Indonesia increased for almost 6% more than the previous periods, or increased 12,946,000 ha in total, while smallholder plantations increased for 70,450,000 ha (BPS 2014).These global market dynamics, followed by the change of national policies, are the examples of driver factors that failed to be modelled and predicted by classical model of land use change.

Reviewing to the third assumption, Indonesian government established several area as national parks that were designated as protected areas, with the assumption that no activities that potentially change land use are allowed. Therefore, in reality, Indonesia encounters the problem of rapid deforestations and forest degradations, including in conservation area (Jepson et al. 2001, Curran et al. 2004, Joppa et al. 2008). As an example, Curran et.al (2004) wrote that in Kalimantan, protected lowland forests declined by more than 56% of its total area, or more than 29,000 km^2 during 1985 to 2001. Tesso Nilo National Park in Riau was also reported experienced intensive deforestation and forest degradation (Braun 2012), and oil palm expansion also reported occurred in this area (McLaughlin 2011). This condition leads to the incapability of the land use modelling to predict deforestation or oil palm expansion in the protected areas, as the model assign that there will be no changes occurred in these areas.

V.2. Artificial Neural Network for modelling land use change

In this research, ANN method was proven to give better result in predicting persistence rather than predicting the change (see Section IV.4). The accuracy of the model in predicting the persistence is 91%, 97% and 95% for Riau, East Kalimantan and West Kalimantan respectively, while in predicting the change, the model can only predict the change correctly of 0.06% for Riau and East Kalimantan, and 0.02% for West Kalimantan. However, this condition was confirmed by Pontius et.al (2004) and Pontius Jr et.al (2008), as they mentioned that in many cases, the accuracy of the model will mainly came from the ability of the model to predict persistence rather than to predict the change. They also explained that this condition occurred since the changes were only occupied a small percentage of the total area, and it was acceptable that the accuracy of the model to predict the changes were lower than to predict the persistence.

Furthermore, some earlier studies have implemented ANN method to model land use change, and some of them have proved that this method performs better rather than other statistical land use modelling methods (Hill et al. 1994, Ahmed and Ahmed 2012, Ahmed et al. 2013). However, this method has been criticized as giving a 'black box' of ignoring the mechanisms that occur within the network, since the method dismissing the contribution of each input variable to the model output and the actual relationship between the input variables and the output layer is unknown (Gevrey et al. 2003, Mas et al. 2004, Olden et al. 2004). Even though the LCM package in IDRISI software provides a classical stepwise method to assess the degree of significance of each input variable (see Section III.3.3 and Appendix 2), the results did not explicitly show the degree of correlation of each variable with the land use change, and the kind of contribution of each input variable to the model output, in which this information can be one of the most important part to understand the process of land use change (Gevrey et al. 2003). Therefore, for example, we cannot identify whether slopes has a negative or positive correlation with deforestation and the strength of the relationship, since the result from LCM only stated that this variable was significant or insignificant in deforestation process.

V.3. Elaborating global open source data to predict land use change

Several issues arise when we elaborated global open source data for land use modelling. The problems arise in relation with spatial resolution, temporal resolution and the accuracy level of the original datasets. The sub-chapter below describe about what problems arise during our research.

V.3.1. Spatial resolution and its implications on land use modelling

In this research, we elaborated several datasets from different source with different format and spatial and temporal resolution (see Table III.1 for the source and description of the datasets). Our datasets have resolutions that vary widely. Choosing the best resolution for land use modelling is somewhat challenging, as the input resolution will affect the result and determine the speed of the computation process. The choice to choose the best resolution is fully depending on the users, with the consideration of the scale and extent of the model and the purpose of land use modelling. So far, there are no specific methods of selecting the most appropriate resolution.

Some studies have proven on how the resolution of the input data will affect the sensitivity of the land use model. De Koning et al. (1998) stated that the patterns of land use/ land cover map can disappear or emerge, going from one scale to the other. He also mentioned that spatial autocorrelation will increase as the spatial resolution increase, and vice versa (De Koning et al. 1998). In terms of modelling land use, this statement has been confirmed by Rajan (2010), in which he proved that finer resolution of the input variable will give better predictive result of land use change, referring to his result in studying land use change in Imbabura Province, Ecuador. However, Pontius & Huffaker (2004) found that this not always be the case. In their work to develop validation techniques for modelling land use change, they mentioned that as the resolution of predictive model increased, it's percent correct generally will also increase, which mean that the accuracy of predictive model increase. Therefore, they introduced the concept of Null model and Null resolution to determine the most fair resolution in predictive model, which originally developed as validation techniques of land use modelling (Pontius et al. 2004, Pontius and Malanson 2005). Even though the concept of Null resolution was aimed to validate predictive land use model, this concept is potentially implemented as an approach to determine the most suitable resolution for input in land use modelling. However, our research was only limited to model land use change in a single resolution (30 m x 30 m), and did not test the effect of the change of resolution on the output as we consider that this chosen resolution can represent the pattern of land use our study area.

V.3.2. The match between the data in describing the real state

Secondly, the match between the data might raise problems. These issues might be sourced from the difference of spatial reference or temporal resolution. As an example, we found some discrepancies between LU maps and PA map. We consider that protected areas are treated as constrain variable, in which the possibility of this area to change into other land use is zero. In fact, we found that there were changes occurred in protected area, that can occurred because probably there were changes in the area (see Section V.1), because the datasets have different spatial references that cause overlapping, or at the time the change occurred, the area has not been designated as protected area (we use 2014 version of protected area map). As an illustration, Tesso Nilo National Park in Riau was established in July 2004 with the area of 38,576 ha, and in 2009 was expanded into 83,064 ha by the government. The change of the extent might explain why there were changes occurred during observation period, and illustrated that the difference of when the data was produced can be the source of error in predicting land use change.

V.3.3. The quality and accuracy of the original data

The level of accuracy in original data might be the issues in elaborating open source data. To what extent of global open source data is reliable and accurate in representing the actual world will also affect the decision on to what extent and what scale of the datasets can be used for land use modelling. Census-based dataset, for example, has been criticized for inaccuracy, since no independent remote sensing survey data can be carried out to validate the data at the time where they compiled (Kongsager and Reenberg 2012). As an example, we used crop map datasets, which in this research we used OP map as part of crop maps as our reference data. These datasets have been used as reference of some analysis of land cover/ land use in global scale. For example, Licker et al. (2010) used the same crop maps to evaluate the yield gap for 18 different crops in global scale in relationship with the climate and agricultural management, and able to concluded the effect of different climate and management practice in contributing the differences of crop yield in global scale. Another example, Avnery et al. (2011) was also successfully counted the global reduction of agricultural yields of three different crops because of the elevated concentrations of surface ozone (O_3) in the year of 2000 using these datasets (Avnery et al. 2011). However, as we tried to use OP map that was the part from crop maps dataset for analysis at sub-national level, the data did not show any significance in the model performance, and only perform significant in forest-other change in East Kalimantan (see Section III.3.5 and Appendix 2). Furthermore, our investigation to look at the possibility of using this dataset to predict the next oil palm area after deforestation also showed no significant correlation (see Section IV.2). Here, Licker et al. (2010) confirmed that crop map datasets should be used to address issues related to their intended purpose, which is to compare regional crop yield patterns and agriculture area patterns, rather than to investigate it as individual grid cell. Moreover, Kongsager & Reenberg (2012) also mentioned that agricultural census-based map might have error that sourced from the under estimating of household-based production for self-consumption. Furthermore, since census data in crop map datasets were extrapolated to a limited biophysical and topographic parameter, it is possible that the datasets will give error due to the distribution of crop yields onto occasional grid cell with unsuitable conditions (Licker et al. 2010), in which our finding proved that the distribution pattern of oil palm dataset did not have any correlation with the existing of oil palm area based on the MoF dataset. Therefore, we found that oil palm dataset cannot be used to predict the next expansion of the oil palm area.

V.4. The role of agricultural expansion in land use change

This research proved that, among debates of whether oil palm plantation has been changing the world's forest, oil palm has proven to be the major driver of deforestation during 2000 – 2009, especially in East Kalimantan and West Kalimantan. In these provinces, deforestation was mainly followed by oil palm. For approximately 47% of the total deforested area in East Kalimantan was converted into oil palm, while in

East Kalimantan was 42%. This means that approximately 315,336 ha and 267,040 ha of forests area were converted into oil palm in East Kalimantan and West Kalimantan respectively. Moreover, our finding in Riau also shown that agricultural expansion can be preceded by intermediate change, as explained in Figure IV.4. We found that once the forest has been cleared, in some cases, it does not necessarily followed by the opening of new oil palm plantation. We found that forest area that was converted into open land in 2003, 14% of them were converted into palm oil in 2006, while 51% of them were still remaining open land. Looking at the same polygons of remaining open land in 2006, 90% of them were still remaining as open land in 2009, while the rest of them were converted into other land use. It is unclear on what kind of land use right that follow the forest clearing in these polygons, but if it was followed by cultivation right, it is most likely that these lands should be converted into oil palm plantations, as this agricultural type is the most common type of commodities. Based on Indonesian Agrarian Law, as the land was not cultivated in accordance with the purpose, this land can be categorized as abandoned land and the right can be revoked by the Government. The occurrence of abandoned land in Indonesia has been claimed to give adverse effect economically and environmentally (Sigit 2012). In Riau itself, there are approximately 55,451 ha of cultivation right that have been abandoned in 2010 (NLA 2010). Therefore, we conclude that not only as the driver of deforestation, agriculture expansion also potentially drives the occurrence of abandoned land, in which this issue is still rarely investigated.

Moreover, look through at Riau case, even though many researches claimed that oil palm plantations were not always converted from forestland and in some area only have a limited contribution to the total forests loss in Indonesia (Koh and Wilcove 2008, Butler 2011, Susanti and Burgers 2012), which will also be found if we directly interpret the result of land use change detection, our study found that further analysis on the history of land use change is necessary to get more insight on direct or indirect process of agriculture expansion in changing the world's forest. As an example, Budidarsono et al. (2013) also mentioned that the establishment of large-scale oil palm plantations strongly influences the rate of land development in a region by speed up the development of infrastructure and stimulates the growth of the local economy. This condition is indirectly affects the emergence of non-agricultural activities in the area as the consequences of economic growth, such as trade, home industry and services, in which agriculture expansions are responsible for (Budidarsono et al. 2013). Our model, unfortunately, have the limitation in explaining this continuous effects of agriculture expansion. Therefore, we suggest that, future study of the land use change history might useful in understanding the process of land use change related to agriculture expansion.

V.5. Future study of modelling land use change

Summarized the explanations above, some improvements might be implemented for future study. First, as we only model land use change in single resolution (30 x 30 m), and even though the original resolution of the datasets determines most of the result, future study on to what extent of input resolution will affect the result of the land use model, and how to choose the best fit resolution of the input model might give benefit for future improvement of modelling land use change. Here we suggest that the concept of null resolution from Pontius et.al (2004) and Pontius and Malanson (2005) can be improved to be used as the method to choose the optimal resolution in land use modelling.

Furthermore, following the result from Licker et al. (2010) on the investigation to analyse the yield gap of crop map, even though the research was carried out on the global scale, might also be used to explain the yield gaps in sub-national level and furthermore to explain the underlying drivers of oil palm expansion. In our study area, we found that the three provinces have relatively large difference of the maximum pixel values. This also indicated that they have different oil palm yields, and ignoring the factor of climate, the gaps can be sourced from the differences of soil quality and human management, including irrigation, fertilization and other planting practices (Licker et al. 2010), and also the different agricultural practises. Looking at these reasons, OP map, even though mostly performed as insignificant variable in our method of land use modelling in our study area, provides an opportunity to compare crop yields at national or sub-national level, and investigating the drivers of oil palm comprehensively.

In relevance with modelling land use change using statistical approach, we suggest that, given the limitation of land use change modelling with several assumptions of some behaviour of the variables, rather than literally interpret the prediction result as an absolute future change, it should be seen as a tool to give the insight of the area that are vulnerable to change. Apart from the issue of accuracy, this interpretation can be used as reference for the next policy and monitoring program of deforestation and land use change.

VI. CONCLUSION

Agricultural activities have had a long history in changing the world's surface. Even though not always be the case, agricultural expansions are often became one of the main causes of deforestation, and has been accused responsible for environmental degradation, in global and local scale. Nonetheless, population growth and the increasing of world food demand put agriculture expansion become inevitable. Therefore, understanding the dynamic of agriculture expansions is important.

Our study shows that oil palm activities were proven responsible for the most change of land use in our study area. We also found that sometimes, deforestations involving agriculture expansion was followed by intermediate change before it were established as plantation areas. Furthermore, our result shows that distance from existing oil palm plantation was the most significant variable in the process of deforestation and agriculture expansion.

Furthermore, we found that land use modelling can be used as an effective way to capture the dynamics of agricultural activities. Artificial Neural Network, integrated with Markov Chains method can be used to model the change of land use following deforestation, particularly for oil palm, and can answer the questions which might arise on what change, where change, the extent of change and what will be the next condition. However, our results show the limitation of a statistical model in predicting land use change. The model cannot capture the dynamic of deforestation drivers that which were not included as model parameters. Here, we found that the rapid expansion of oil palm that mainly driven by global market cannot be captured by the model. Therefore, land use change is a complex system that cannot be captured only by a single method of land use modelling.

Our study also proved that integrating open source datasets for modelling land use change purpose is possible to be implemented, even though integrating global scale data for local analysis may arise several problems. The issue of aggregation level caused by the resolution of the datasets raises the problem of loss of information, which at some point made the dataset perform insignificantly when it was applied for the analysis at a more detailed level. Furthermore, we conclude that the statistical modelling method should be considered as a tool, which cannot be interpreted as the real representation of the state, but rather to give the introduction on what process that occurred in land use change and cannot be solely considered as an absolute explanation. To get the more comprehensive explanation on the real process of land use change, an interdisciplinary analysis should be implemented.

REFERENCES

(1998). Forest Area Criteria for Oil Palm Plantation Cultivation Decree of Ministry of Forestry and Agriculture Nr: 376/Kpts-II/1998. Indonesia http://www.dephut.go.id/index.php/ news /otresults/ 152.

Aaviksoo, K. (1995). "Simulating vegetation dynamics and land use in a mire landscape using a Markov model." Landscape and Urban Planning**31**(1): 129-142.

Ahmed, B. and R. Ahmed (2012). "Modeling Urban Land Cover Growth Dynamics Using Multi‑Temporal Satellite Images: A Case Study of Dhaka, Bangladesh." ISPRS International Journal of Geo-Information**1**(1): 3-31.

Ahmed, B., R. Ahmed and X. Zhu (2013). "Evaluation of model validation techniques in land cover dynamics." ISPRS International Journal of Geo-Information**2**(3): 577-597.

Avnery, S., D. L. Mauzerall, J. Liu and L. W. Horowitz (2011). "Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage." Atmospheric Environment**45**(13): 2284-2296.

Barbier, E. B., N. Bockstael, J. C. Burgess and I. Strand (1995). "The linkages between the timber trade and tropical deforestation—Indonesia." The World Economy**18**(3): 411-442.

Bell, E. J. (1974). "Markov analysis of land use change—an application of stochastic processes to remotely sensed data." Socio-Economic Planning Sciences**8**(6): 311-316.

Blennow, K. and P. Persson (1998). "Modelling local-scale frost variations using mobile temperature measurements with a GIS." Agricultural and forest Meteorology**89**(1): 59-71.

Bohling, G. C. and M. K. Dubois (2003). An integrated application of Neural Network and Markov Chain techniques to prediction of lithofacies from well logs Kansas, USA, Kansas Geological Survey

Boyd, D. S. (1996). "A review of “Modelling Deforestation Processes:” A review. Trees Series B Report n◦ l. By EF Lambin.(Luxembourg: European Commission, 1994)[Pp. 113 “rsqb;. Free, on request." International Journal of Remote Sensing**17**(5): 1061-1062.

BPS (2014). Kalimantan Barat Dalam Angka Indonesian Statistical Bureau

BPS (2014). Riau Dalam Angka Indonesia, Indonesian Statistical Bureau

BPS, I. S. B. (2014). "Statistics Indonesia." Retrieved 11 April 2015, 2015, from http://www.bps.go.id/ linkTabelStatis/ view/id/ 1665.

Braun, S. (2012). "National Park Management between Rhetoric and Results: The Failure of Indonesia's Mainstream Conservation Model in Tesso Nilo and the Advantages of Private Property in Nature Conservation Efforts." Civil society conflicts-Environment & Social Conflicts**2**(1).

Budidarsono, S., A. Susanti and A. Zoomers (2013). "Oil palm plantations in Indonesia: The implications for migration, settlement/resettlement and local economic development."

Butler, R. A. (2011). "First large-scale map of oil palm plantations reveals big environmental toll." Retrieved 09 April 2015, 2015, from http://news.mongabay.com/2011/0307-koh_palm_oil.html.

BV, R. (2010). Map Comparison Kit 3 User Manual The Netherlands, RIKS BV.

Camacho Olmedo, M., M. Paegelow and J.-F. Mas (2013). "Interest in intermediate soft-classified maps in land change model validation: suitability versus transition potential." International Journal of Geographical Information Science**27**(12): 2343-2361.

Carter, S. and J. Stuiver (2014). Slope derived from SRTM data, v1, Wageningen University and CIFOR.

Civco, D. L. (1993). "Artificial neural networks for land-cover classification and mapping." International Journal of Geographical Information Science**7**(2): 173-186.

Curran, L. M., S. N. Trigg, A. K. McDonald, D. Astiani, Y. Hardiono, P. Siregar, I. Caniago and E. Kasischke (2004). "Lowland forest loss in protected areas of Indonesian Borneo." Science**303**(5660): 1000-1003.

Davis, G. and R. Ackermann (1990). "Indonesia: Sustainable development of forests, land, and water."

De Koning, G., A. Veldkamp and L. Fresco (1998). "Land use in Ecuador: a statistical analysis at different aggregation levels." Agriculture, ecosystems & environment**70**(2): 231-247.

Dybowski, R., V. Gant, P. Weller and R. Chang (1996). "Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm." The Lancet**347**(9009): 1146-1150.

Eastman, J. and J. Toledano (2000). Markov Chain and Cellular Automata Approaches to Land Cover Change Modeling. Workshop presentation (invited), 4th International Conference on Integrating GIS and Environmental Modeling, Banff, Alberta.

Eastman, J. R. (2006). IDRISI Andes guide to GIS and image processing Worcester, USA, Clark Labs, Clark University

Etter, A., C. McAlpine, K. Wilson, S. Phinn and H. Possingham (2006). "Regional patterns of agricultural land use and deforestation in Colombia." Agriculture, ecosystems & environment**114**(2): 369-386.

Fearnside, P. M. (1997). "Transmigration in Indonesia: lessons from its environmental and social impacts." Environmental Management**21**(4): 553-570.

Fearnside, P. M. (2000). "Global warming and tropical land-use change: greenhouse gas emissions from biomass burning, decomposition and soils in forest conversion, shifting cultivation and secondary vegetation." Climatic change**46**(1-2): 115-158.

Fitzherbert, E. B., M. J. Struebig, A. Morel, F. Danielsen, C. A. Brühl, P. F. Donald and B. Phalan (2008). "How will oil palm expansion affect biodiversity?" Trends in ecology & evolution**23**(10): 538-545.

Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily and H. K. Gibbs (2005). "Global consequences of land use." science**309**(5734): 570-574.

Fraser, A. I. (1998). Social, economic and political aspects of forest clearance and land-use planning in Indonesia. Human activities and the tropical rainforest, Springer**:** 133-150.

Geist, H. J. and E. F. Lambin (2001). "What drives tropical deforestation." LUCC Report series**4**: 116.

Gelling, P. (2007). "Forest loss in Sumatra becomes a global issue." The New York Times.

Gevrey, M., I. Dimopoulos and S. Lek (2003). "Review and comparison of methods to study the contribution of variables in artificial neural network models." Ecological Modelling**160**(3): 249-264.

Gilks, W. R. (2005). Markov chain monte carlo, Wiley Online Library.

Hagen-Zanker, A., B. Straatman and I. Uljee (2005). "Further developments of a fuzzy set map comparison approach." International Journal of Geographical Information Science**19**(7): 769-785.

Hagen, A. (2003). "Fuzzy set approach to assessing similarity of categorical maps." International Journal of Geographical Information Science**17**(3): 235-249.

Hassan, R. M., R. Scholes and N. Ash (2005). Ecosystems and human well-being: current state and trends: findings of the Condition and Trends Working Group, Island Press.

Hepner, G. F., T. Logan, N. Ritter and N. Bryant (1990). "Artificial neural network classification using a minimal training set-comparison to conventional supervised classification."

Hill, T., L. Marquez, M. O'Connor and W. Remus (1994). "Artificial neural network models for forecasting and decision making." International Journal of Forecasting **10**(1): 5-15.

Hosonuma, N., M. Herold, V. De Sy, R. S. De Fries, M. Brockhaus, L. Verchot, A. Angelsen and E. Romijn (2012). "An assessment of deforestation and forest degradation drivers in developing countries." Environmental Research Letters **7**(4): 044009.

Hsu, K. l., H. V. Gupta and S. Sorooshian (1995). "Artificial Neural Network Modeling of the Rainfall-Runoff Process." Water resources research **31**(10): 2517-2530.

Hunt, C. (2010). "The costs of reducing deforestation in Indonesia." Bulletin of Indonesian Economic Studies **46**(2): 187-192.

Indonesia, F. W. and G. F. Watch (2002). "The state of the forest: Indonesia." Forest Watch Indonesia, Bogor, Indonesia, and Global Forest Watch, Washington DC.

Iremonger, S., C. Ravilious and T. Quinton (1997). "A global overview of forest conservation."

IUCN and UNEP-MCMC. (2014). "The World Database on Protected Areas (WDPA) ", 2014, from http://www.protectedplanet.net/.

Jarvis, A., H. I. Reuter, A. Nelson and E. Guevara (2008). Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT).

Jat, M. K., P. K. Garg and D. Khare (2008). "Monitoring and modelling of urban sprawl using remote sensing and GIS techniques." International journal of Applied earth Observation and Geoinformation **10**(1): 26-43.

Jepson, P., J. K. Jarvie, K. MacKinnon and K. A. Monk (2001). "The end for Indonesia's lowland forests?" Science **292**(5518): 859-861.

Johnes, P. J. (1996). "Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach." Journal of hydrology **183**(3): 323-349.

Johnston, B. F. and J. W. Mellor (1961). "The role of agriculture in economic development." The American Economic Review: 566-593.

Jokar Arsanjani, J., M. Helbich, W. Kainz and A. Darvishi Boloorani (2013). "Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion." International Journal of Applied Earth Observation and Geoinformation **21**(0): 265-275.

Joppa, L. N., S. R. Loarie and S. L. Pimm (2008). "On the protection of “protected areas”." Proceedings of the National Academy of Sciences **105**(18): 6673-6678.

Kanellopoulos, I., A. Varfis, G. Wilkinson and J. Megier (1992). "Land-cover discrimination in SPOT HRV imagery using an artificial neural network—a 20-class experiment." International Journal of Remote Sensing **13**(5): 917-924.

Kehutanan, K. (2014). Statistik Kementerian Kehutanan 2013, Kementerian Kehutanan Republik Indonesia

Khan, J., J. S. Wei, M. Ringner, L. H. Saal, M. Ladanyi, F. Westermann, F. Berthold, M. Schwab, C. R. Antonescu and C. Peterson (2001). "Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks." Nature medicine **7**(6): 673-679.

Kissinger, G., M. Herold and V. De Sy (2012). Drivers of deforestation and forest degradation: a synthesis report for REDD+ policymakers, Lexeme Consulting Vancouver.

Koh, L. P. and D. S. Wilcove (2008). "Is oil palm agriculture really destroying tropical biodiversity?" Conservation letters**1**(2): 60-64.

Kongsager, R. and A. Reenberg (2012). Contemporary land-use trasitions: the global oil palm expansion, Department of Geography and Geology, University of Copenhagen.

Koomen, E. and J. Borsboom-van Beurden (2011). Land-use modelling in planning practice, Springer Science & Business Media.

Lambin, E., B. Turner, H. Geist, S. Agbola, A. Angelson, J. Bruce, O. Coomes, R. Dirzo, G. Fischer and C. Folke (2001). "Our emerging understanding of the causes of land use and cover change." Global Environmental Change**11**(4): 261-269.

Lambin, E. F. (1997). "Modelling and monitoring land-cover change processes in tropical regions." Progress in physical geography**21**(3): 375-393.

Lambin, E. F., M. D. A. Rounsevell and H. J. Geist (2000). "Are agricultural land-use models able to predict changes in land-use intensity?" Agriculture, Ecosystems & Environment**82**(1–3): 321-331.

Land, S.-D. M. o. A. (2010). Abandoned Land Database S.-D. M. o. A. Land. Indonesia Ministry of Agraria and Land Use Planning/ National Land Agency Republic of Indonesia.

Lee, S., J.-H. Ryu, J.-S. Won and H.-J. Park (2004). "Determination and application of the weights for landslide susceptibility mapping using an artificial neural network." Engineering Geology**71**(3–4): 289-302.

Lee, T. M., N. S. Sodhi and D. M. Prawiradilaga (2007). "The importance of protected areas for the forest and endemic avifauna of Sulawesi (Indonesia)." Ecological Applications**17**(6): 1727-1741.

Li, H. and J. Reynolds (1997). Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: A combined Markov and cellular automaton approach, New York: Lewis Publishers**:** 211-230.

Li, X. and A. G.-O. Yeh (2002). "Neural-network-based cellular automata for simulating multiple land use changes using GIS." International Journal of Geographical Information Science**16**(4): 323-343.

Licker, R., M. Johnston, J. A. Foley, C. Barford, C. J. Kucharik, C. Monfreda and N. Ramankutty (2010). "Mind the gap: how do climate and agricultural management explain the ‘yield gap’of croplands around the world?" Global ecology and biogeography**19**(6): 769-782.

Ligtenberg, A., A. K. Bregt and R. Van Lammeren (2001). "Multi-actor-based land use modelling: spatial planning using agents." Landscape and urban planning**56**(1): 21-33.

Locke, A. and G. Henley (2013). "Scoping report on biofuels projects in five developing countries." Overseas Development Institute, London.

Manurung, E. T. (2000). Mengapa Konversi Hutan Alam Harus Dihentikan. Makalah disampaikan pada acara Seri Lokakarya Kebijakan Kehutanan, Topik. **1:** 8-9.

Mas, J.-F., H. Puig, J. L. Palacio and A. Sosa-López (2004). "Modelling deforestation using GIS and artificial neural networks." Environmental Modelling & Software**19**(5): 461-471.

Mas, J. F., H. Puig, J. L. Palacio and A. Sosa-López (2004). "Modelling deforestation using GIS and artificial neural networks." Environmental Modelling & Software**19**(5): 461-471.

Matthews, E. (1983). "Global vegetation and land use: New high-resolution data bases for climate studies." Journal of climate and applied Meteorology**22**(3): 474-487.

McLaughlin, D. W. (2011). "Land, food, and biodiversity." Conservation Biology**25**(6): 1117-1120.

Mertz, O., S. J. Leisz, A. Heinimann, K. Rerkasem, W. Dressler, V. C. Pham, K. C. Vu, D. Schmidt-Vogt, C. J. Colfer and M. Epprecht (2009). "Who counts? Demography of swidden cultivators in Southeast Asia." Human Ecology**37**(3): 281-289.

Michalski, F., C. A. Peres and I. R. Lake (2008). "Deforestation dynamics in a fragmented region of southern Amazonia: evaluation and future scenarios." Environmental Conservation**35**(02): 93-103.

Monfreda, C., N. Ramankutty and J. A. Foley (2008). "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000." Global biogeochemical cycles**22**(1).

Morgan, R., J. Quinton, R. Smith, G. Govers, J. Poesen, K. Auerswald, G. Chisci, D. Torri and M. Styczen (1998). "The European Soil Erosion Model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments." Earth surface processes and landforms**23**(6): 527-544.

Morton, D. C., R. S. DeFries, Y. E. Shimabukuro, L. O. Anderson, E. Arai, F. del Bon Espirito-Santo, R. Freitas and J. Morisette (2006). "Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon." Proceedings of the National Academy of Sciences**103**(39): 14637-14641.

Muller, M. R. and J. Middleton (1994). "A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada." Landscape Ecology**9**(2): 151-157.

Murata, N., S. Yoshizawa and S.-I. Amari (1994). "Network information criterion-determining the number of hidden units for an artificial neural network model." Neural Networks, IEEE Transactions on**5**(6): 865-872.

Myint, S. W. and L. Wang (2006). "Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach." Canadian Journal of Remote Sensing**32**(6): 390-404.

Nagy, H., K. Watanabe and M. Hirano (2002). "Prediction of sediment load concentration in rivers using artificial neural network model." Journal of Hydraulic Engineering**128**(6): 588-595.

Nawir, A. and L. Rumboko (2007). "History and state of deforestation and land degradation." Forest Rehabilitation in Indonesia: Where to after More Than Three Decades: 11-32.

Olden, J. D., M. K. Joy and R. G. Death (2004). "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data." Ecological Modelling**178**(3): 389-397.

Overmars, K., G. De Koning and A. Veldkamp (2003). "Spatial autocorrelation in multi-scale land use models." Ecological modelling**164**(2): 257-270.

Palmer, C. (2001). "The extent and causes of illegal logging: An analysis of a major cause of tropical deforestation in Indonesia."

Pijanowski, B. C., D. G. Brown, B. A. Shellito and G. A. Manik (2002). "Using neural networks and GIS to forecast land use changes: a Land Transformation Model." Computers, Environment and Urban Systems**26**(6): 553-575.

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

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

Pontius Jr, R. G., W. Boersma, J.-C. Castella, K. Clarke, T. de Nijs, C. Dietzel, Z. Duan, E. Fotsing, N. Goldstein and K. Kok (2008). "Comparing the input, output, and validation maps for several models of land change." The Annals of Regional Science**42**(1): 11-37.

Pontius, R. G., D. Huffaker and K. Denman (2004). "Useful techniques of validation for spatially explicit land-change models." Ecological Modelling**179**(4): 445-461.

Prentice, I. C., M. T. Sykes, M. Lautenschlager, S. P. Harrison, O. Denissenko and P. J. Bartlein (1993). "Modelling global vegetation patterns and terrestrial carbon storage at the last glacial maximum." Global Ecology and Biogeography Letters: 67-76.

Rajan, D. (2010). Understanding the drivers affecting land use change in Ecuador: An application of the Land Change Modeler software. Masters University of Edinburgh

Ramankutty, N., A. T. Evan, C. Monfreda and J. A. Foley (2008). "Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000." Global Biogeochemical Cycles**22**(1).

Razavi, B. S. (2014). "Predicting the trend of land use change using Artificial Neural Network and Markov Chain Model (Case study: Kermanshah City)." Research Journal of Environmental and Earth Sciences **6**(4): 215-226.

Rist, L., L. Feintrenie and P. Levang (2010). "The livelihood impacts of oil palm: smallholders in Indonesia." Biodiversity and conservation**19**(4): 1009-1024.

Romijn, E., J. H. Ainembabazi, A. Wijaya, M. Herold, A. Angelsen, L. Verchot and D. Murdiyarso (2013). "Exploring different forest definitions and their impact on developing REDD+ reference emission levels: a case study for Indonesia." Environmental Science & Policy**33**: 246-259.

Rumelhart, D. E., G. E. Hinton and R. J. Williams (1985). Learning internal representations by error propagation, DTIC Document.

Sader, S. A. and A. T. Joyce (1988). Deforestation rates and trends in Costa Rica, 1940 to 1983. Mississippi, USA, National Aeronautics and Space Administration, National Space Technology Laboratories, Earth Resources Laboratory.

Sandker, M., A. Suwarno and B. M. Campbell (2007). "Will forests remain in the face of oil palm expansion? Simulating change in Malinau, Indonesia." Ecology and Society**12**(2): 37.

Schneider, L. C. and R. G. Pontius (2001). "Modeling land-use change in the Ipswich watershed, Massachusetts, USA." Agriculture, Ecosystems & Environment**85**(1): 83-94.

Sigit, T. S. (2012). Penertiban Dan Pendayagunaan Tanah Hak Guna Usaha (Hgu) Perkebunan Di Jawa Tengah (Studi Analisis Terhadap Tanah Terlantar), Universitas Muhammadiyah Surakarta.

Sklar, F. H. and R. Costanza (1991). "10. The Development of Dynamic Spatial Models for Landscape Ecology: A Review and Prognosis." Quantitative methods in landscape ecology: the analysis and interpretation of landscape heterogeneity**82**: 239.

Sousa, S., S. Caeiro and M. Painho (2002). "Assessment of map similarity of categorical maps using Kappa statistics." ISEGI, Lisbon.

Strahler, A. H., L. Boschetti, G. M. Foody, M. A. Friedl, M. C. Hansen, M. Herold, P. Mayaux, J. T. Morisette, S. V. Stehman and C. E. Woodcock (2006). "Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps." European Communities, Luxembourg**51**.

Sumargo, W., S. Nanggara, F. Nainggolan and I. Apriani (2009). "Potret keadaan hutan Indonesia: periode tahun 2000-2009." Forest Watch Indonesia, Bogor, Indonesia.[online] URL: http://fwi. or. id/wp-content/uploads/2013/02/PHKI_2000-2009_FWI_lowres. pdf.

Sunderlin, W. and I. A. P. Resosudarmo (1999). "The effect of population and migration on forest cover in Indonesia." The Journal of Environment & Development**8**(2): 152-169.

Sunderlin, W. D. (1996). "Rates and causes of deforestation in Indonesia: towards a resolution of the ambiguities."

Suprayogi, S. (2003). "Prediksi Ketersediaan Air Menggunakan Tank Model dan Pendekatan Artificial Neural Network (Studi Kasus Sub DAS Ciriung Kabupaten Serang)."

Susanti, A. and D. P. Burgers (2012). Oil palm expansion in Riau Province, Indonesia: Serving people, planet, profit? Netherland Overseas Development Institute.

Tang, J., L. Wang and Z. Yao (2007). "Spatio-temporal urban landscape change analysis using the Markov chain model and a modified genetic algorithm." International Journal of Remote Sensing **28**(15): 3255-3271.

Thirumalaiah, K. and M. Deo (1998). "River stage forecasting using artificial neural networks." Journal of Hydrologic Engineering **3**(1): 26-32.

Tokar, A. S. and P. A. Johnson (1999). "Rainfall-runoff modeling using artificial neural networks." Journal of Hydrologic Engineering **4**(3): 232-239.

USDA (2010). INDONESIA: Rising Global Demand Fuels Palm Oil Expansion USA United States Department of Agriculture Foreign Agriculture Service. **2015**.

Veldkamp, A. and E. F. Lambin (2001). "Predicting land-use change." Agriculture, Ecosystems & Environment **85**(1–3): 1-6.

Veldkamp, A. and P. Verburg (2004). "Modelling land use change and environmental impact." Journal of Environmental Management **72**(1): 1-3.

Verburg, P. H., G. H. J. de Koning, K. Kok, A. Veldkamp and J. Bouma (1999). "A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use." Ecological Modelling **116**(1): 45-61.

Verburg, P. H., K. Neumann and L. Nol (2011). "Challenges in using land use and land cover data for global change studies." Global Change Biology **17**(2): 974-989.

Verburg, P. H., K. P. Overmars, M. G. Huigen, W. T. de Groot and A. Veldkamp (2006). "Analysis of the effects of land use change on protected areas in the Philippines." Applied Geography **26**(2): 153-173.

Verburg, P. H., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon and S. S. Mastura (2002). "Modeling the spatial dynamics of regional land use: the CLUE-S model." Environmental management **30**(3): 391-405.

Vermeulen, S. and N. Goad (2006). Towards better practice in smallholder palm oil production, Iied.

Visser, H. and T. de Nijs (2006). "The Map Comparison Kit." Environmental Modelling & Software **21**(3): 346-358.

Waddell, P. (2002). "UrbanSim: Modeling urban development for land use, transportation, and environmental planning." Journal of the American Planning Association **68**(3): 297-314.

Warlina, L. (2007). "Land use change modeling for spatial use management in term of regional sustainable development (with special reference to Bandung Region)."

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**(3): 273-284.

Wicke, B., R. Sikkema, V. Dornburg and A. Faaij (2011). "Exploring land use changes and the role of palm oil production in Indonesia and Malaysia." Land Use Policy **28**(1): 193-206.

Wu, Q., H.-q. Li, R.-s. Wang, J. Paulussen, Y. He, M. Wang, B.-h. Wang and Z. Wang (2006). "Monitoring and predicting land use change in Beijing using remote sensing and GIS." Landscape and Urban Planning **78**(4): 322-333.

Yang, X., X.-Q. Zheng and L.-N. Lv (2012). "A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata." Ecological Modelling **233**(0): 11-19.

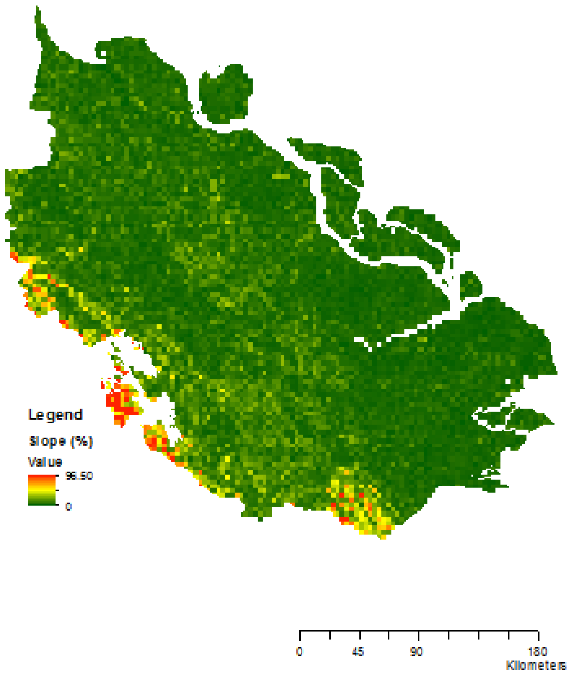
Zander, P. and H. Kächele (1999). "Modelling multiple objectives of land use for sustainable development." Agricultural Systems **59**(3): 311-325.

INPUT VARIABLES OF LAND USE CHANGE MODELLING

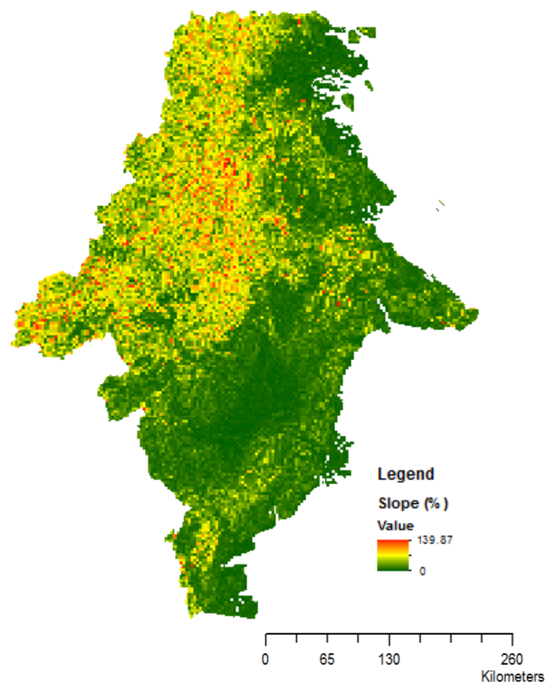
Slope Maps

Slope map was downloaded from ftp://xftp.jrc.it/pub/srtmV4/tiff/. The legend represents the value of slope in percentage.

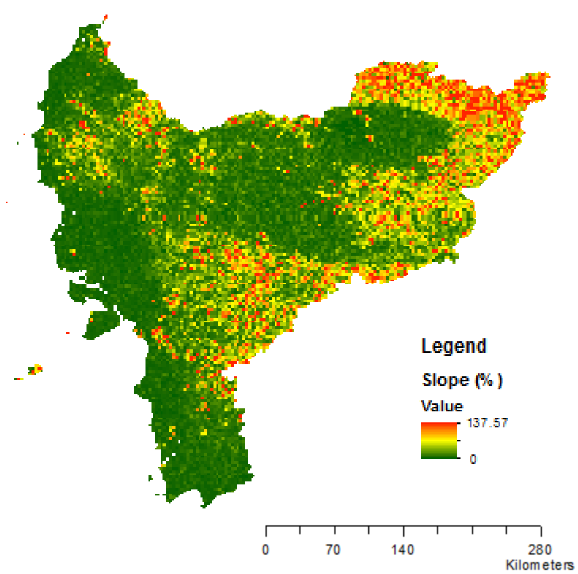
Riau



East Kalimantan



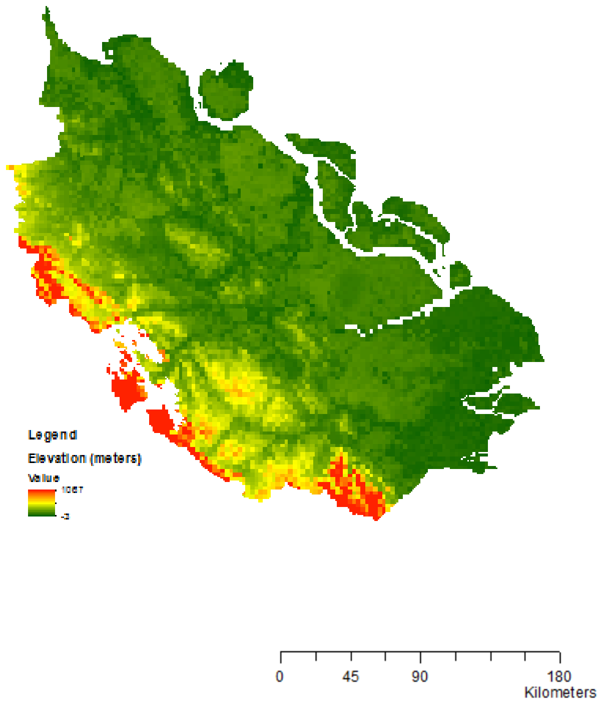
West Kalimantan



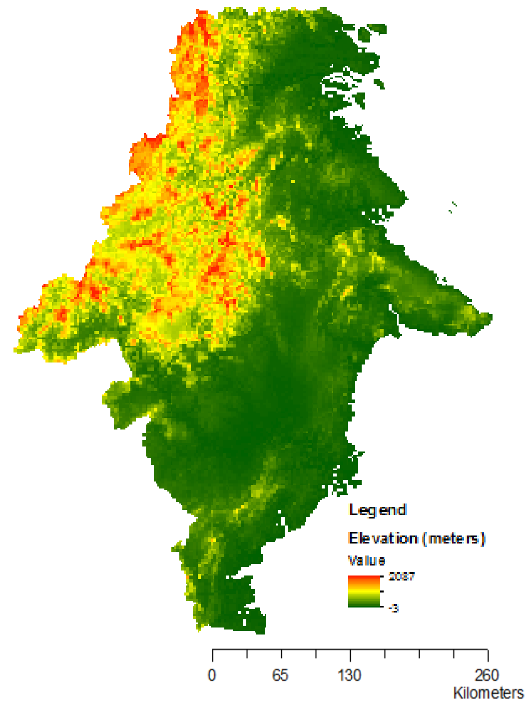
Elevation Maps

Elevation map was downloaded from http://srtm.csi.cgiar.org. The legend represents the elevation in meter.

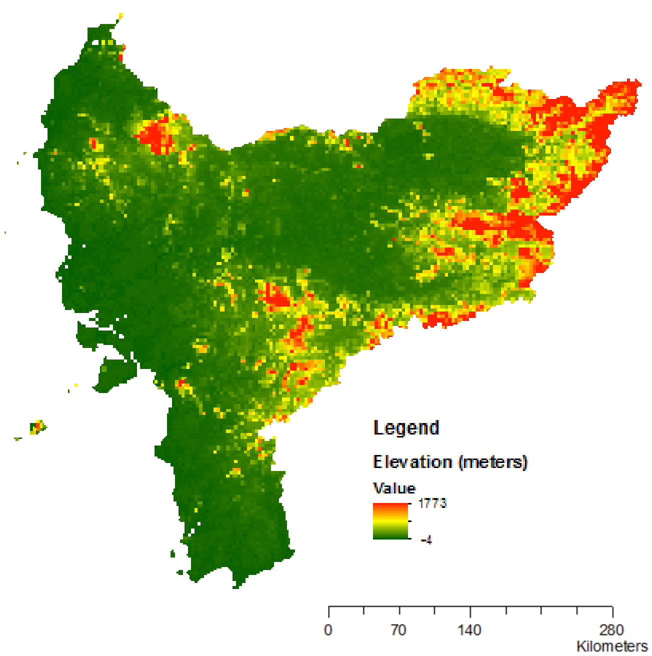
Riau



East Kalimantan



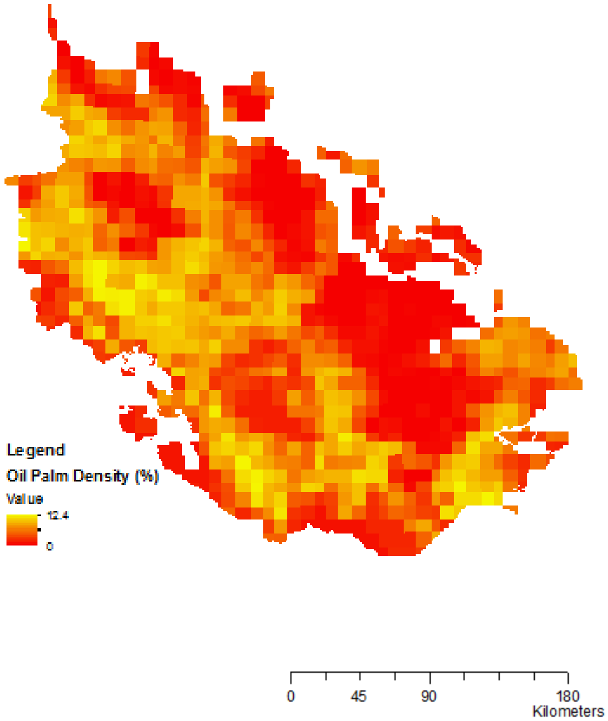
West Kalimantan



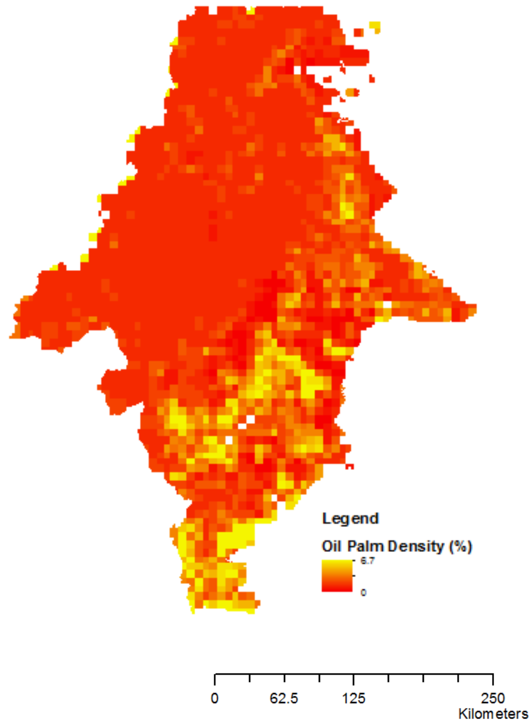
Oil palm maps

Oil palm dataset was downloaded from http://www.geog.mcgill.ca/~nramankutty/Datasets/Datasets.html. The legend values represent the percentage of oil palm area in each pixel.

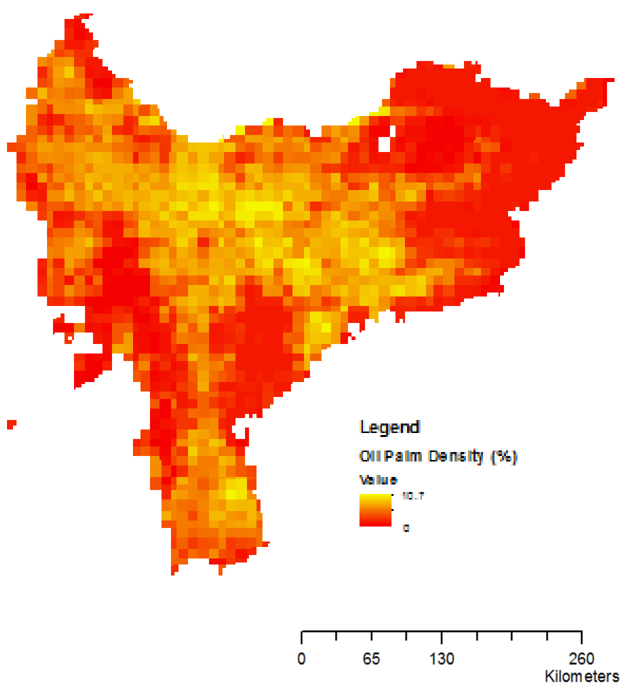
Riau



East Kalimantan



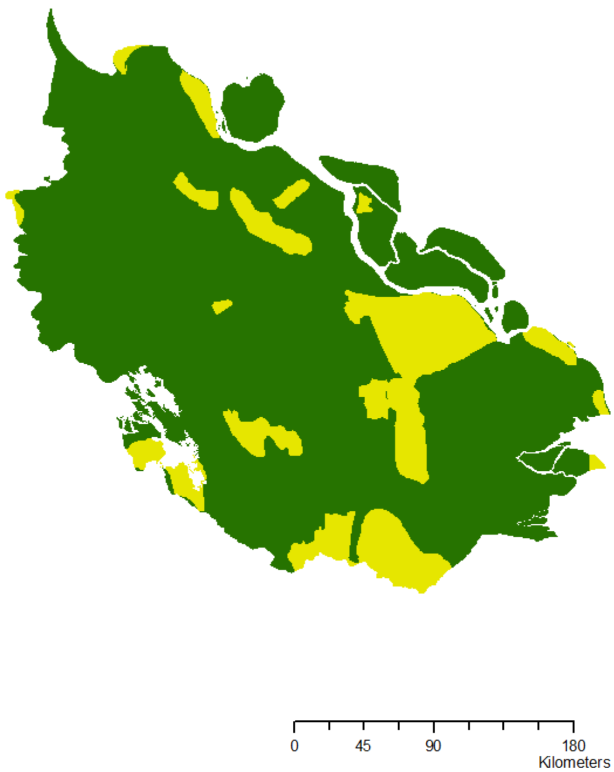
West Kalimantan



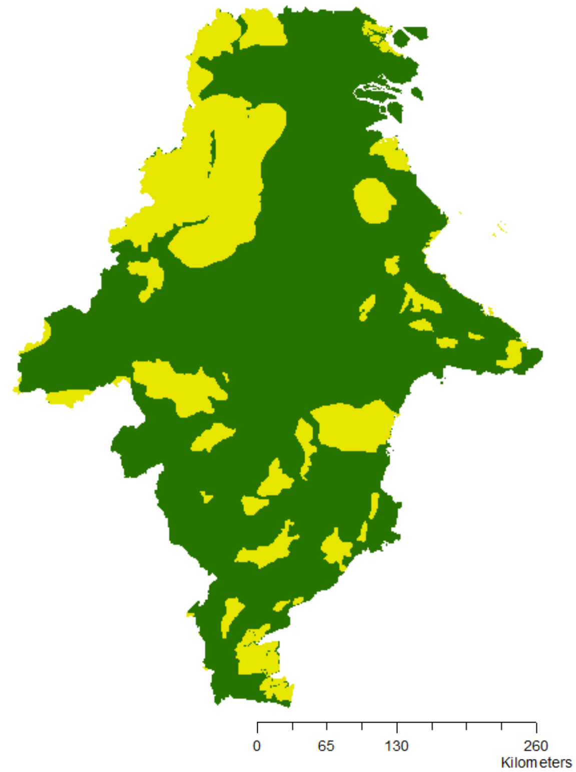
Protected area maps

World protected area map was downloaded from www.protectedplanet.net. Original map was available in vector format. The map was reclassified into 0 as protected area and 1 as non-protected area.

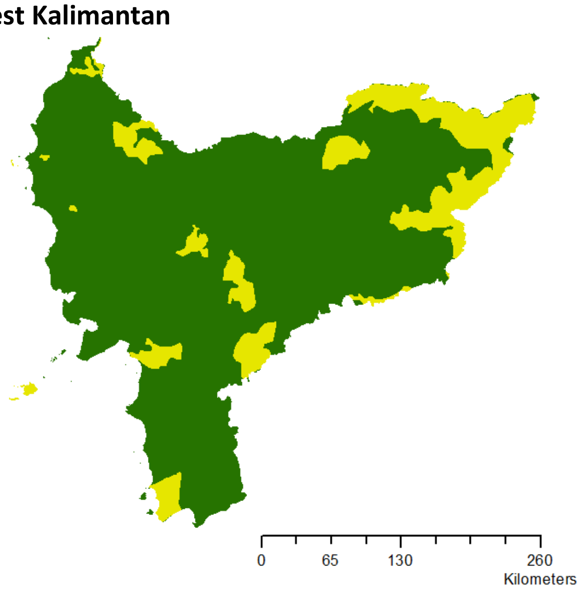
Riau



East Kalimantan



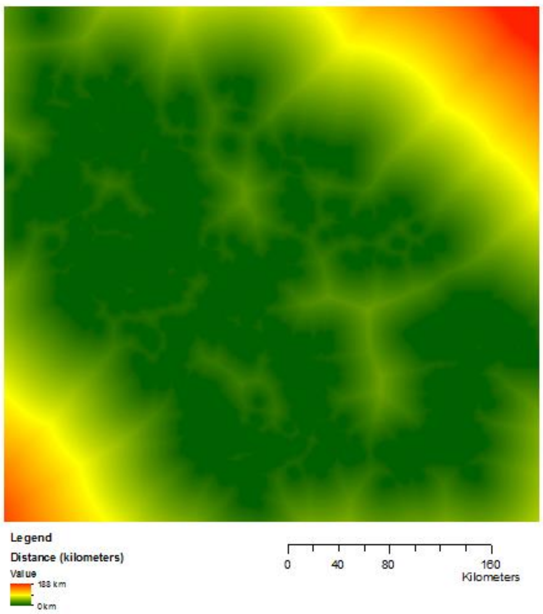
West Kalimantan



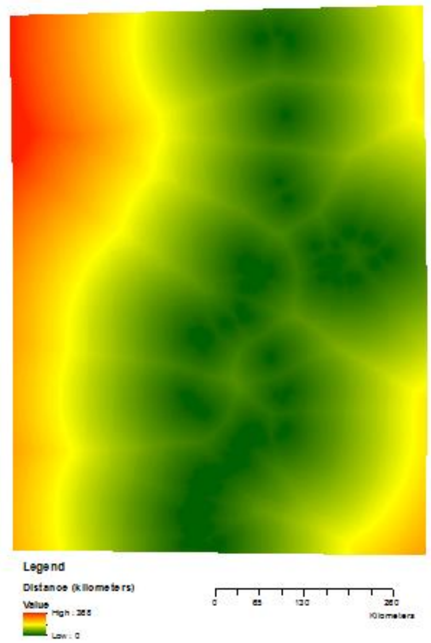
Distance from existing oil palm area maps

Distance from existing oil palm area was calculated using Euclidean distance from LU map from the MoF for the year of 2000.

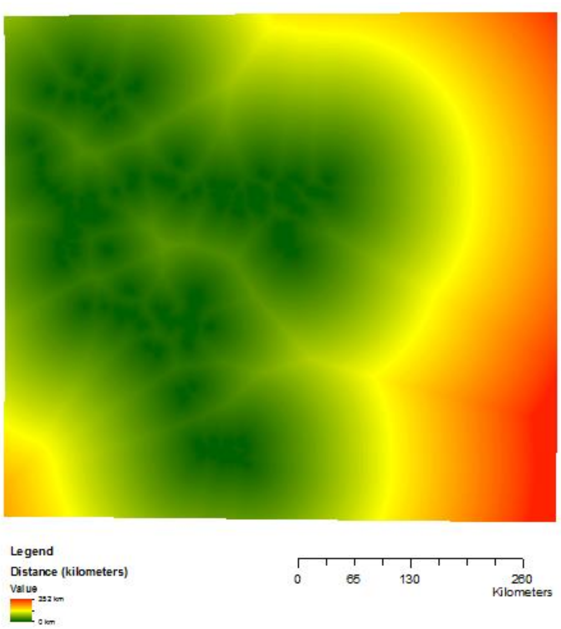
Riau



East Kalimantan



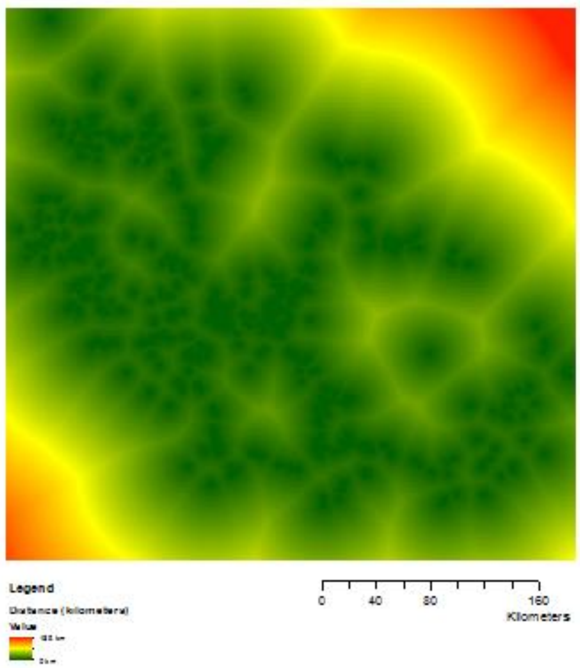
West Kalimantan



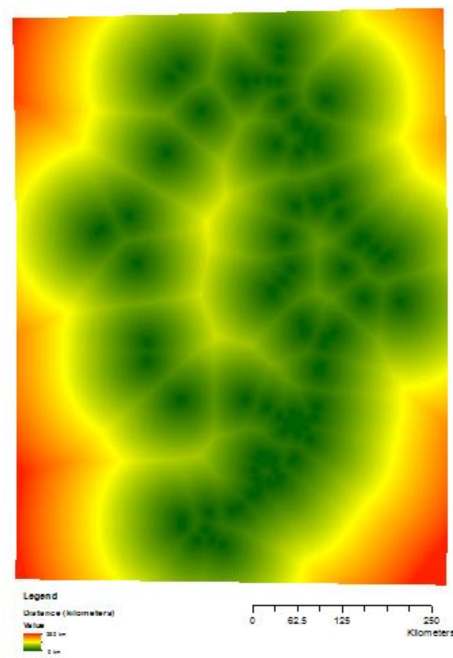
Distance from existing settlements area maps

Distance from existing settlement area was calculated using Euclidean distance from LU map from the MoF for the year of 2000.

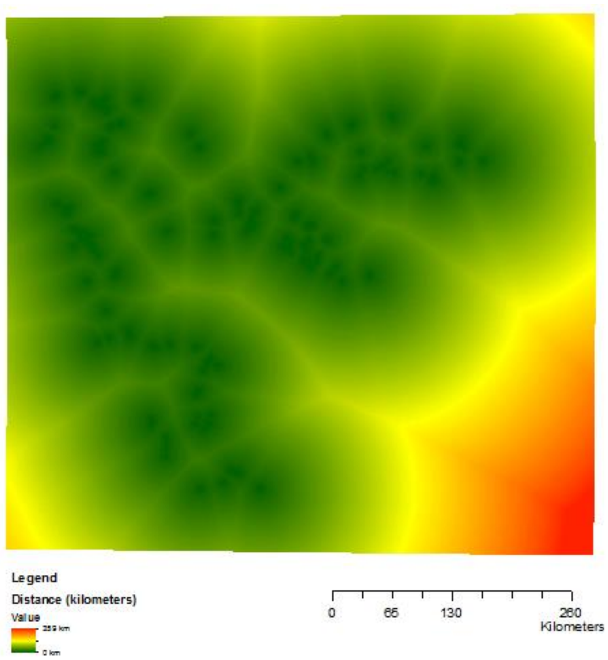
Riau



East Kalimantan



West Kalimantan



APPENDIX 2

Variables performance in sub-model status

This appendix describes the performance on each independent variable listed in Table III.4 to contribute on the accuracy of the sub-model, and part of sub-model calculation result. The explanatory power of each independent variable is explained in the description below, while Table 2.1 – 2.3 shows the performance of each independent variable in sub-model structure, given the accuracy of the sub-models when the referred variable was held constant, as the result of first statistical test conducted in transitional sub-model structure step (see Section III.3.5). The column heading represent the sub-model class that was evaluated. The absence of values in certain sub-model class occurred because there were no class of change of that sub-model in the area.

Riau

Table 2. 1 – 2.3 The performance of each independent variable in transitional sub-model structure. The value represents percentage of model accuracy when the variable referred was held constant. The higher differences between accuracy when all variables were included (in the topmost of the row) and when the correlated variable was held constant, the higher the degree of importance of the variable is.

	F – AG	F – OP	F – OT	OT – OP	OT – AG	AG – OP
With all variables	76.14	74.85	63.09	50.24	96.81	-
Distance from existing oil palm plantation	61.30	62.80	60.29	50.24	94.18	-
Distance from settlement	57.72	65.48	58.72	50.24	91.89	-
Oil palm	75.70	74.53	60.22	50.24	94.40	-
Slope	75.91	74.46	62.98	50.24	96.69	-
Elevation	74.77	73.72	61.70	50.24	88.24	
Protected area	73.61	73.27	60.02	50.24	93.84	-

East Kalimantan

	F – AG	F – OP	F – OT	OT – OP	OT – AG	AG – OP
With all variables	73.34	87.25	78.03	91.58	50.45	91.38
Distance from existing oil palm plantation	73.91	49.74	80.37	57.03	50.45	61.60
Distance from settlement	50.27	86.49	77.78	89.87	50.45	88.99
Oil palm	72.85	87.19	57.35	91.34	50.45	91.48
Slope	73.36	87.26	62.83	91.30	50.45	91.37
Elevation	72.81	87.26	62.17	91.34	50.45	91.37
Protected area	73.82	87.32	77.65	89.48	50.45	65.02

West Kalimantan

	F – AG	F – OP	F – OT	OT – OP	OT – AG	AG – OP
With all variables	71.58	87.65	80.10	-	91.19	84.56
Distance from existing oil palm plantation	51.23	50.00	72.53	-	49.04	44.95
Distance from settlement	70.05	87.51	70.75	-	90.63	84.17
Oil palm	70.59	87.39	79.36	-	91.21	84.50
Slope	70. 99	87.55	79.78	-	91.48	84.57
Elevation	71.41	87.23	79.88	-	91.19	84.56
Protected area	71.05	84.50	79.85	-	88.72	84.76

Distance from the existing oil palm plantations

Distance from existing oil palm plantation performed differently in three study area. In Riau, this variable only significant in conversion from forest to agriculture and forest to oil palm only, while in the other sub-model, this variable only give less than 3% of difference if it held constant. In East Kalimantan, this variable has high contribution in all changes related with oil palm, which are forest to oil palm, other to oil palm and agriculture to oil palm, by reducing the accuracy for more than 25%. Meanwhile, in West Kalimantan, this variable was significant in all sub-models, with accuracy difference for more than 30%.

Distance from the existing settlements

This variable performs significantly in all changes in Riau, except for other to oil palm sub-model. In East Kalimantan, distance from settlement only significant in the change from forest to agriculture class, while in other changes it only contribute to change the overall accuracy for less than 3%. Meanwhile, in West Kalimantan, this variable only significant in forest to other change, with the difference for almost 10% by keeping the variable to be constant in the calculation.

Oil palm dataset

Oil palm dataset perform less significant in almost all sub-models in all provinces. This variable only performs significant in forest-other class in East Kalimantan, showed by the decrease of accuracy for 20.68% when this variable was held constant.

Slope

This variable only give small contribution to overall performance for all sub-models, except in forest to other in East Kalimantan, by giving the differences of accuracy for almost 8% when this variable was held constant.

Elevation

As well as slope, elevation only performs significant in forest to other class in East Kalimantan, by decreasing the accuracy for 15.86% when this variable was held constant. In other sub-models, this variable only contributes to decrease accuracy for less than 3% when it held constant.

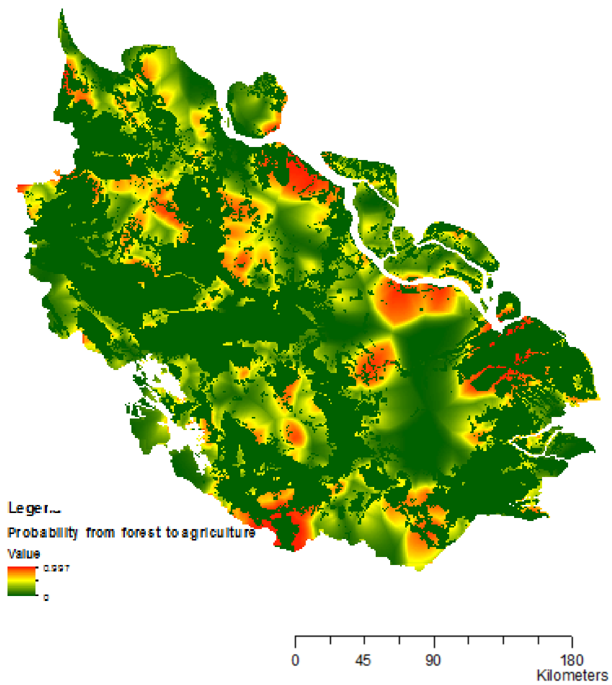
Protected area

Overall differences from total accuracy while keeping PA map to be constant in Riau will only give range of value between 1% - 3%, therefore we consider that this variable only have low significance in all sub-models in Riau. Meanwhile, in East Kalimantan, protected area contributes a significance value of accuracy in agriculture-oil palm. Keeping this variable to be constant will reduce the accuracy of sub-model of 26.36%, while in West Kalimantan, protected area did not give significant influences for all sub-model classes.

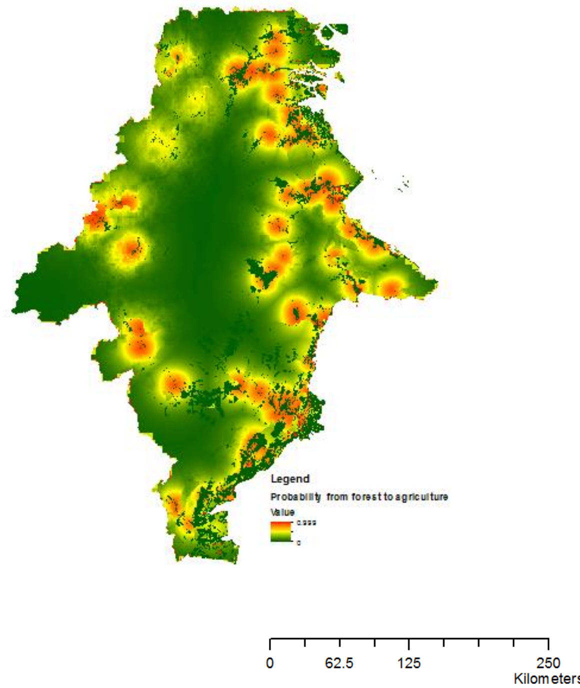
Transitional Probability Maps

Probability Maps from Forest to Agriculture

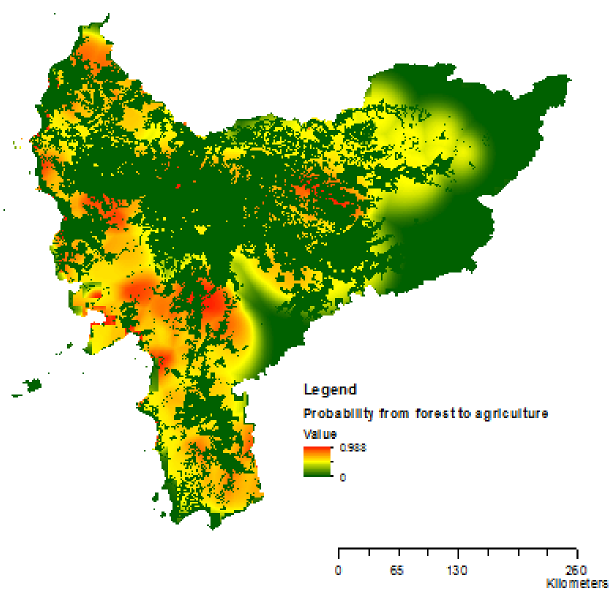
Riau



East Kalimantan

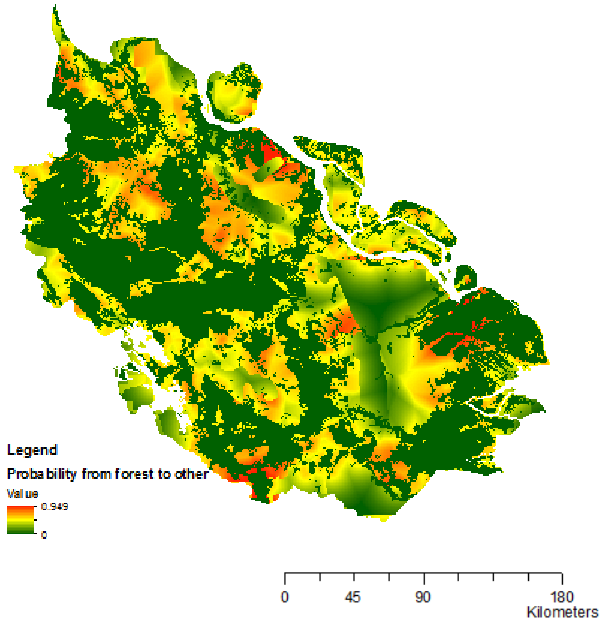


West Kalimantan

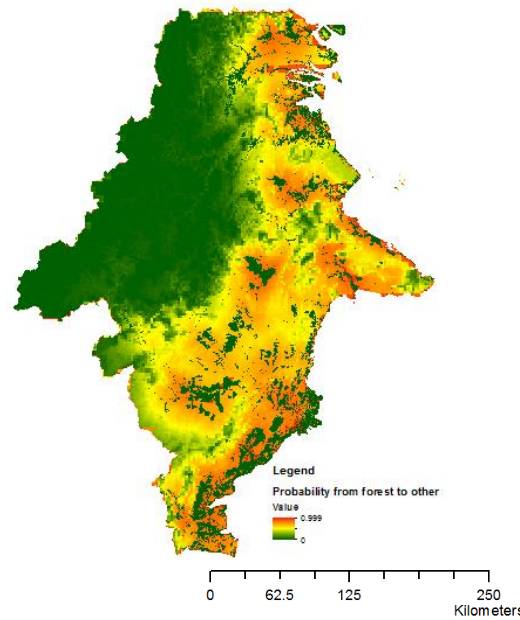


Probability Maps from Forest to Other

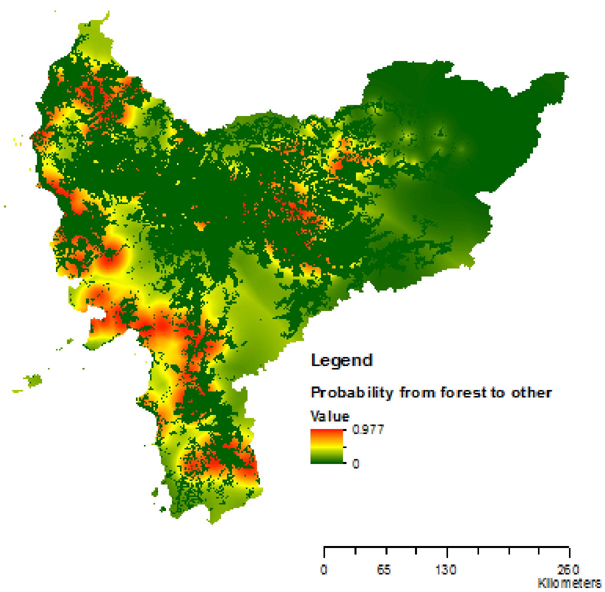
Riau



East Kalimantan

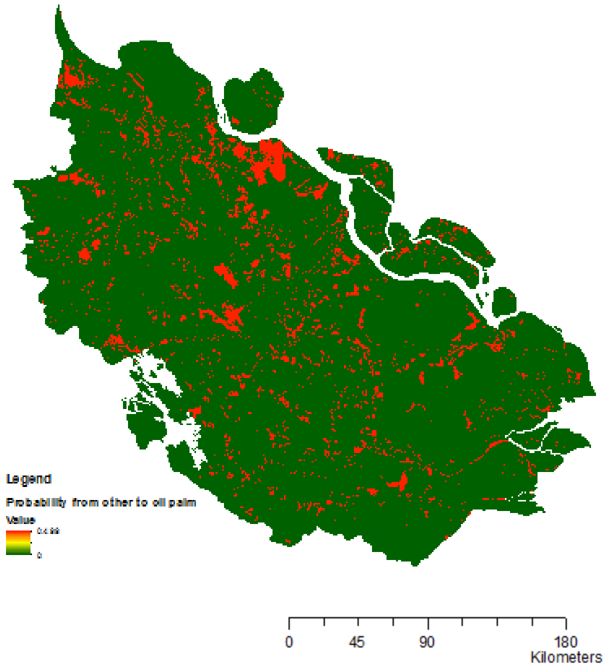


West Kalimantan

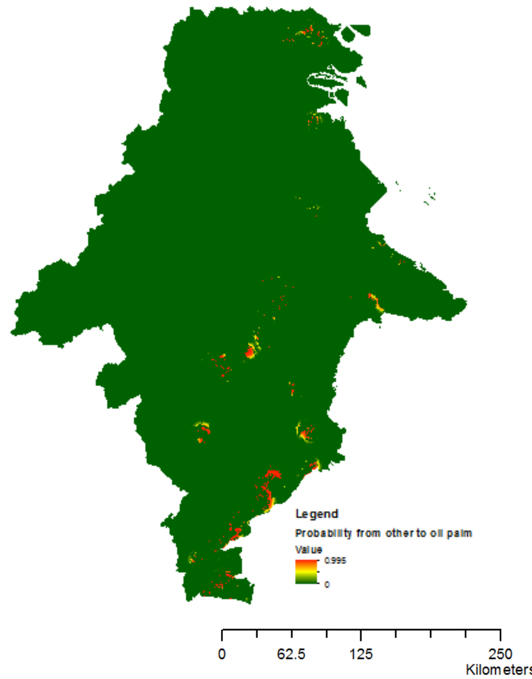


Probability Maps from Other to Oil Palm

Riau

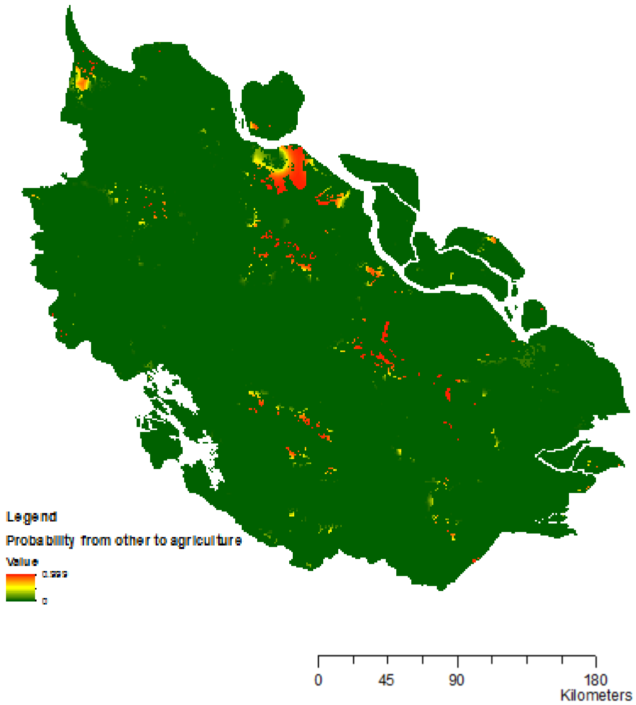


East Kalimantan

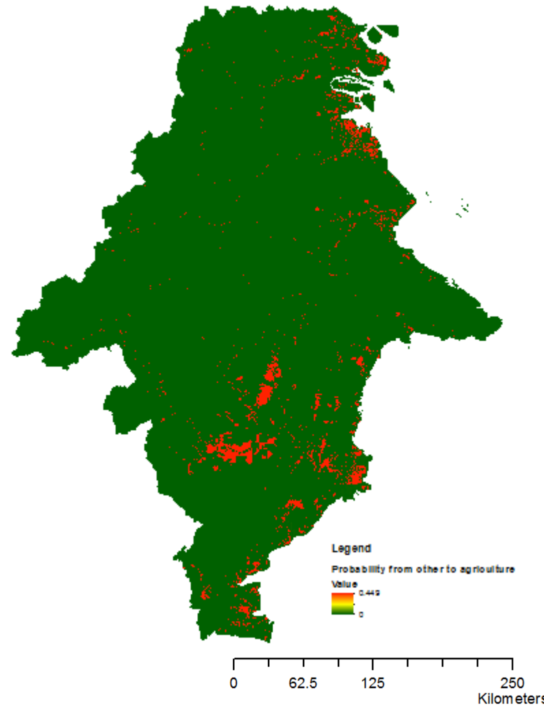


Probability Maps from Other to Agriculture

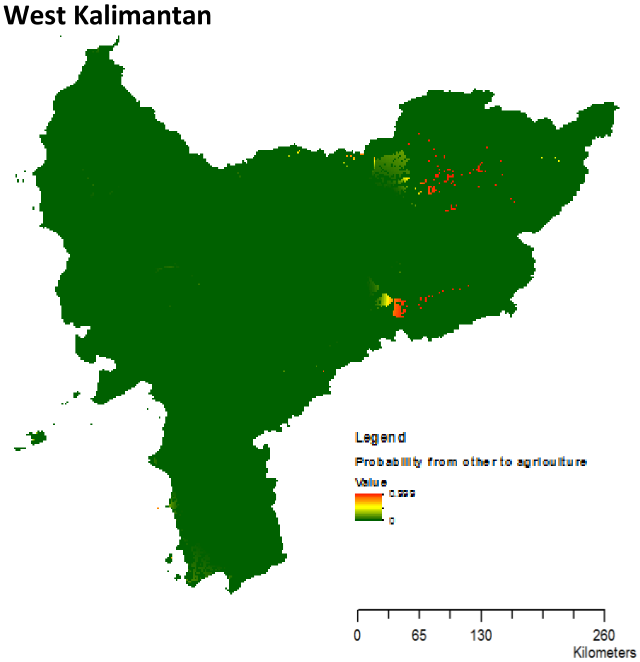
Riau



East Kalimantan

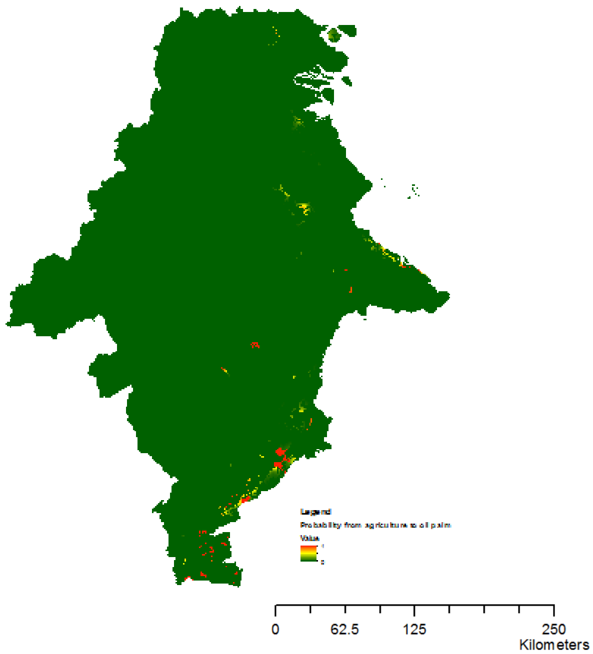


West Kalimantan

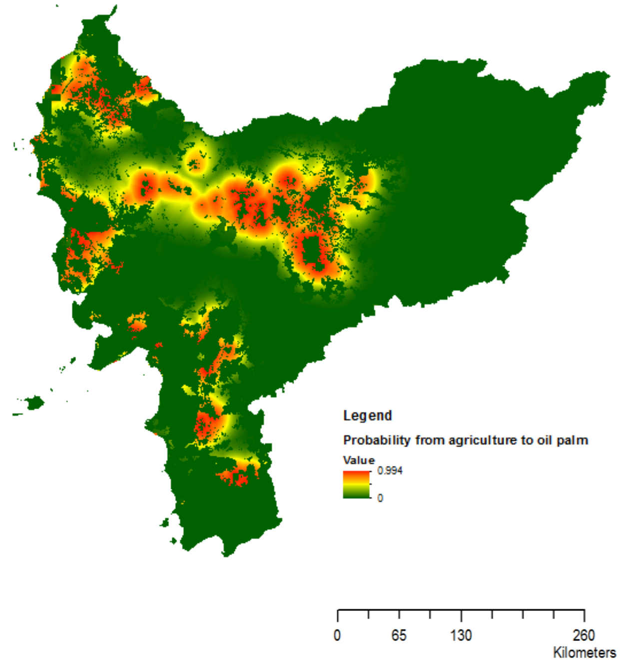


Probability Maps from Agriculture to Oil Palm

East Kalimantan



West Kalimantan

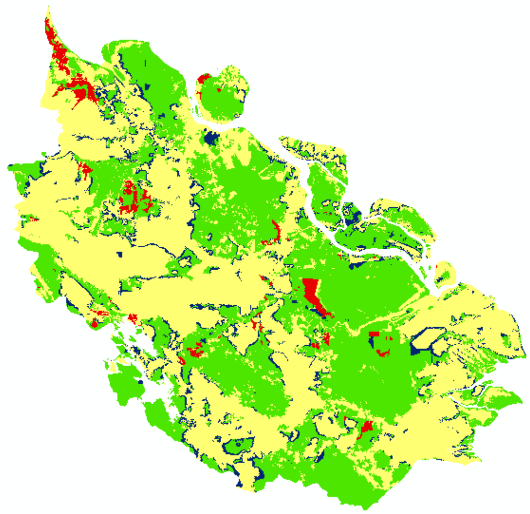


Per Category Map Comparison.

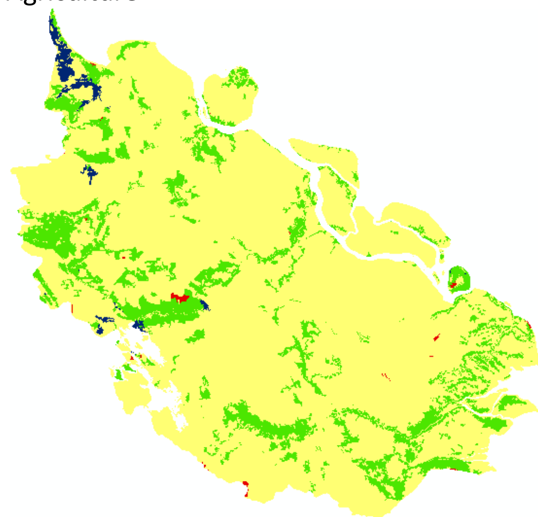
The maps were produced by overlaid 2009 initial land use map from the MoF (map 1) with 2009 predicted map generated from the model (map 2).

Riau

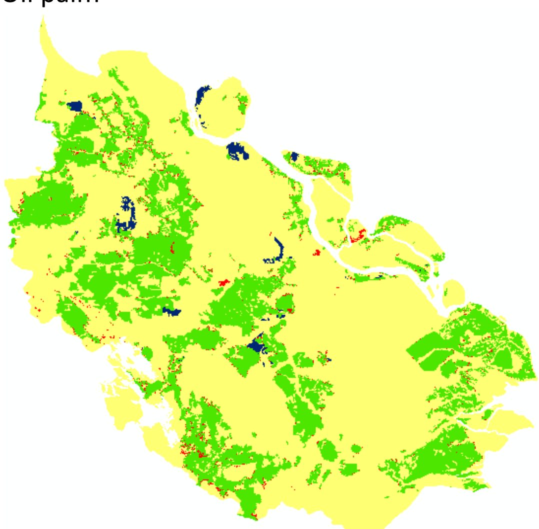
Forest



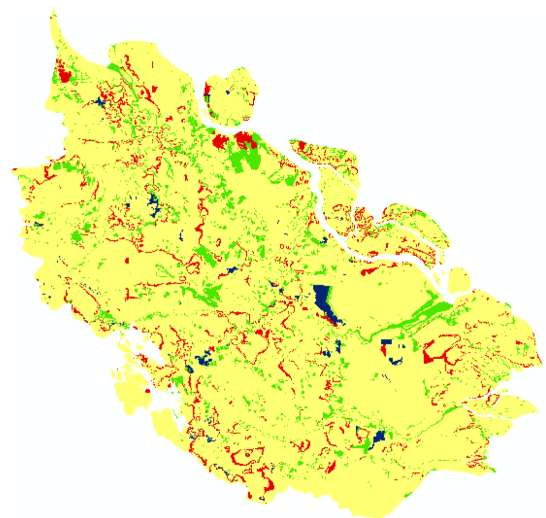
Agriculture

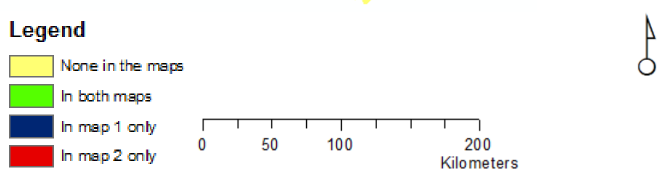


Oil palm



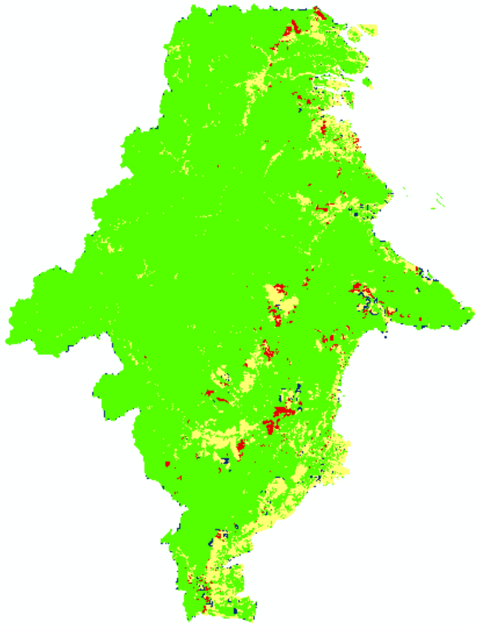
Other



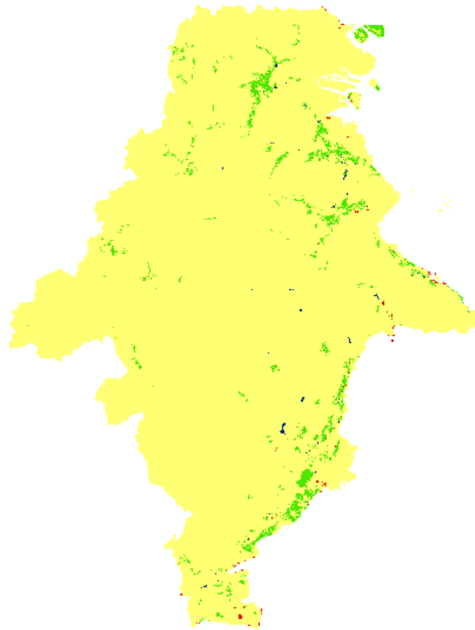


East Kalimantan

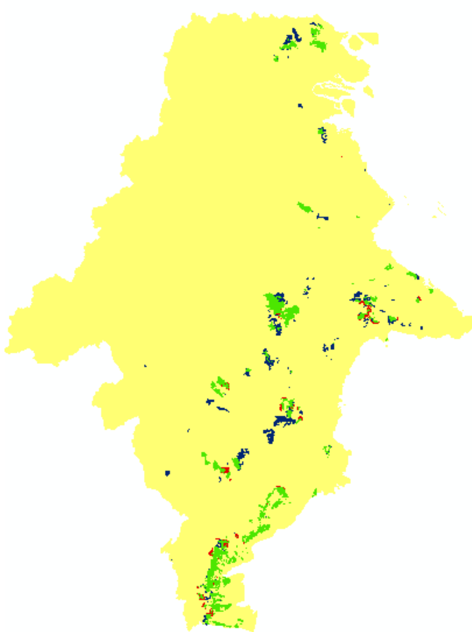
Forest



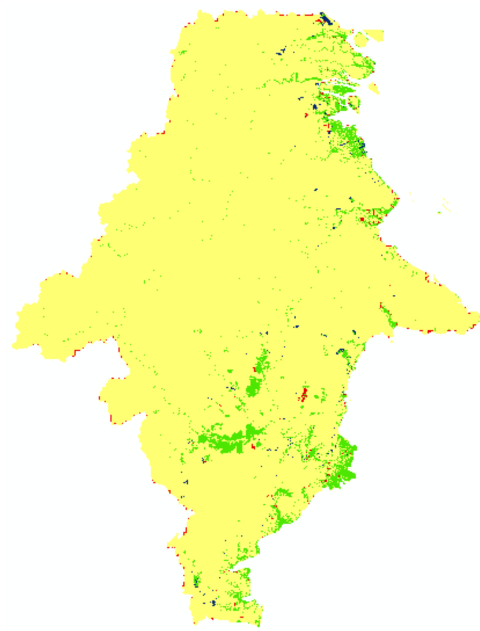
Agriculture

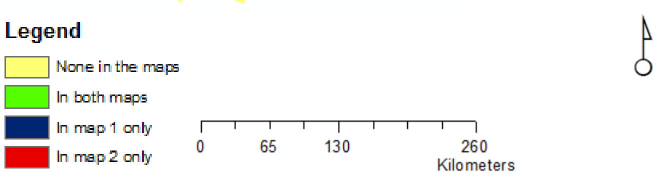


Oil palm



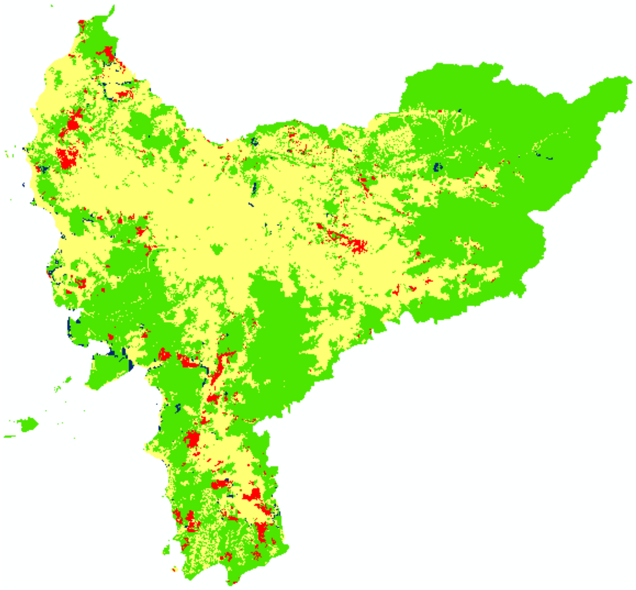
Other



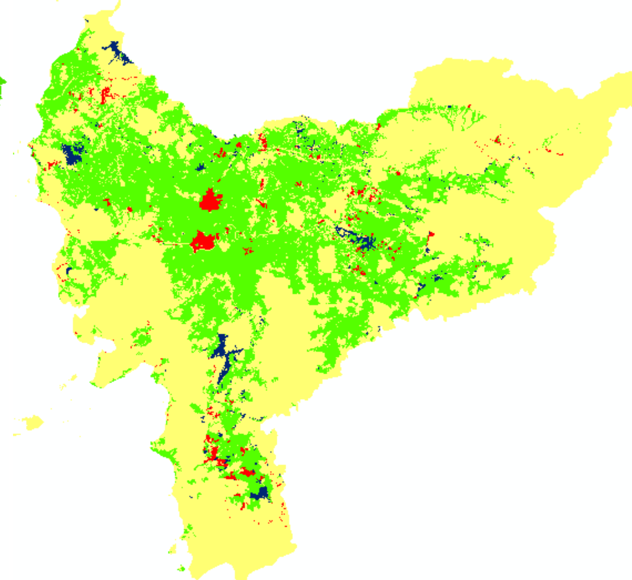


West Kalimantan

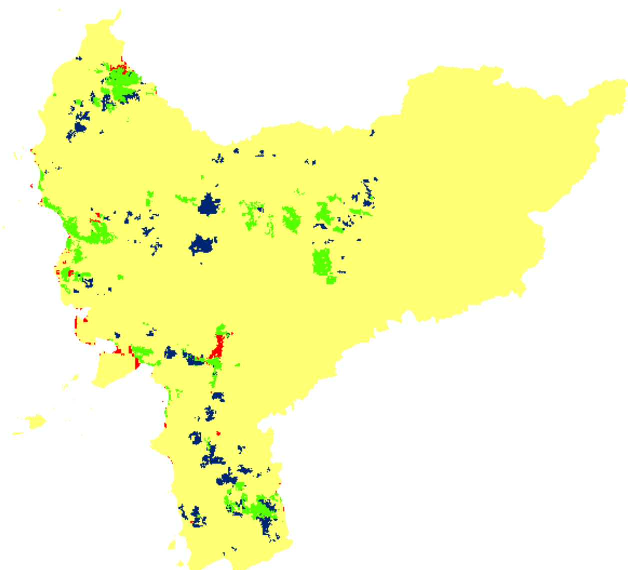
Forest



Agriculture



Oil palm



Other