GIMA

Geographical Information Management and Applications

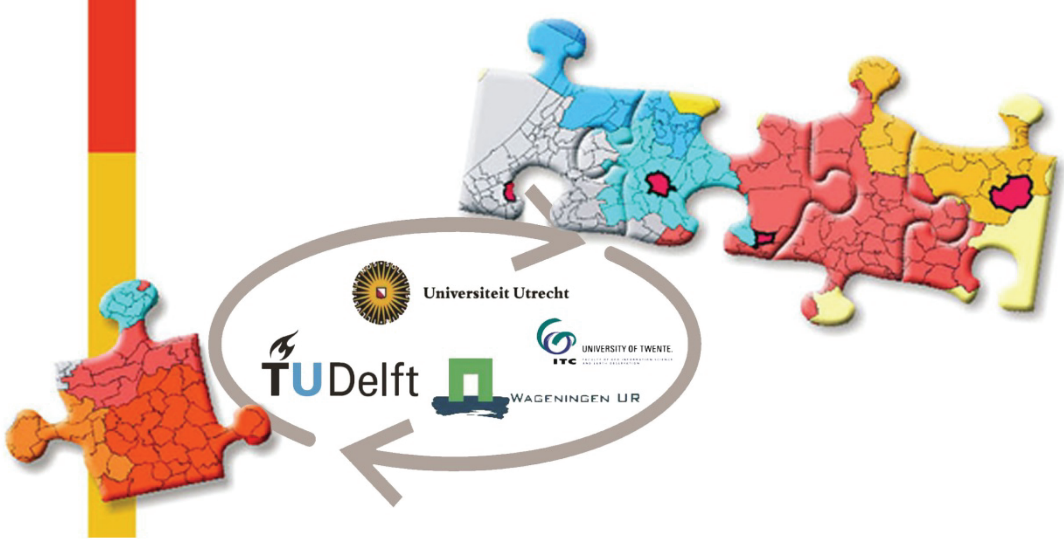
To see the forest for the trees

Understanding drivers and processes involved in deforestation and modelling forest change dynamics in central Vietnam

Master thesis

Astrid Brigitte Bos

November 2013



GIMA

Geographical Information Management and Applications

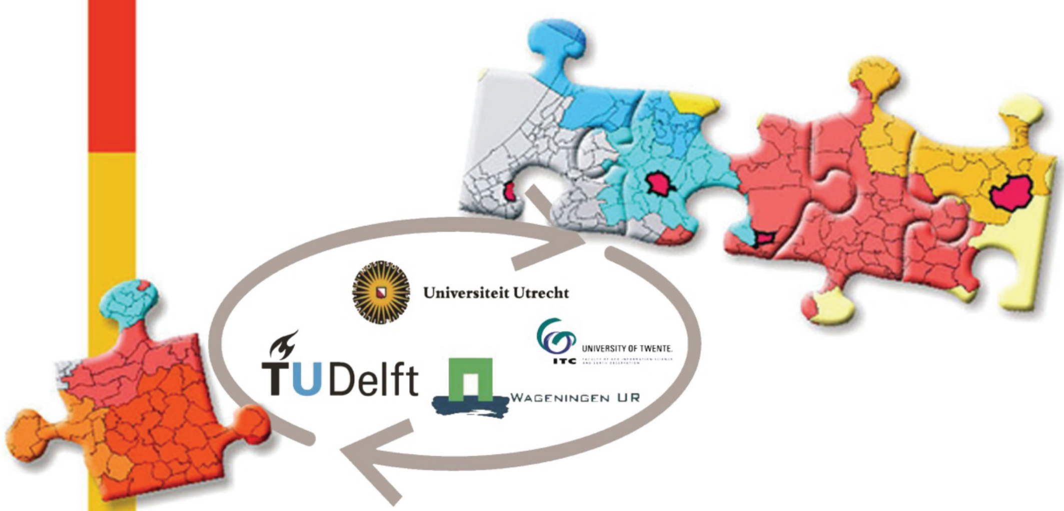
To see the forest for the trees

Understanding drivers and processes involved in deforestation and modelling forest change dynamics in central Vietnam

Master thesis

Astrid Brigitte Bos

November 2013



To see the forest for the trees

Understanding drivers and processes involved in deforestation and modelling forest change dynamics in central Vietnam

MSc thesis Geographical Information Management and Applications (GIMA)

Version 1.0
8 November 2013

Student:	A.B. Bos MSc
Email:	astridbos[at]gmail.com
ITC/UTwente student number:	s6000614
Utrecht University student number:	3475107
Wageningen University student number:	870508-102-100
Supervision:	Dr. V. Avitabile (WUR)
	Dr. ir. A. Ligtenberg (WUR)
Co-supervision:	G. Salvini MSc (WUR)
	M. Schultz MSc (WUR)
Professor:	Prof. dr. ir. A.K. Bregt (WUR)

The landscape we see is not a picture frozen in time only to be cherished and protected. Rather it is a continuing story of the earth itself where man, in concert with the hills and other living things, shapes and reshapes the ever changing picture which we now see. And in it we may read the hopes and priorities, the ambitions and errors, the craft and creativity of those who went before us. We must never forget that tomorrow it will reflect with brutal honesty the vision, values, and endeavours of our own time, to those who follow us.

Author unknown,
as cited in Teahan (2010)

Abstract

Concerns about the scarcity of our natural resources and the widespread effects of climate change are part of the most common debates in the world nowadays. The Land Use and Climate Change interactions (LUCCi) project in the Vu Gia- Thu Bon (VGTB) river basin (Central Vietnam), aims at developing strategies for sustainable land use.

Land use and land cover changes involve complex interconnected processes. The objective of this research is to understand the drivers and processes involved in deforestation and forest degradation and to model these processes in order to provide insights in future deforestation risk areas.

The research is founded on the complex adaptive systems (CAS) theory. Socio-ecological systems can be considered CAS, as they involve many variables, are highly dynamic and its different components adapt or learn as they interact.

The methodology of this research is characterised by two main phases. In the first phase, land cover data from Landsat TM satellite imagery from 2001, 2005 and 2010 was used in combination with spatial data on eight potential correlating factors to deforestation. The eight factors considered are elevation, slope, and six distance factors (distance to cropland, grassland, small settlements, large settlements, all roads and paved roads). Statistical tests compared forest change cells with unchanged forest cells (i.e. the control group). It was found that all tested factors showed a significant correlation with deforestation. The most important factors were found to be distance to cropland and distance to small settlements.

In the second phase of the research, the insights from the previous phase were used as input for the model design of an agent-based model, called SoDRA LUCCi. The model simulates future deforestation risk areas under a business-as-usual scenario and the effects of REDD measures on the projected deforestation for 2010-2020. The agents represent rural households, who are considered to be the key decision-making entities regarding land use and land use change.

The model was calibrated and parameterised by using the correlation and other statistical results from the previous phase. The best-fit calibration method resulted in a good quantitative representation of modelled deforestation when compared to the measured deforestation of 2001-2010. A qualitative pattern check between modelled and measured deforestation showed that the scattered deforestation in remote areas is underrepresented in the model, causing false negatives in the model results. Still, modelled deforestation cells were mostly within an acceptable distance from measured deforestation. According to the local sensitivity analysis, the most sensitive parameters are the one related to the distance to cropland and the parameter that defines the threshold below which the location factors of a forest patch are unsuitable for deforestation, called deforestation-potential-point.

For the 2010-2020 era, the SoDRA LUCCi model predicts a scattered pattern of deforestation in the VGTB area, with the highest concentrations in the north west and centre of the region. With regards to the REDD scenarios, it can be concluded that measures that implement a quota defining the maximum deforestation per household have the largest impact compared to the modelled business-as-usual scenario. Measures that enforce prohibition or reduction of deforestation in existing protected areas or in areas with high carbon stocks are expected to have only limited influence on reducing deforestation. In order to achieve the projected effects as modelled in SoDRA LUCCi, particular REDD measures may focus on (financial) incentives, capacity building and technology transfer for stimulating (alternative) sustainable livelihood activities and strategies.

The first version of the SoDRA LUCCi model is still relatively simple. The model can be improved by distinguishing between agent types. Therefore, it is necessary to have proper socio-economic data on agent behaviour.

The research concludes that agent-based modelling provides a tool for revealing large-scale patterns that are induced by micro-level actions. Rather than getting lost in a forest of details, it offers an instrument for greater understanding of the bigger picture while acknowledging that those details form the backbone of the system. To see the forest for the trees…

Keywords:
agent-based modelling; complex adaptive systems; deforestation; land use/land cover change; livelihoods; REDD; spatial analysis; Vietnam

Acknowledgements

This thesis is the product of nine months of work. However, this piece of paper is just one of the outcomes. In addition, the study has given me some “intangible” results, insights such as: it is never too late to learn something completely new; willingness to learn is more important than previously acquired skills; self-discipline is not something you acquire, but something that is –unfortunately- tested every single day; and sharing experiences, doubts and successes with those around me is more important than I imagined.

Therefore, I would like to seize the opportunity to thank everyone who helped me during this research.

First, I would like to thank everyone in Wageningen for sharing their knowledge, for their cooperation, for their advices and for their company. In particular, I would like to thank Michael Schultz for his patience and for helping me in the first phases of the research, Giulia Salvini for her concrete ideas during the sometimes abstract discussions, Valerio Avitabile for all his time and guidance and for the practical insights in times of doubt, Arend Ligtenberg for his practical tips in the final research phase I –unjustly- feared the most, and Arnold Bregt for reviewing this thesis and for his support as professor of the Laboratory of Geo-information Science and Remote Sensing. Furthermore, I would like to thank the people from the LUCCi project for making this research possible and for providing most of the data that was needed for this research.

Unlike previous research projects, this thesis was fully executed in the Netherlands. Still, I guess for some of the people around me it must have felt like I was living on an island for the past nine months. I would like to thank all my friends and family for their patience, and for motivating and supporting me.

Linda & Iris, thank you for all the nice diners and lunches and for those times I needed a place to stay. It is such a nice experience to start and end my student “career” in Wageningen, especially in your company.

I would like to thank my parents and Janine and Jasper for their endless support, and for pushing me to keep going. Finally, Mark, for all your help, advice, confidence, patience, motivating words and last but not least your encouraging post-it!

8 November 2013

Astrid Bos

Table of contents

List of figures

List of tables

List of appendices

Abbreviations & acronyms

ABM	Agent-based model
BAU	Business as usual
CA	Cellular automata
CAS	Complex Adaptive Systems
DEM	Digital Elevation Model
FIPI	Forest Inventory and Planning
GIS	Geographic information system
GSO	General Statistics Office Vietnam
LUCC	Land use and land cover change
LUCCi	Land Use and Climate Change Interactions in Central Vietnam
N/A	Not applicable
ODD	Overview, Design concepts and Details
POM	Pattern-oriented modelling
REDD	Reducing Emissions from Deforestation and Forest Degradation
SoDRA	Simulation of Deforestation Risk Areas
UNFCCC	United Nations Framework Convention on Climate Change
VGTB	Vu Gia-Thu Bon River Basin

1. Introduction

1.1 Background

1.1.1 Problem definition & introduction to the LUCCi project

Concerns about the scarcity of our natural resources and the widespread effects of climate change are part of the most common debates in the world nowadays. These issues are high on the agenda of the government of Vietnam as well. Population growth, economic development and changing climate conditions have put a pressure on Vietnam's land and water resources, especially during the last decade. Furthermore, Vietnam is one of the countries expected to be most severely affected by climate change impacts, and the central part of the country is currently one of the areas with the highest human-induced impacts (Ribbe, 2010). For these reasons, the Land Use and Climate Change interactions (LUCCi) project in Central Vietnam, started in 2010, is aimed at developing strategies for sustainable land use considering the regional socio-economic development, national planning as well as climate change predictions, greenhouse gas (GHG) emission estimates and natural land and water resources. This Vu Gia- Thu Bon river basin (VGTB) was chosen to be the focus area for the LUCCi project, because it is considered a critical region. Its land use and climate change characteristics are highly dynamic, it is one of the most important catchment areas of Vietnam, and it is a highly populated area, which comprises a wide range of patterns related to climate change that the project aims to address (Ribbe, 2010).

Although the LUCCi project focuses on both land and water resources, this study will first and foremost focus on forest-related changes in the study area. In order to understand the forest change dynamics and drivers in the area, one need to study the land cover and land use. Following Veldkamp and Fresco (1996, p. 254), land use "is determined by the interaction in space and time of biophysical factors (constraints) such as soils, climate, topography, etc., and human factors like population, technology, economic conditions, etc." Although the terms *land use* and *land cover* are often used interchangeably, there is a clear distinction between the two. Land *cover* includes the attributes of the Earth's land surface and its immediate subsurface, including biota, soil, topography, surface and groundwater, and human (mainly built-up) structures. Land *use*, however, is about the purposes for which humans exploit the land cover (Lambin, Geist & Rindfuss, 2006). Together, land cover and land use change are often referred to as *land change* (see for example Bakker & Veldkamp, 2008) or denoted as land-use/cover change (LUCC).

Undeniably, land use and land cover processes are interconnected in a complex way. Therefore, predicting future land changes demands for a clear understanding of these processes. In a sense, modelling the future is often a backward route, in which one first has to understand the past land cover patterns and land use characteristics in order to define rule sets as input for the model. As Lambin et al. (2006, p. 1) put it "predicting how land-use changes affect land degradation, the feedback on livelihood strategies from land degradation, and the vulnerability of places and people in the face of land-use/cover changes requires a good understanding of the dynamic human-environment interactions associated with land-use change".

This implies that land *use* data can be derived from census and survey data and expert knowledge whereas land *cover* data can be deduced from remote sensing data (Bakker & Veldkamp, 2008). For prediction modelling, it is necessary to distinguish between these two issues, which appear to be not always straightforward. When one wants to know at *what rates* changes take place, one needs to focus on the land use commodity quantities. These are based on the changes in the demand for these commodities. The *location* of the changes is linked to the land cover patterns and can be predicted based on the knowledge on local proximate causes directly linked to the land changes (Bakker & Veldkamp, 2008).

Studying these LUCC patterns however, is a complex task. LUCC is a process that includes actors and factors at different social and spatial levels (Valbuena, Verburg, Bregt & Ligtenberg, 2010). Understanding these interrelationships related to forest changes in the VGTB area is therefore a true challenge.

1.1.2 Study area

The Vu Gia Thu Bon River Basin is an area situated in Central Vietnam. The region covers an area of over 10,000 km^2 and is home to almost two million inhabitants (Huong & Viet, 2009) although the population density varies widely over the area with large urban settlements in the East and only few rural settlements in the highland forest areas bordering Laos.

Figure 1 shows a map of the study area. The semi-transparent areas bordering the South China Sea represent the communes that are included in the LUCCi project, but will not be considered in the modelling part of this MSc thesis research. The reason for this is that these areas have a rather different environmental and socio-economic setting, which is characterised by lowlands, fishery livelihoods and a large amount of urban settlements. As this research aims to study land change patterns with a special focus on deforestation, it was decided that these communes in the East of the study area should be excluded[1]. The result is a rather homogenous socio-economic area, which enables further analysis and modelling of land change with a focus on deforestation.

Hereafter, for clarity reasons the study area *without* the excluded communes will be referred to as focus area. As will be elaborated in Chapter 4 and Chapter 5, the spatial analyses were executed on the complete VGTB area, whereas the agent-based model (ABM) simulates changes in the focus area only.

---

[1] The criteria that have been used to select or exclude communes are as follows:
- Include all communes that are characterized as highland areas;
- Exclude all communes East of the main National Highway;
- After making a buffer of 5km on both sides (East and West) of the National Highway, exclude all remaining communes that intersect this buffer.

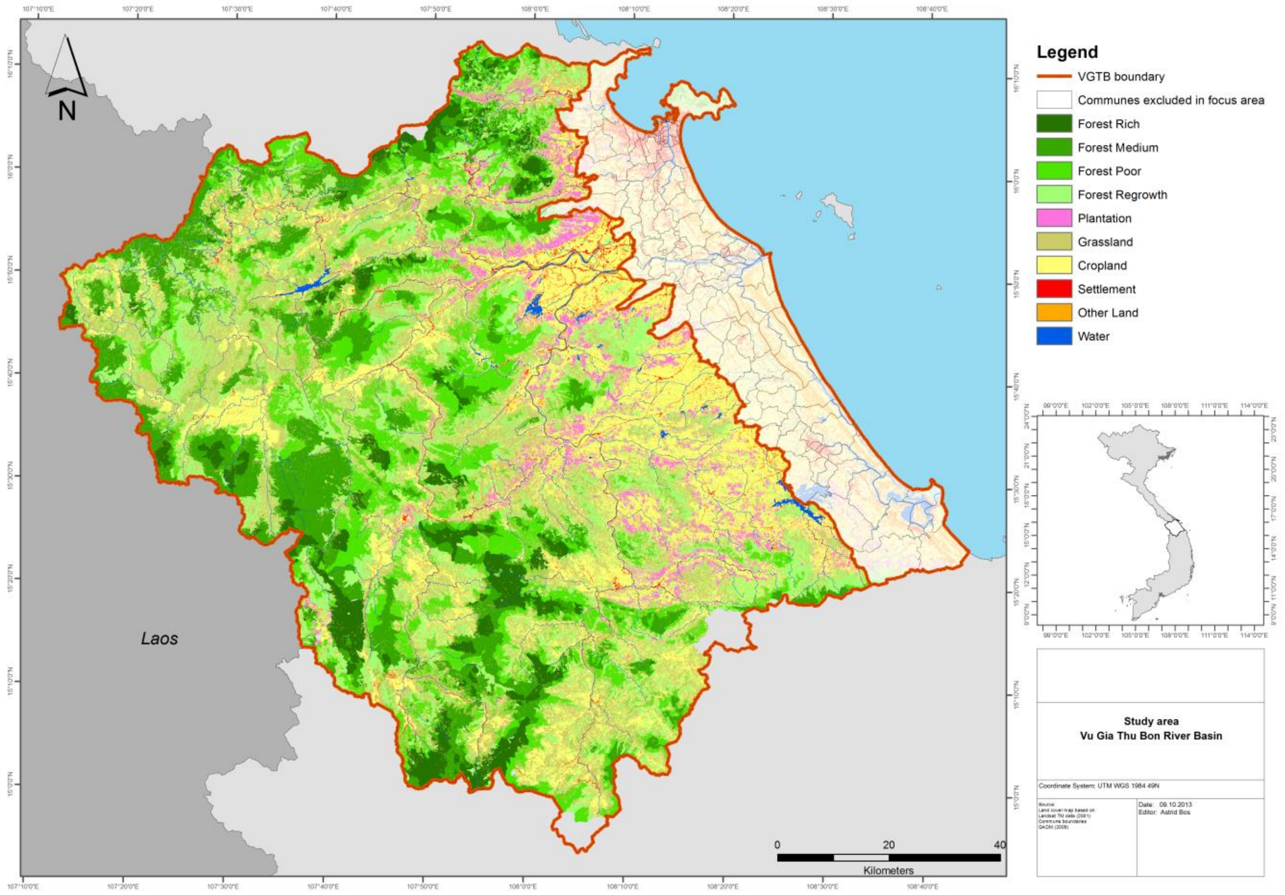


Figure 1 Vu Gia- Thu Bon study area and marked focus area

In Vietnam, the administrative division consists of –in hierarchical order- the national government, provinces, districts and a fourth level which includes communes (Figure 2). In practice, communes consist of several villages but villages are not registered as administrative zones. The boundaries of the VGTB area and of the focus area of this research are based on the commune boundaries. The VGTB area contains communes of four different provinces, that is, Da Nang, Quang Nam, Kon Tum and Quang Ngai. By far, the province of Quang Nam covers most of the area. An exact overview of the provinces, districts and communes included in the focus area can be found in Appendix A.

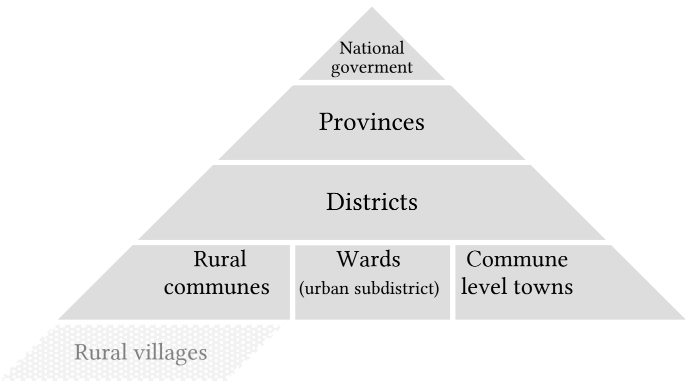


Figure 2 Vietnam's administrative divisions

1.1.3 Land change and forest transition processes in Vietnam

In pre-colonial eras, Vietnam was almost completely covered by forest. In 2005, almost 40% of the area of Vietnam was forest land (FAO, 2009). Vietnam has an annual net forest growth increase of 2%, which makes it one of the few countries in South East Asia in which forests are actually expanding. However, these numbers conceal the ever continuing trends of deforestation and forest degradation.

The Vietnam War played a considerable role in past deforestation. During this war between communist North Vietnam and government-led South Vietnam supported by the United States, the US sprayed approximately 72 million litres of herbicides (e.g. Agent Orange) on more than 1.5 million hectares of land (about 10% of South Vietnam) (Pesticide Action Network North America, 1998). Many forests were destroyed and few, if any, recovered fully after the war because environmental protection and restoration was not put high on the agenda (Adley & Grant, 2006).

Driven by the need to reconstruct the country and to feed the fast growing population, deforestation continued after the war. The new socialist *Doi Moi* (renovation) strategy of the government focused on liberalisation, decentralisation and individualisation of rights, including changing tenure systems (Lambin & Meyfroidt, 2010). Through forest land allocation, individual households were allocated forest land for forestry and agroforestry purposes in order to put bare and unused lands back into use. Together with the liberalisation of the markets this boosted the wood processing industry and led to a large increase of small and large scale forest plantations. Since the late 1990s there is a net forest increase in Vietnam (Minh & Warfvinge, 2002). At the same time, the policies, liberalisation, paddy land allocation and market integration increased the profitability of cultivation in the lowlands. "The rapid deforestation in mountainous regions combined with a high population density led to a reduction in fallows, soil erosion on hillsides and a shortage of suitable land for shifting cultivation" (Lambin & Meyfroidt, 2010, p. 114).

1.1.4 REDD and REDD+ scenarios

To understand what would happen with the land use and corresponding land cover with and without certain (policy) interventions, this research incorporates several scenarios for the future. These scenarios are formulated based on different implementation measures using REDD strategies. REDD (Reducing Emissions from Deforestation and Forest Degradation) is a multilateral initiative proposed under the United Nations Framework Convention on Climate Change (UNFCCC) in which developing countries are provided financial rewards for effectively avoiding deforestation and forest degradation.

More recently, REDD+ reflects the broadening of the approach, by focusing on the role of conservation, sustainable management of forests and enhancement of forest carbon stocks (Parker, Mitchell, Trivedi, Mardas & Sosis, 2009). The implementation strategies include a whole range of possible methods, including the provision of (financial) incentives, capacity building and technology transfer.

In practice, these measures are adjusted to the local circumstances. That is, the measures taken to counter deforestation and forest degradation are linked to the causes of deforestation in that particular area. The underlying thought is that measures can only be effective if they provide an alternative for the actions linked to the causes of deforestation and forest degradation. These actions and thus the corresponding measures may differ between countries, provinces, districts or even villages. Therefore, it may not be realistic to model REDD measures at river basin level. Hence, the model simulates only the projected *effects* of REDD scenarios, rather than the measures themselves. For example, the model simulates what would happen if authorities succeed in a fully protection of conservation areas, prohibiting any form of deforestation in those areas. The research does not consider the means needed for this measure. The scenarios considered in the model will be explained in more detail in Section 5.2.4.

1.2 Scope

1.2.1 Research objectives

As explained in the previous section, this research forms part of the LUCCi project. The main aim of this study is already illustrated in the title of this thesis, that is,

understanding drivers and processes involved in deforestation and modelling of forest change processes in Vietnam.

The research contributes to the LUCCi project by:

- Performing a **spatial analysis** of **land cover change patterns** in the study area, using land cover maps which are based on Landsat satellite data from 2001, 2005 and 2010;
- **Modelling of land change scenarios** for the next decade through Agent Based Modelling, in order to map future land use patterns and areas under particular deforestation risk.

1.2.2 Research questions

The central question for this research is:

Which drivers and processes are involved in forest change in the Vu Gia Thu Bon region in the past decade and how can Agent-Based Modelling provide insights in future trends of forest change for the next decade?

The central question is divided into sub questions, which are linked to the two different objectives and corresponding phases of the study.

1. For the past decade, which land use and land cover change patterns can be distinguished in the study area?
2. How do the land cover and land use change patterns in the study area relate to forest change in particular?
3. Which policy interventions, REDD strategies and corresponding land use change scenarios should be tested in the model?
4. Based on these scenarios, which areas are most likely to be deforested and are thus under particular deforestation risk?

1.2.3 Research limitations

The previous section described the aims of this study. Because these are still quite broad, it is also important to define what the research is *not* about, as the given timeframe naturally entailed limitations.

When it comes to modelling, there is a trade-off between analysis at broad scales and fine scales (Agarwal, Green, Grove, Evans & Schweik, 2002; Mertens & Lambin, 1997). The former is characterised by high levels of aggregation of data which may, unjustly, ignore the variability of geographic situations and diluting causal relationships. To work with fine scales is rather impractical, which makes scaling up to larger regions often quite difficult.

The strength of Agent-Based Modelling, as will be explained in more detail in the Chapter 3 on Methodology, is that it gives the user the opportunity to test the effects of different types of behaviour. Therefore, the emphasis will be on different options that agents in the area have and their effects on the land use change, rather than on all kinds of ecological or climatological processes. This does not mean that the latter do not influence the land cover and land use in the area, but they are of limited focus in this research.

For a more detailed reflection on the research process itself and the corresponding results, please refer to Chapter 6.

1.3 Looking forward

The remainder of this thesis is structured as follows. The next chapter deals with the Theoretical Framework, which is based around the theory of complex adaptive systems. In Chapter 3 the Methodological Framework is introduced, which covers the conceptual framework, an introduction to the methods and techniques applied and a description of the corresponding data and software. Chapter 4 describes the first method dealing with Spatial Analysis from the methodology in detail to the results. In Chapter 5 the agent-based model called SoDRA LUCCi, which was developed for this research, is described. The model has been tested based on the known deforestation rates and patterns of the 2001-2010 era. The final version of the model projects expected future deforestation risk areas for the 2010-2020 era as well as projected effects of several (REDD) interventions. Chapter 6 presents a critical reflection upon the research process and its results, followed by some general conclusions (Chapter 7) and recommendations (Chapter 8) for LUCCi project members, the Vietnamese authorities as well as for future research.

2. Theoretical Framework

This research is funded on theories and concepts from multiple sciences, ranging from ecology, environmental sciences and social sciences to computer sciences. This chapter is divided into sections that represent different theories, concepts and frameworks. It starts very general and moves to applied theories and concepts.

2.1 Complex adaptive systems

Dynamic system theory, complex theory and complex adaptive systems theory are three closely related notions that have their foundations in neuroscience and computer sciences (artificial intelligence). Furthermore, they are used in both environmental (ecology and biology) and social sciences.

Systems can be considered complex when they are dynamic and involve a large amount of variables (Gros, 2013). In addition, they can be considered adaptive when there are many components involved that adapt or learn as they interact (Holland, 2006). According to Holland (2006) complex adaptive systems (CAS) can be characterised by four distinctive features:

- Parallelism; numerous agents can act simultaneously
- Conditional action; agents' actions depend on signals they receive (IF/THEN)
- Modularity; agents' rules can be defined through "subroutines", executing sequences of rules
- Adaptation and evolution; agents and their behaviour can change over time.

In addition, important concepts in CAS are self-organisation and emergent properties (Ligtenberg, 2006). Self-organisation is often prominent in biological systems, where adaptive behaviour emerges from interactions among autonomously operating agents (Munthali, 2012). However, using concepts from complexity science in ecology science is a development that is still relatively new. In the last two decades, concepts from complexity science have been adopted in ecology science, together with the emerging acceptance that ecosystems can be considered complex adaptive systems themselves. These paradigm shifts in ecology sciences have affected the way renewable resources are managed. That is, the interactions between ecological and social elements have gained a central role in thinking about natural resource management. The heterogeneity and interdependent dynamics of these elements are increasingly being acknowledged (Le Page et al., 2013).

2.2 Modelling socio-ecological systems

In general, a model can be defined as a "formal, theoretical and/or physical system intended to bear specified similarities with a given natural [i.e. real world] system" (Gross & Strand, 2000, p. 28). As Le, Park, Vlek, and Cremers (2008, p. 136) acknowledge "(…) modelling of LUCC involves the complexity of both its human drivers and natural constraints". Therefore, modelling socio-ecological systems, which form a crucial role in LUCC, requires a thorough understanding of the individual elements of the system and their interrelationships.

According to Agarwal et al. (2002, p. 28) one central and three corresponding sub questions should be considered when developing a model that should represent a socio-ecological system:

How did the social-ecological system develop into its current state, and how might it change in the future?

- How have ecological processes influenced the social patterns and processes that have emerged?
- How have social patterns and processes influenced the use and management of resources?
- How are these interactions changing, and what implications do these changes bring to the state of the social-ecological system?

These questions already underline the bidirectional relationships between social entities and their ecological environment. As Munthali (2012) notes, there are different options to model these socio-ecological systems. First, cellular automata (CA) are models in which each cell exists in one of a finite set of states. Transition rules based on the cell's spatio-temporal neighbourhood define the future state of the cell. Another option are Markov models, in which "cell states depend probabilistically on temporally lagged cell state values" (Munthali, 2012, p. 144). Combinations of these model types are also possible.

Both modelling options acknowledge the central role of (human) decision-making, but fail to represent these decision processes in their simulations (Munthali, 2012; Parker, Manson, Janssen, Hoffmann & Deadman, 2003). "When the focus is on human actions, agents become the crucial components in the model. While cellular models are focused on landscapes and transitions, agent-based models (ABMs) primarily focus on humans and their actions" (Munthali, 2012, p. 144). ABM therefore acknowledges the humans as central decision-making actors. Still, environmental processes that operate to a greater or lesser extent independently from human decision-making should not be ignored when designing an ABM.

2.3 Decision-making & the Sustainable Livelihoods Approach

The previous section concluded by putting the humans central in modelling socio-ecological systems. Which human entity to use is also an important point for consideration. Opting for individuals, households, or some sort of organisational entity has consequences for the model design.

The sustainable livelihoods approach puts individual households at the centre of decision-making processes. "A livelihood comprises the capabilities, assets (including both material and social resources) and activities required for a means of living. A livelihood is sustainable when it can cope with and recover from stresses and shocks, maintain or enhance its capabilities and assets, while not undermining the natural resource base" (Chambers & Conway, 1991, p. 6).

A household bases its decisions and actions on their abilities (defined by their assests) and strategies (Figure 3). The assets consist of different forms of capital: human, natural, physical, financial and social capital (Carloni & Crowley, 2005; Chambers & Conway, 1991; Scoones, 1998). The vulnerability context and the natural capital of a livelihood are directly linked to the contextual socio-ecological systems. For example, degraded soils caused by unsustainable shifting cultivation practises in which the fallow period is too short, limit the natural capital of the livelihood. This might lead to a shift in or adjustment of the livelihood strategies (e.g. a shift to extensification of agricultural land) and thus a shift in livelihood outcomes. Regarding rural households, the livelihood outcomes are thus directly linked to the way the land is used which may impose LUCC.

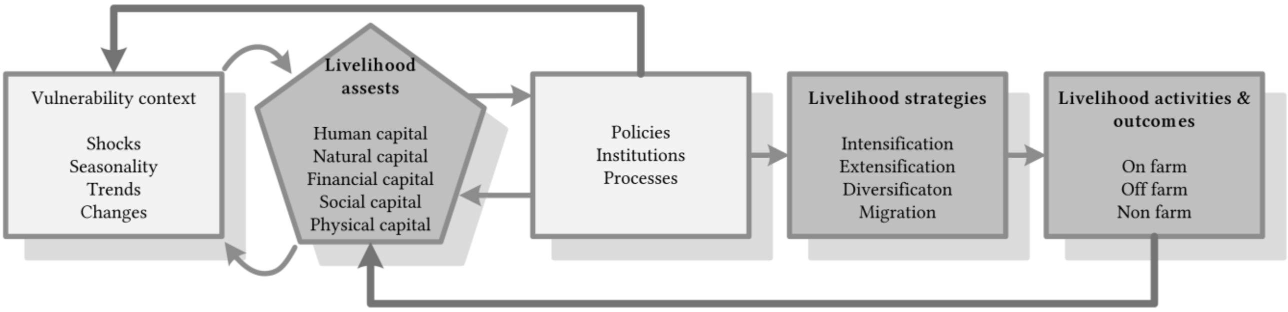


Figure 3 The sustainable livelihoods approach, adapted from Carloni and Crowley (2005)

2.4 Modelling Complex Adaptive Systems using Agent-Based Models

2.4.1 Introducing Agent-Based Models

The next consideration regards how the agent's behaviour within its complex adaptive system should be modelled. In this research, an agent-based modelling (ABM) approach was chosen. The concept behind ABMs is based on computer sciences, and is founded on the idea that human- or animal-like agents can be simulated at the micro-scale in a computer simulation in order to study how their aggregation leads to complex macro-behavior and phenomena (Berger, 2001; Munthali, 2012). An ABM is a model in which "a system's dynamic behaviour is represented through rules governing the actions of a number of autonomous agents" (Maguire, Batty & Goodchild, 2005, p. 8). It "(…) offers a way of incorporating the influence of human decision-making on land use in a mechanistic, formal, and spatially explicit way, taking into account social interaction, adaptation, and decision-making at different levels" (Matthews, Gilbert, Roach, Polhill & Gotts, 2007, p. 1447). In previous studies, ABMs have demonstrated that some qualitative features of complex (natural) systems can be reproduced and explained using relatively simple ABMs (Gross & Strand, 2000). In this sense, complexity may emerge from simple rule sets.

2.4.2 The distinctive character of ABMs

ABMs differ from other models in several ways. First, ABMs are able to capture emergent phenomena. Emergent phenomena are "system dynamics that arise from how the system's individual components interact with and respond to each other and their environment" (Railsback & Grimm, 2012, p. 10). These system dynamics may be unexpected, as they emerge from underlying processes. Second, ABMs are flexible and provide a natural, intuitive description of the system (Munthali, 2012). Third, besides modelling individual entities' decision-making, ABMs are also capable of incorporating social processes and non-monetary elements that influence decision-making and capable of dynamically linking social and environmental processes (Matthews et al., 2007). Consequently, this makes ABMs often more complex compared to more conventional modelling approaches. "ABMs are less simplified (…), they represent a system's individual components and their behaviours. Instead of describing a system only with variables representing the state of the whole system, we model its individual agents" (Railsback & Grimm, 2012, p. 10). Finally, ABMs are characterised by the agent's adaptive behaviour. Agents in an ABM have the ability to "adjust their behaviour to the current states of themselves, of other agents, and of their environment" (Railsback & Grimm, 2012, p. 10). These distinctive abilities make ABMs valid choice for many types of environmental modelling.

However, the use of ABMs is still not being used on a wide scale, and its usefulness for scientific research is being criticised by some authors, leading to a lively debate in academic spheres. As Munthali (2012, p. 152) acknowledges, an often heard critique is that "[ABMs] cannot be sufficiently deductive to give confidence in the outcomes from the model parameters". Proper sensitivity analyses are therefore key to understand the system that is being modelled. Furthermore, at the micro-scale ABMs tend to be sensitive to small perturbations in model parameter values. These unwanted effects on the model can be avoided by limiting the focus of the simulation on emergent patterns and on general trends on a larger scale (Parker & Meretsky, 2004). Validation of ABMs is often found to be difficult. Furthermore, since actual decision-makers are often not involved in the process of designing and experimenting with the model, and researchers often lack proper understanding of the actual decision makers process, the value of ABMs as decision support tools has tend to be rather limited (Munthali, 2012).

Still, ABMs are valued more and more because they offer the opportunity to incorporate the influence of human decision-making on land use in a mechanistic, formal and spatially explicit way (Matthews et al., 2007; Munthali, 2012). In that sense, if not for predicting the future, they do provide us a way to understand the world around us a little bit better.

2.4.3 ABMs as science

ABMs are used in scientific research to try and understand non-linear systems (Munthali, 2012). They represent a "third way" of doing science (Axelrod, 1997; Matthews et al., 2007). The first way of doing research refers to the inductive approach, in which patterns are discovered in empirical data. The deductive way of doing research entails the process of proposing hypothesis, making observations of the real world and thus proving or falsifying predictions derived from the hypothesis. Hence, ABM represents a third way of science. "ABM (...) is an amalgamation of these two approaches – like deduction, it starts with a set of explicit assumptions derived from perceptions of the way the world works, but uses these to generate simulated data that can be analysed inductively" (Matthews et al., 2007, p. 1457). Thus, empirical data for the inductive process does not originate from direct observation of the real world, but emerges from a defined set of rules. In turn, these set of rules are based on and validated through past real-world observations.

2.5 Looking forward

The theoretical framework presented in this chapter underpins the methods used for this research. The conceptual model functions as a bridge between the theories and the methods (see Section 3.1, and in particular Figure 4 on page 11). The next chapter describes the methodological framework of this research.

3. Methodological Framework

This research is executed using different methods and techniques, which can be divided into two main phases. In this chapter, first the conceptual model will be presented, followed by an introduction to the two main methods. The methods will be discussed in more detail in the upcoming two chapters.

3.1 Conceptual model

The conceptual model of this research is presented in Figure 4. It is partly based on the research of Wood and Porro (2002) who studied land changes, in particular deforestation, in Latin America. Although through large scale exploitation also the land use of commercial companies and government agencies may have an influence on the land cover, in this research the focus is on individual agents, and thus local people's livelihoods.

The assumption is that people's behaviour, i.e. resource allocation, can be influenced by both socio-economic and bio-physical drivers. These drivers may operate at different levels. An example of a macro socio-economic driver could be the global market price of cash crops. Soil quality could be an example of a micro bio-physical driver. While institutional arrangements like land rights or bio-physical issues like soil quality or climate characteristics define the opportunities and constraints for land use, it is hypothesised that it is the decisions of the household, and thus the livelihoods' resource allocation, that determines the final land use for that particular area. Land use and land use changes have an effect on the land cover, triggering land cover change.

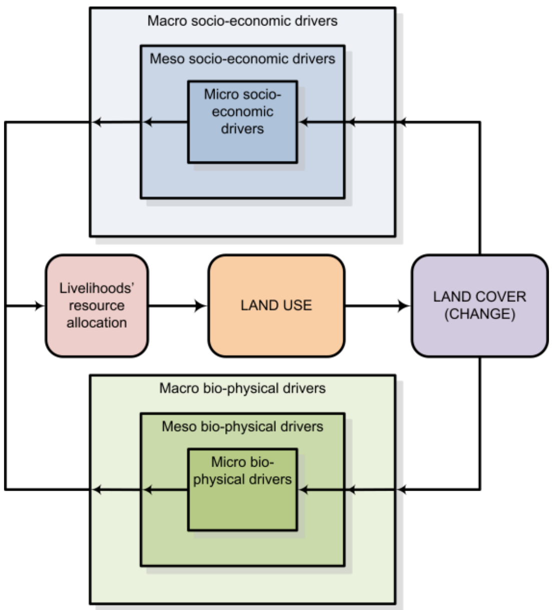


Figure 4 Conceptual model

3.2 Methods & techniques

This study can be roughly divided into two -more or less chronological- phases. These phases are characterised by their own methods and techniques and are closely linked to the objectives and sub questions explained earlier in Chapter 1.

Besides a chronological order, these methods also follow a logical order, as the results from prior methods will form the input for the succeeding methods with the aim to generate a sound model for the study area.

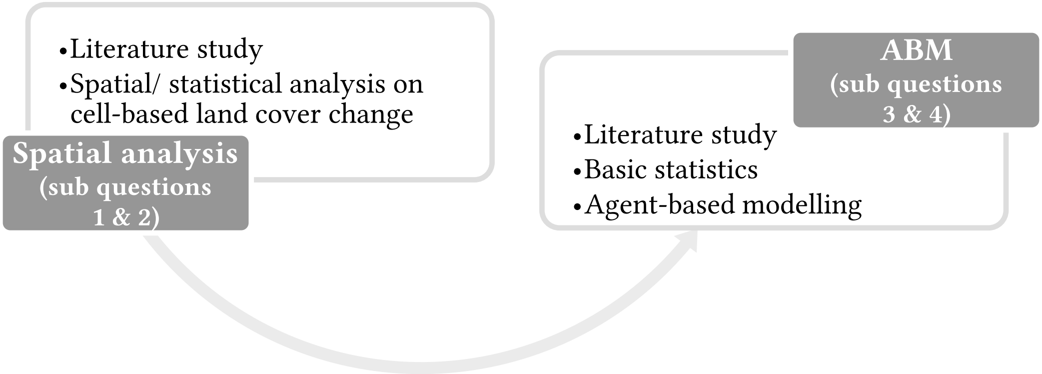


Figure 5 Research phases and corresponding methods

3.2.1 Literature study

The first phase concerned a literature study, in which the background of the research setting was studied, as well as the methods and techniques needed for the remaining phases of the research. In order to check whether all parts of the literature needed for this research were covered, a literature map (Meth & Williams, 2006) was created (see Figure 6). As the literature study formed the basis for the remainder of the research, it covered not only background topics (upper part of the circle), but also methodology related literature (lower part of the circle).

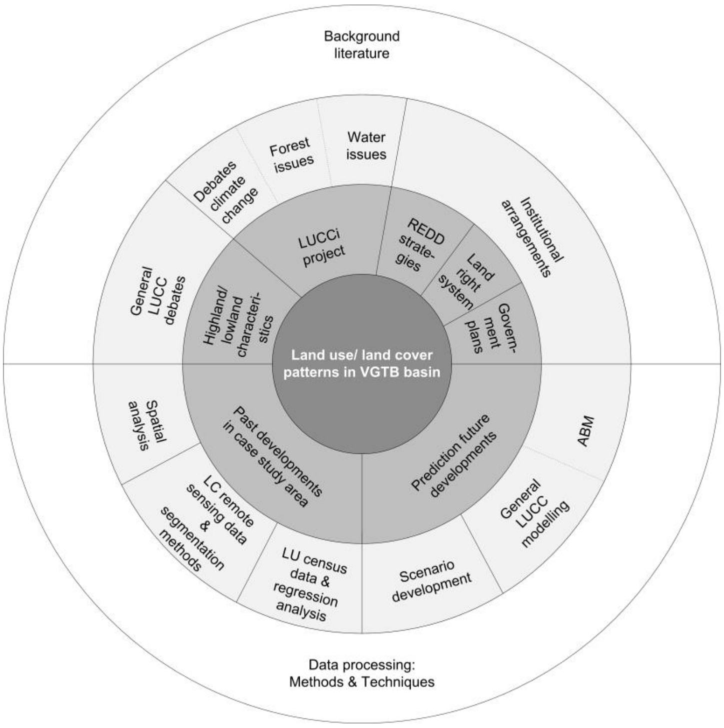


Figure 6 Literature map

3.2.2 Spatial analysis

The land cover information for this study is based on land cover data derived from 30m resolution Landsat TM satellite imagery. Data was available from 2001, 2005 and 2010, and land cover data was classified in six classes (i.e. forest, cropland, grassland, water, urban and other land), allowing analysis of land cover change for the eras 2001-2005 and 2005-2010. Additional thematic and topographic data was made available through the LUCCi project, allowing a context analysis on the deforested cells that focused on elevation (based on a Digital Elevation Model (DEM)), slope, and several proximity factors. The acquired information was added as attribute information to the deforested cells.

For a more detailed report on the spatial analysis process, see Section 4.2.1 and 4.2.2.

3.2.3 Statistical analysis

The attribute information attained through the spatial analysis formed the input for statistical analysis to reveal correlations between the potential topographic and thematic factors and deforestation. The statistical analysis compared the focus group, that is, all deforested cells, with a control group, existing of randomly selected unchanged forest cells for the eras 2001-2005 and 2005-2010.

For a more detailed report on the statistical analysis process, see Section 4.2.3 and 4.2.4.

3.2.4 Basic statistics

Ideally, the agent behaviour in an ABM is founded on detailed information on socio-economic factors like land use, livelihood assets, livelihood strategies and tenure systems. For this research, there were few socio-economic datasets available with all different advantages and disadvantages. An overview of these datasets and their pros and cons can be found in Appendix B. It was found that none of the (available) datasets was of sufficient quality (adequate spatial level and level of detail) to distinguish agent types and thus define differences in agent behaviour. It had been decided that for this first version of an ABM for the LUCCi project, only one agent type was defined, that is, a rural household. Its behaviour in the model was based on calibrations using the known land cover changes (2001-2010) (see also Section 5.4).

An important parameter in the model is the number of agents, since they define to a large extent how much forest area is being deforested. The number of rural households was calculated and derived from different datasets. The exact number of rural households per lowest administrative level (commune) can be found in Appendix A.

For a more detailed report on the basic statistics process for agent definition, see Section 5.2.

3.2.5 Agent-based modelling

The final part of the research aimed at converting the insights acquired in the previous phases into agents and action rules for modelling deforestation and the effect of certain interventions for the coming decade. The value of modelling is often manifold. It can give the researcher or end user new insights or guide further analysis; real-life experiments in land-use systems are difficult, if not impossible, and thus no option; models can play a role in communication between researchers; and models enable the user to explore possible future developments in the land-use system (using "what-if" scenarios) (Verburg, Kok, Pontius Jr. & Veldkamp, 2006). However, this is often more difficult than it appears. As Agarwal et al. (2002, p. 26) acknowledge "the utility of a land-use change model can be measured primarily by its ability to demonstrate emergent patterns in land-use change processes and, secondarily, as a predictive tool".

In this research, an agent-based modelling approach was used. An ABM exists of several key elements:

- Agent – The agent can be an individual or an (aggregated) set of individuals. In fact, an agent can represent any form and any level of organisation (Verburg, 2006).
- It's behaviour – To model an agent's behaviour requires understanding of agents' actions, and information about when the environment and other agents change based on these actions (Munthali, 2012). Decision making structures of agents can be divided into two broad categories: heuristic and optimising. Heuristic agents have "neither the information to compare all feasible alternatives nor the computational power to select the optimum" (Munthali, 2012, p. 149). Optimising decision making requires the ability to process large amounts of information about all feasible alternatives and always select the best one (Munthali, 2012). As the same author acknowledges, agents in LUCC models often follow a heuristic decision making process, which can be framed in a heuristic decision making tree (Figure 7).

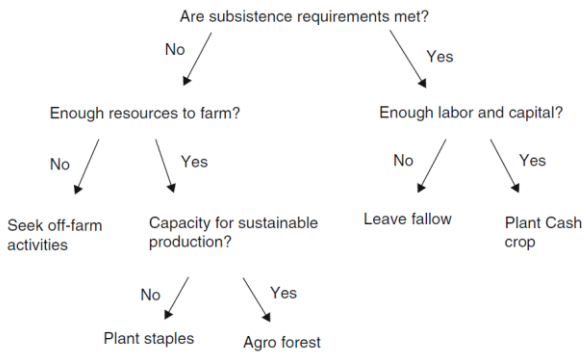


Figure 7 Example of a heuristic decision making tree, source: Munthali (2012)

  Agents' behaviour can be both influenced by internal (related to the agent itself, e.g. household) and external factors (institutions and social networks) (Valbuena et al., 2010).
- Environment – The environment consists of different elements that can involve themselves, or change as a response to agent behaviour. In geographic, real life simulations, the environment is represented by a raster grid (Deadman, Robinson, Moran & Brondizio, 2004).
- Interaction between agents and environment - Not only the environment can be influenced by agents' behaviour, but agents' behaviour in itself may also be based on the state of environment. Changes in the system can therefore be considered bidirectional. Changes in the environment may enforce changes in human behaviour and vice versa. This is why ABMs can be regarded as *across-level* models (Railsback & Grimm, 2012).

No models are completely accurate and correct, but they may be of valuable use in understanding the processes at stake. Here, the aim was to get insights in the areas that are under particular deforestation risk for the next decade. Moreover, the model incorporates certain REDD measures and visualises its effects with regards to the projected "business-as-usual" deforestation.

For a more detailed report on the ABM process, see Section 5.2.

3.3 Looking forward

In this chapter, a first overview of the methodology has been given. The next two chapters focus on the two main phases of the research, that is, the spatial/statistical analysis and the agent-based modelling.

4. Spatial analysis

4.1 Introduction

In order to say something about land cover trends and possible threats in the future, we must learn from the lessons of the past. Therefore, it is important to investigate the land cover change patterns of the past decade. Four main questions are central while doing these analyses, that is, *which* changes took place, in what *quantities*, *where* did they take place, and, probably the most difficult one to answer, *why* did these changes occur?

This chapter describes the process of analysing the land cover change patterns of the past decade (2001-2010). This is done in a more or less chronological manner, starting with the land cover change data and the thematic layers which are used as input, via several processing stages to the final statistical analysis in SPSS. The final sections of this chapter reveal the main results and a critical discussion and conclusion.

4.2 Methodology & process

4.2.1 Pre-processing in ArcGIS

The starting point of this analysis is three land cover maps, of 2001, 2005 and 2010 (Figure 8) and two corresponding forest change maps of 2001-2005 and 2005-2010 (Figure 10). As part of previous stages of the LUCCi project, these maps were generated based on Landsat TM satellite data obtained in the corresponding years. Figure 9 shows the land cover maps of 2001, 2005 and 2010 with detailed forest classes. Here, the Landsat TM data was combined with the forest classes of the FIPI (Forest Inventory and Planning) forest maps of the Vietnamese government.

For the analysis, the focus is only on forest change or, more specifically, on change from forest land to another land cover class. Reforestation, i.e. from any land cover class to forest land, is not included in this research. The forest change cells were extracted from the different change raster files using the ArcGIS tool “reclassify”. All other changed cells and non-changed cells were given the value 0.

The goal of the analysis was to examine whether some factors have a correlation with forest change. Based on literature research, previous experience and data availability, the following factors were chosen to be included in the analysis:

	Factor	Type	Original file raster/ vector	Source:
1	Digital Elevation Model (DEM)	Topographic	Raster	DEM LUCCi
2	Slope	Topographic	Raster	Calculated from DEM
3	Distance to All roads	Thematic	Vector	LUCCi vector file
4	Distance to Paved roads	Thematic	Vector	Based on All Roads file
5	Distance to Small settlements	Thematic	Raster	Based on land cover data (Landsat TM)
6	Distance to Large settlements	Thematic	Vector	Based on land use data (Landsat TM)
7	Distance to Cropland	Thematic	Raster	Based on land cover data (Landsat TM)
8	Distance to Grassland	Thematic	Raster	Based on land cover data (Landsat TM)

Table 1 Factors included in spatial analysis

Factor description

1. Digital Elevation Model (DEM)

   The DEM data at 30m resolution was provided for this research and required no other pre-processing.

2. Slope

   The slope layer was based on the DEM data and was generated using ArcGIS' spatial analysis "slope" tool. As a result, a continuous raster was created showing the slope in degrees.

3. Distance to All roads

   Road data was provided by the Vietnamese government. Although it was hard to check the quality of the data as no proper metadata was provided, the data showed a sufficient overlap with the roads (classified as urban area) recognised using the Landsat TM data.

   The road data were already clipped to the VGTB area. It is expected that the distance to all roads of forest change cells near the borders of the VGTB region may show an overestimation compared to reality, as roads from neighbouring regions were not included in the analysis. However, the general trend is likely to be not affected, as the north-western, western and south-western border areas are remote areas in which only few roads exist.

   The data, consisting of polylines, were rasterised to allow an overlay with the forest change raster data.

4. Distance to Paved roads

   For this layer the same data was used as the "Distance to All roads" layer. The data contains attribute information showing the type of road, although many lines did not have this information. Therefore, it was decided to include both the complete data (the "Distance to All roads" layer) as well as a separate layer showing the lines which were certain to be paved. Lines with the attribute values like "unpaved", "dirt" or "track" and lines without any value were excluded from this layer.

   Again, the polylines were rasterised to allow an overlay with the forest change raster data later on.

5. Distance to Small settlements

   In Vietnam, the lowest administrative level is the commune level (McElwee, 2008). A commune often consists of one small rural town and some surrounding villages. However, as these villages are officially not recognised as administrative unit (see Section 1.1.2), it is hard to find data on the exact location of these villages. Therefore, it was decided to take all cells from the land cover maps classified as settlement and assume that these include at least all small villages. However, these cells also include all other non-vegetative and non-water cells, such as roads. As in Vietnam rural people also may reside outside the villages and along the roads, this was not considered to be having a considerable influence on the results.

6. Distance to Large settlements

   Like with the distinction between all and paved roads, it was considered to be important to distinguish between small and large settlements. It was assumed that larger settlements offer more services like markets for both forest and crop products. However, data on the availability of markets was lacking, making it difficult to generate a "Distance to Markets" layer. To work around this absence of data, it was decided to make a distinction between small and large settlements.

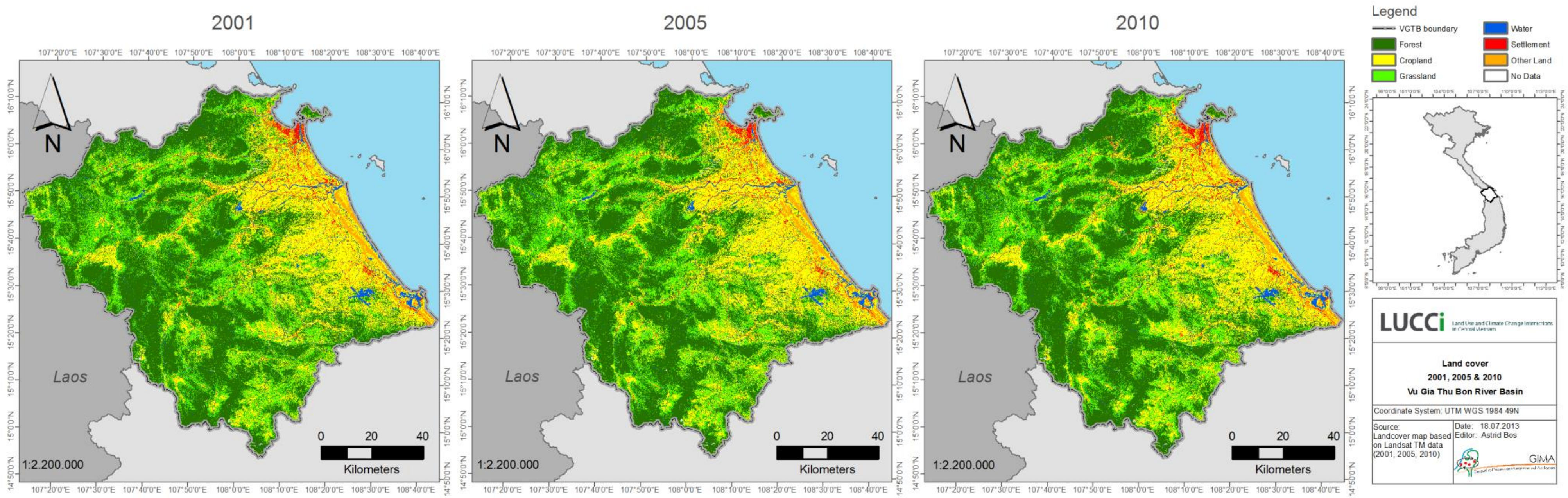


Figure 8 Land cover 2001, 2005 & 2010

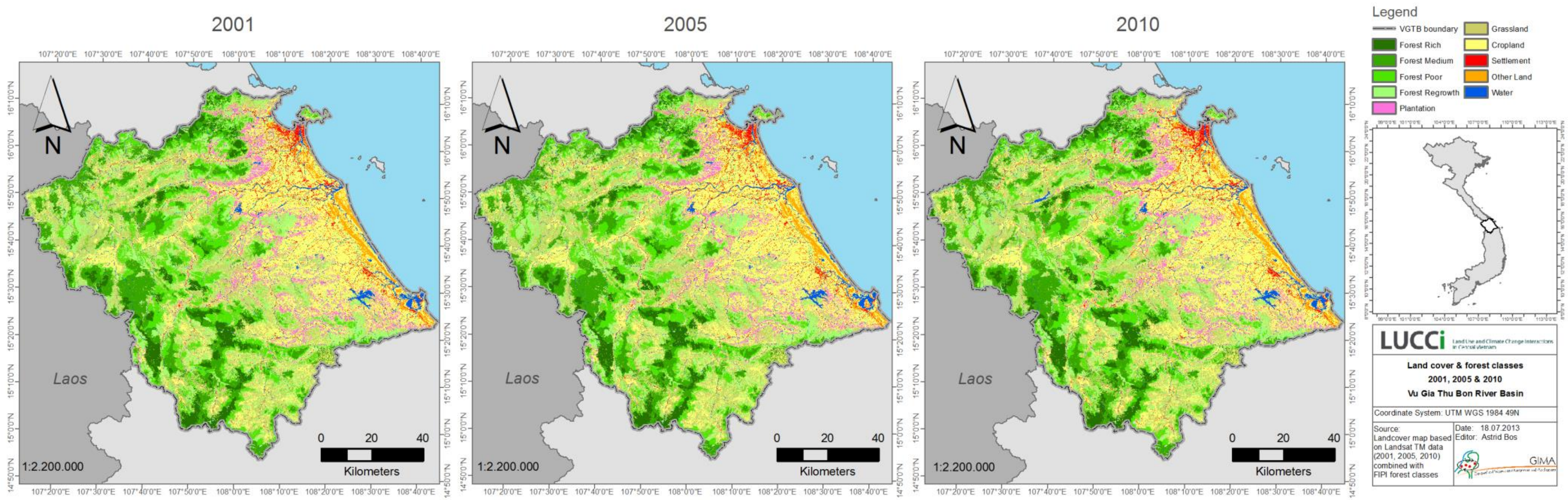


Figure 9 Land cover & forest classes 2001, 2005 & 2010

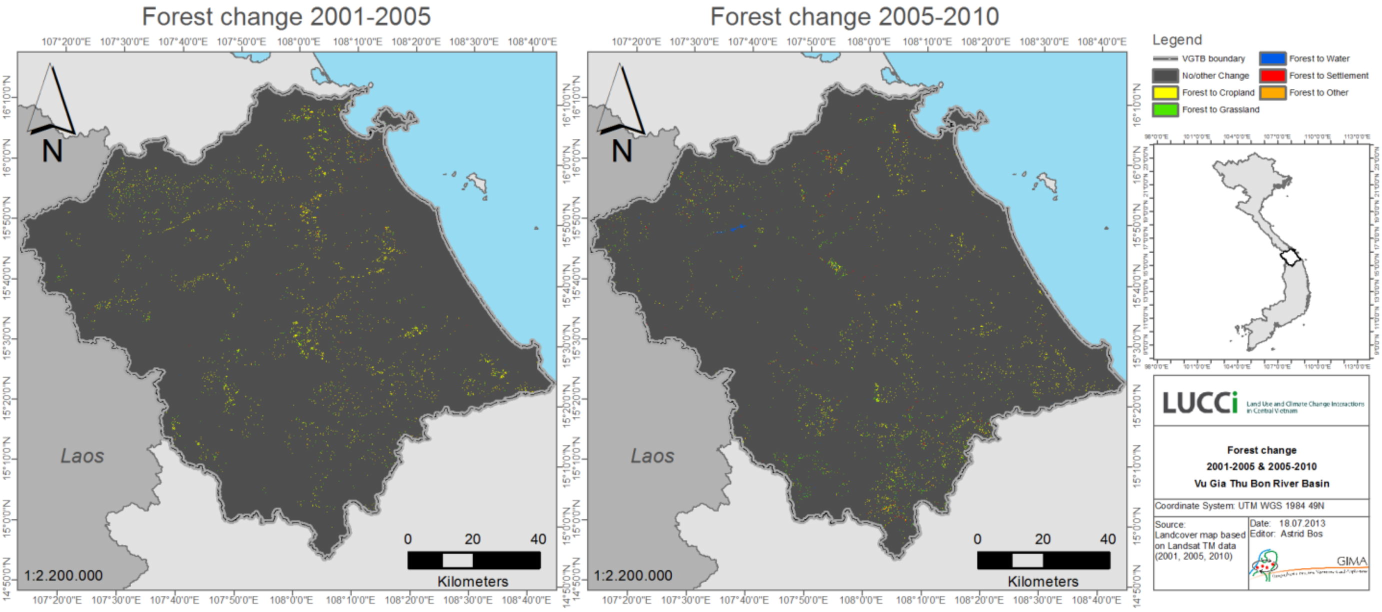


Figure 10 Forest changes 2001-2005 & 2005-2010

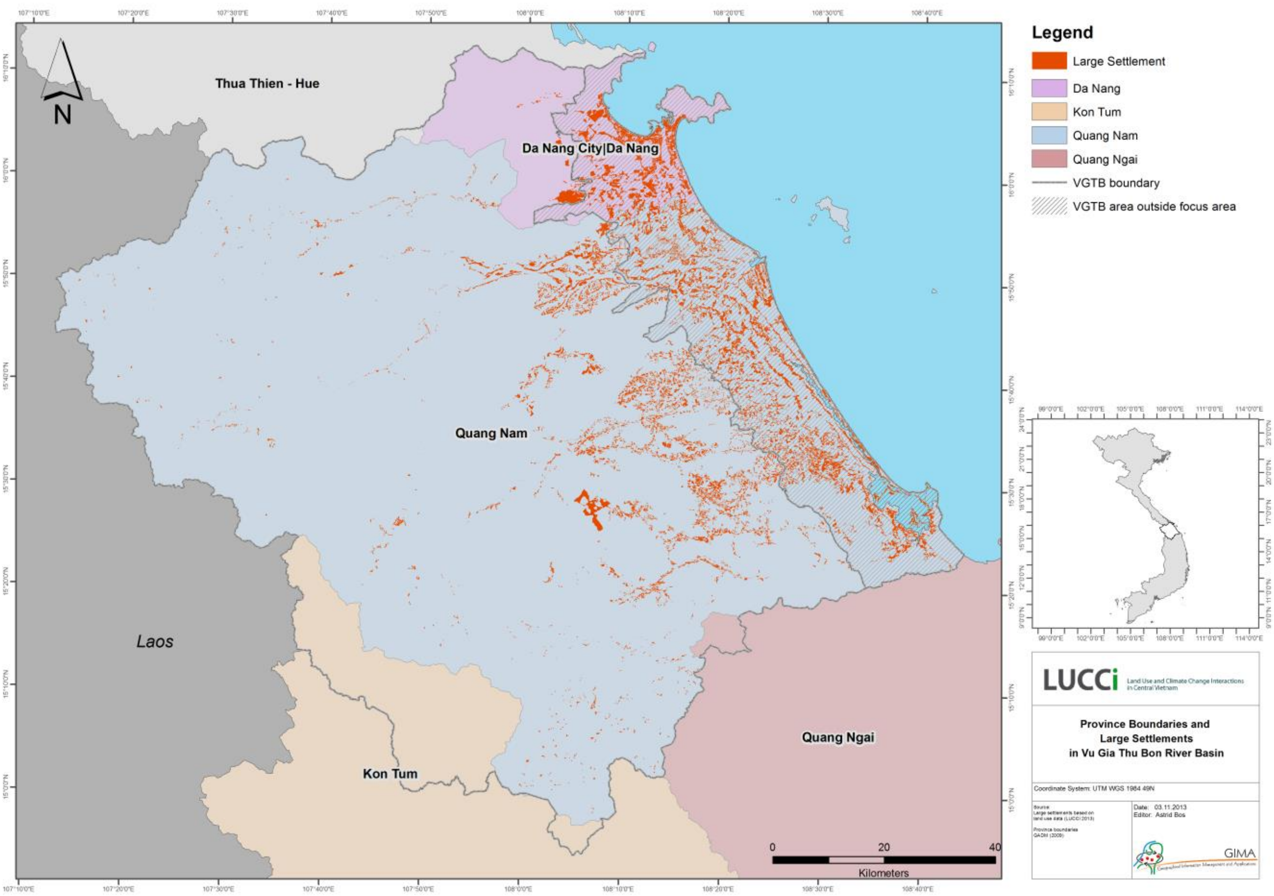


Figure 11 Province boundaries and large settlements in the Vu Gia- Thu Bon river basin

The data on large settlements was based on land use shapefiles generated by the Vietnamese government and edited by other project members of the LUCCi project. These shapefiles only contained data from the provinces of Da Nang and Quang Nam, so any large settlements in the small parts of the provinces of Kon Tum and Quang Ngai which are also part of the VGTB region may be overlooked. Again, this was not considered to be having a major influence as Kon Tum and Quang Ngai are only for a very small portion part of the VGTB region (see Figure 11). Furthermore, these areas are part of the remote rural areas where little or no large settlements are expected to be situated.

The land use shapes classified as "urban settlement" or "rural settlement" were extracted and rasterised to allow an overlay with the forest change raster data in a later stage of the process.

7. Distance to Cropland

   The major forest change in both eras was the change from forest to cropland, counting for over 70% of forest change in 2001-2005 and almost 60% of forest change in 2005-2010 (Schultz & Avitabile, 2012). As a hypothesis, it was thought that future forest change may appear close to existing cropland. For the 2001-2005 forest change era, cropland cells were extracted from the 2001 land cover raster file. Likewise, for the 2005-2010 forest change era, all cells classified as cropland were extracted from the 2005 land cover raster file.

8. Distance to Grassland

   Another substantial change is from forest to grassland. Therefore, the distance of forest change cells to existing grassland was also investigated in similar ways as distance to cropland.

To allow proper segmentation in eCognition (see Section 4.2.2), the raster values of the thematic layers were multiplied by 1000. Next, the forest change raster was overlaid with the different thematic layers using ArcGIS' raster calculator. As for most of the layers, the data come from different sources, and therefore unrealistic overlap of classes is possible. For example, a cell classified as "forest to cropland" (based on the land cover data) can be overlapped with "paved roads". All these possible overlaps are assigned unique values through the raster calculator and treated as a separate class within eCognition in the next step of the process.

Control group

In order to state that the correlations and regressions found in the analyses are unique for the forest change cells only and not, for example, coincidence or distinctive for both changed and unchanged forest cells, a control group was installed and incorporated in all further analyses.

For both change eras, the total amount of forest change cells was approximately 150,000 (Schultz & Avitabile, 2012), with a cell size of 30 square metres. The two control groups existed of non-changed forest cells for the eras 2001-2005 and 2005-2010. Per control group, 150,000 cells were randomly selected using ArcGIS' tool "create spatially balanced points" (Theobald et al., 2007). All unchanged forest cells were given the maximum inclusion probability value, i.e. 1, all other cells were given the value 0. The resulting “balanced points” were rasterised again and overlaid with the different thematic layers like the changed forest cells.

4.2.2 eCognition object oriented image analysis

The second step in the spatial and statistical analysis involved the object oriented image analysis of the changed and unchanged (i.e. control group) forest areas. This was done using eCognition Developer software.

General project settings and Rule Set for all layers

As its name already suggests, eCognition Developer for object oriented image analysis allows the user to analyse images (raster files) in an object oriented way.

Figure 12 eCognition Process Tree, example of DEM Forest Change 2005-2010

This process consists of three main steps (Figure 12). First, segments, i.e. objects, of spatially connected cells were generated. In eCognition, this is done via a general segmentation algorithm based on homogeneity definitions together with local and global optimisation techniques (Baatz & Schäpe, 2000). The parameters for the heterogeneity criterion were set to a minimum (Shape parameter = 0.1; Compactness parameter = 0.5) to allow small object generation. As the differences between the raster values of the different classes were at least 1000 in the previous step of the process (see Section 4.2.1), erroneous segmentation in which objects consist of more than one forest change class, was avoided (Figure 13).

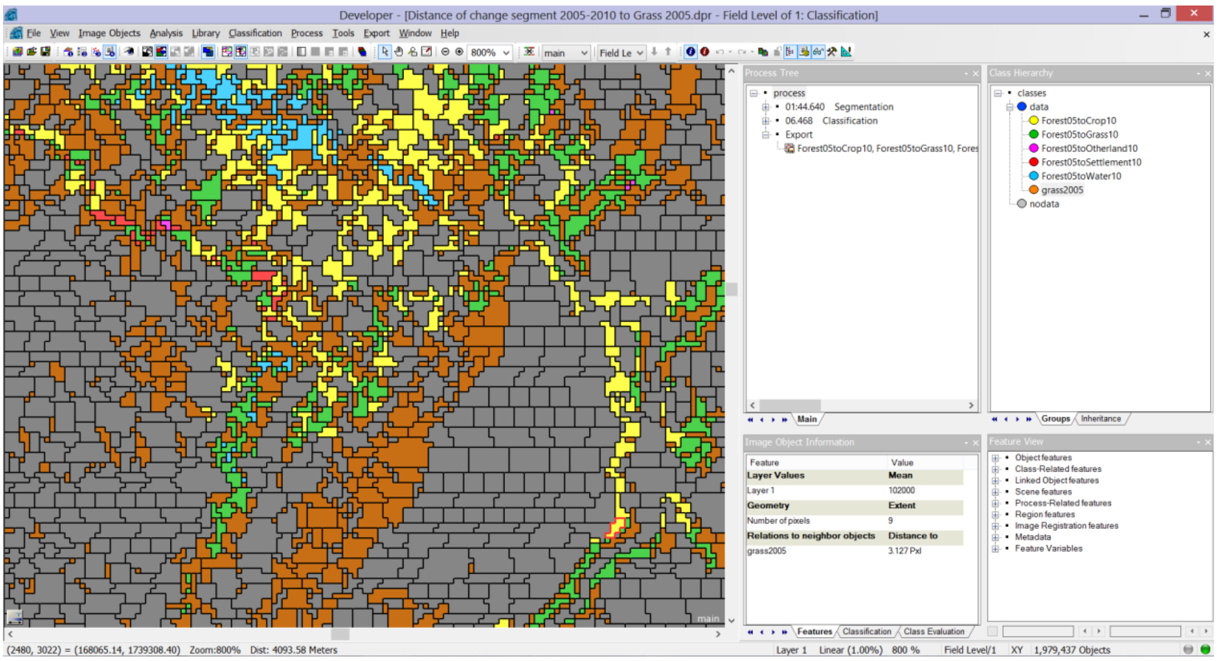


Figure 13 Screenshot of eCogntition project after segmentation and classification

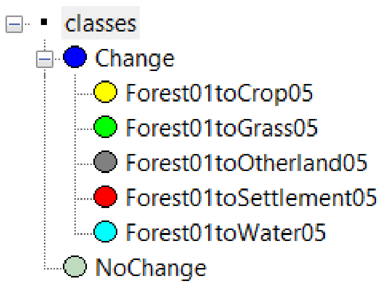


Figure 14 Classification of Forest Change in eCognition

Second, the original raster values formed the basis for a twofold classification (Figure 14). Note that here "NoChange" refers to both unchanged forest as well as non-forest land cover changes.

One of the biggest values of eCognition is its window on Image Object Information. The user can select information to be shown in this window ranging from basic object features such as the basic layer value to class-related features such as distance to certain types of neighbours. With the latter option, it allows the user to analyse the distance in pixels of forest change to, for example, segments classified as Paved Road.

Third, during the export phase segments were converted into polygons and an attribute table showing the original classification value and corresponding image object information (e.g. average DEM of that segment, minimum distance to All Roads, etc.). After exportation, the polygons can be adapted and analysed in other GIS software such as ArcGIS.

Specific project settings and Rule Set for topographic and thematic layers

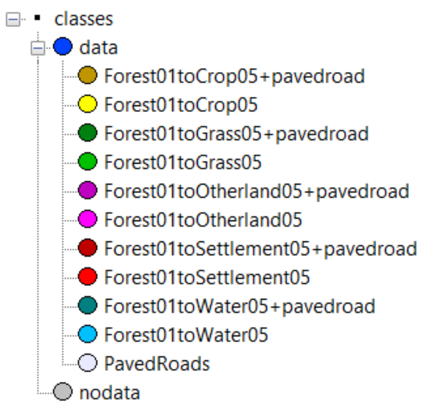


Figure 15 Extended classification allowing overlapping features

Feature	Value
Layer Values	**Mean**
Layer 1	102000
Geometry	**Extent**
Number of pixels	4
Relations to neighbor objects	**Distance to**
Forest01toCrop05+pavedroad	112.24 Pxl
Forest01toGrass05+pavedroad	112.43 Pxl
Forest01toOtherland05+pavedroad	123.19 Pxl
Forest01toSettlement05+pavedroad	754.56 Pxl
Forest01toWater05+pavedroad	1066.30 Pxl
PavedRoads	113.71 Pxl

Figure 16 Example of Image Object Information window

In general, the project settings and Rule Set (i.e. process tree) do not vary too much between the different layers. The main difference between the topographic and thematic layers is that for the thematic layers the thematic data is overlaid with the forest change layer or control group layer, resulting in a single layer project in eCognition, whereas the forest change layer (or control group layer) and the continuous topographic layer are kept as separate layers in eCognition.

As explained in Section 4.2.1, for some of the thematic layers overlapping classes are occurring. In these cases, the classification step in eCognition (Figure 14) was extended to include these overlapping classes (Figure 15). In the case of, for example, Paved Roads, the distances in pixels of forest change to all features that include Paved Roads are shown in the Image Object Information window and exported to the shapefile (Figure 16).

As mentioned above, the eCognition projects of the topographic layers consist of two separate layers. Here, only the values of the change segments (original forest change classes) were exported together with their corresponding value of the second layer, such as the DEM value.

All steps in the eCognition process were repeated for the sampled control group.

4.2.3 ArcGIS post-processing

Before the polygons generated in eCognition can be used in SPSS for statistical analysis, a few post-processing steps within ArcGIS were needed.

For the change segments of the thematic layers that contain segments with values that show an overlap of forest change and the thematic data (i.e. distance to all roads; paved roads; and large settlements) it needs to be investigated what the minimum distance is and thus which distance should be taken for further analysis. A new field (Dis_min_px) was added and the minimum distance was calculated using the following function in Python

$$\text{Dis\_min\_px}_i \; = min([!field1!, !field2!, !field3!, ! \ldots !])$$

For all attributes that show distance information, eCognition has calculated these distances in pixels. For all layers, the resolution was 30 meters. A new field was added to all thematic layers and the corresponding distances in meters was calculated using VB script

$$\text{Distance\_}m_i = Dis\_min\_px_i * 30$$

As a result of the eCognition projects, each information layer (topographic and thematic) was stored separately. To allow proper statistical analysis, these files were combined using ArcGIS Union tool, resulting in four main files:

- Forest change 2001-2005
- Unchanged cells 2001-2005 (control group)
- Forest change 2005-2010
- Unchanged cells 2005-2010 (control group)

SPSS treats every case (here: polygon) equally. These polygons need to have the same size, otherwise large polygons and small polygons will be treated as having the same weight. Therefore, using the Fishnet tool a file with equally sized (30m) polygons was created using the study area extent. This file was combined with the Union files using the Intersect tool, consisting of equally sized polygons (Figure 17).

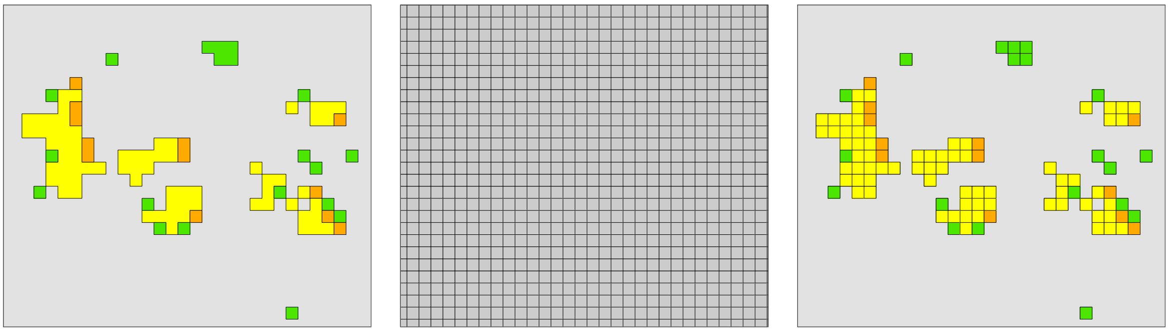


Figure 17 Unequal forest change polygons (I), Fishnet (II) and resulting equal polygons (III)

Until this point, the whole VGTB region was included in the process. However, as explained in Section 1.1.2, the topographic and socio-economic characteristics of the Eastern coastal area differs considerably from the inland midlands and highlands. Furthermore, the area that is suffering from deforestation and forest degradation is first and foremost situated in the inlands. Therefore, the modelling part of this research concentrates only on the inland area (i.e. the *focus area*). From this point in the analysis onwards, only the focus area is considered. Therefore, only the polygons that fall completely within the focus area boundaries were selected and copied into new files. A new field "Superclass" was added to the four files, in order to make a distinction between changed (1) and unchanged (0) forest cells. Also, a field with the era was added, allowing distinction between 2001-2005 (value 1) and 2005-2010 (value 2) in SPSS. Next, the files were combined per era using the Merge tool, resulting in two files:

- 2001-2005 forest change & unchanged area
- 2005-2010 forest change & unchanged area

Finally, one more step was taken. The two files were combined with the data on administrative zones using the tool Spatial Join. This enabled analysis on province, district and commune level.

These resulting files were then imported in SPSS and used for statistical analysis.

4.2.4 Spatial statistics

Although ArcGIS is also able to perform simple spatial statistical analyses, SPSS provides more robust methods that allow for more complex analysis on correlations and regressions. First, extreme outliers as a consequence of "no data" in the DEM layer were deleted. The two era files were combined into one data file, although treated separately during most statistical analyses based on the newly added "era" attribute. This section will describe the steps taken in SPSS. The results per factor will be presented in the next section.

Histograms

Histograms help us as a first step into analysing the data and discovering main differences between the forest change cells and the control group of unchanged forest cells. They show us also whether the distribution of values is normal or not. An overview of all histograms can be found in Appendix C.

Frequency statistics

Whereas histograms provide us a visual overview of the distribution of the values, frequency statistics are able to quantify the shape of the distributions (Field, 2009). As can be derived from the skewness and kurtosis values, which should be zero or close to zero for normal distributions, only the DEM and Slope distributions are more or less normally distributed. This has an effect on the options for further statistical analyses. The complete frequency statistics table is presented in Appendix D and Appendix E.

Boxplots

The boxplots go a step further than the histograms and frequency statistics by allowing quick comparison between different forest change classes and between the forest change classes and the control group. The centre of the boxes show the median of the values, the box itself shows the upper and lower boundary in between 50% of all values fall. The boxplots are presented by potential forest change factor (DEM, slope, distances) and within the boxplots a distinction is made between the two eras. All boxplots can be found in Appendix F.

Independent-samples T test

In most statistical analysis, one wants to know whether the variances of two different sample groups are equal. Contrastingly, here a difference in variances of the two sample groups (changed and unchanged forest) might indicate a potential relation between the factor tested and forest change. Therefore, an Independent-samples T test was performed.

First, it is necessary to state the hypotheses for the different factors:

1. Lowlands are more prone to deforestation than highlands (DEM)
2. Forest change is more likely to happen on flat areas rather than on steep areas (Slope)
3. Forest change is more likely to happen near roads (All Roads)
4. Forest change is more likely to happen near main roads (Paved Roads)
5. Forest change is more likely to happen near settlements (Small Settlements)
6. Forest change is more likely to happen near large towns (Large Settlements)
7. Forest change is more likely to happen near areas that were classified as cropland in the starting year of the era (Cropland)
8. Forest change is more likely to happen near areas that were classified as grassland in the starting year of the era (Grassland)

The Levene's test was used to check whether the assumption on equal variances (the null hypothesis) should be rejected or not. With an alpha level of 0.01, all factors were tested to be significant, which means that the null hypothesis on equal variances can be rejected. A complete overview of the outcomes of the Independent-samples T test is presented in Appendix G.

Bivariate correlation

Next, it should be tested whether there is a relationship between on the one hand forest change and on the other hand the different factors. Moreover, it would be interesting to see the direction of these correlations.

To find a relationship between forest change (dichotomous) and the topographic and thematic layers (interval/ratio), a bivariate correlation test using Spearman's correlation coefficient is needed, which is a non-parametric statistic (Bryman, 2008; Field, 2009). The coefficient ranges from -1 (complete negative relation) to +1 (complete positive relation). Values close to zero represent a weak or no relationship. All factors showed a significant negative relationship with the forest status of the cells. Unchanged forest was given the value 0 and Changed forest the value 1. Thus, a negative relationship indicates that the higher the values of the different factors (altitude, slope, and distances), the lower the chance was that a cell was classified as being deforested. The strongest relations were found for distance to cropland (-0,595 for 2001-2005) and distance to small settlements (-0,568 for both eras). Although still significant, the weakest relations were distance to paved roads (-0,091 for 2005-2010) and slope (-0,174 for 2005-2010).

SPSS generates pivot tables, showing the results in both the rows and columns. Therefore, superfluous rows were deleted from the table. The table can be found in Appendix H.

Logistic Regression

Logistic regression is useful in situations where the outcome variable (here: forest change) is categorical (here: dichotomous) and the predictor variables (here: DEM, slope, distances) are continuous (Field, 2009). It allows the user to predict the outcome given certain other information.

The classification tables of the binary logistic regression in Appendix I show that with the current values available, the model its predictive value can already increase from 52.1% (null model) to 81.4% (predictive model). The values that play the largest roles in the prediction equation are Slope, Distance to Cropland and Distance to Grassland for both eras[2].

4.3 Results

4.3.1 Hypothesis testing part I – factors that influence the deforestation potential

The first part the results regards whether or not the factors influence the chance that a particular forest area becomes deforested, i.e. the *deforestation potential* of forest areas. As said before, all test results can be found in Appendix C to I. In this section, the most important results are highlighted.

Digital Elevation Model (DEM)

Based on the histogram, it can be concluded that forest change is not normally distributed over DEM but occurs mostly in the lowlands. As can be derived from the boxplot, within the forest change group, the forest-to-cropland class deviates the most compared to the control group (Figure 18). One may conclude that cropland is highly sensitive to DEM and is started mostly in areas with a low altitude. Based on the correlation coefficients, one may notice that the influence of DEM on forest change decreases slightly over the years, indicating that forest change is more and more also occurring on higher altitudes. This is also reflected by an increasing mean for the DEM of forest change cells when comparing the two eras.

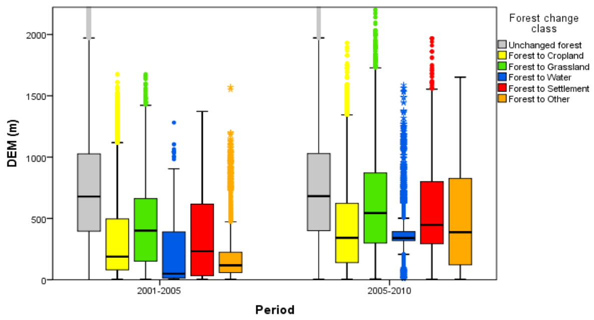


Figure 18 Boxplot - DEM

Slope

The values for slope of forest change cells are normally distributed. Again, cropland is most sensitive for the slope values, whereas grassland does not differ considerably from the control group.

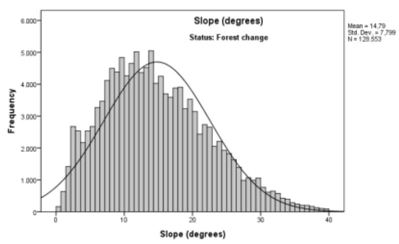


Figure 19 Histogram – Forest change 2001-2005 – Slope

[2] In terms of their own units, so degrees for slope and meters for distance to cropland and distance to grassland.

Distance to All Roads

For distance to all roads, all forest change classes differ greatly from the control group. The frequency statistics (Appendix D and Appendix E) reveal high kurtosis values which indicates a pile-up of values to the left of the distribution. In other words, forest change occurs mostly within a short range along roads. As the distance to roads increases, the number of forest change cells decreases drastically.

Distance to Paved Roads

For distance to paved roads, the trends are similar to the previous factor. There are far less paved roads when compared to all roads, which explains the much higher values, but also higher ranges. Especially for the 2005-2010 era, distance to paved roads does not seem to have a large correlation with forest change, that is, -0.091, although the negative relationship is still significant (Appendix H).

Distance to Small Settlements

Again, the values for distance to small settlements differ considerably from the control group. As an average, it can be stated that deforestation occurs approximately within 1,6km from small settlements. The boxplot clearly shows that forest to settlement mainly occurs near existing small settlements, although this does not reflect any absolute numbers (Figure 20).

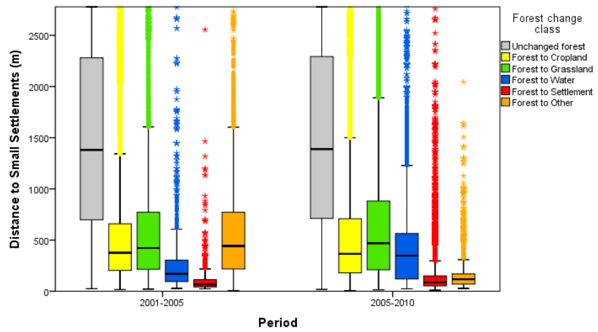


Figure 20 Boxplot – Distance to Small Settlements

Distance to Large Settlements

The trends of distances to large settlements are similar to those of small settlements, but the ranges are much wider. Although the variable in the equation of the logistic regression (Appendix I) might suggest otherwise (-0.00004), distance to large settlements does have a relatively high correlation value (-0.567) (Appendix H). The low B value in the logistic regression model may be due to the fact that in general, the values for distances to large settlements are high (unlike for example the distance to small settlements). The B values must be weighted using mean values before use for correlation and regression purposes.

Distance to Cropland

Change from forest to cropland has the highest absolute value (Schultz & Avitabile, 2012), which is why it was presumed that the distance to existing cropland would be highly correlated with forest change. The results confirm this hypothesis with a correlation value of -0.595 (Appendix H). The boxplot for the 2005-2010 era shows that in comparison to the other classes, forest to cropland occurs closest to existing cropland (Figure 20). However, for the 2001-2005 era, these numbers do not vary considerably. As an average, all forest change classes occur within approximately 150m from existing cropland.

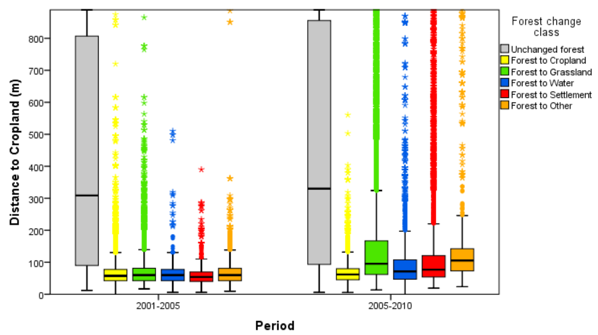


Figure 21 Boxplot – Distance to Cropland

Distance to Grassland

As expected, the boxplot shows that forest to grassland occurs close to existing grassland (Appendix F). What is striking is that, especially in 2001-2005, the distance to grassland is not as highly correlated to forest change as distance to cropland (Appendix H). Still, the histogram shows that most of the deforestation occurs within 200m from existing grassland.

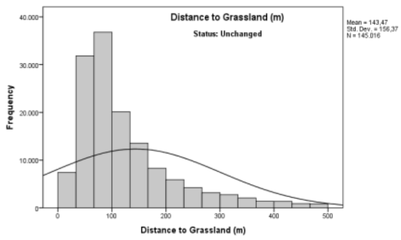


Figure 22 Histogram – Forest change 2001-2005 – Distance to Grassland

4.3.2 Regional differences & time era differences

As said before, the area of the VGTB river basin is characterised by large topographic and socio-economic differences. In general, the area varies from mountainous regions in the East to low and flat areas in the West. As Figure 23 shows, there are also large differences regarding the population density between East and West.

To check whether there are spatial differences regarding *where* forest changes occur and whether these possible spatial differences are time dependent, a final analysis was done based on the means of all topographic and thematic layers.

This analysis was done by calculating the means of each factor for each commune in Excel and visualising the results in ArcGIS. The results can be found in Appendix J, Appendix K and Appendix L. The most prominent differences per factor will be discussed here.

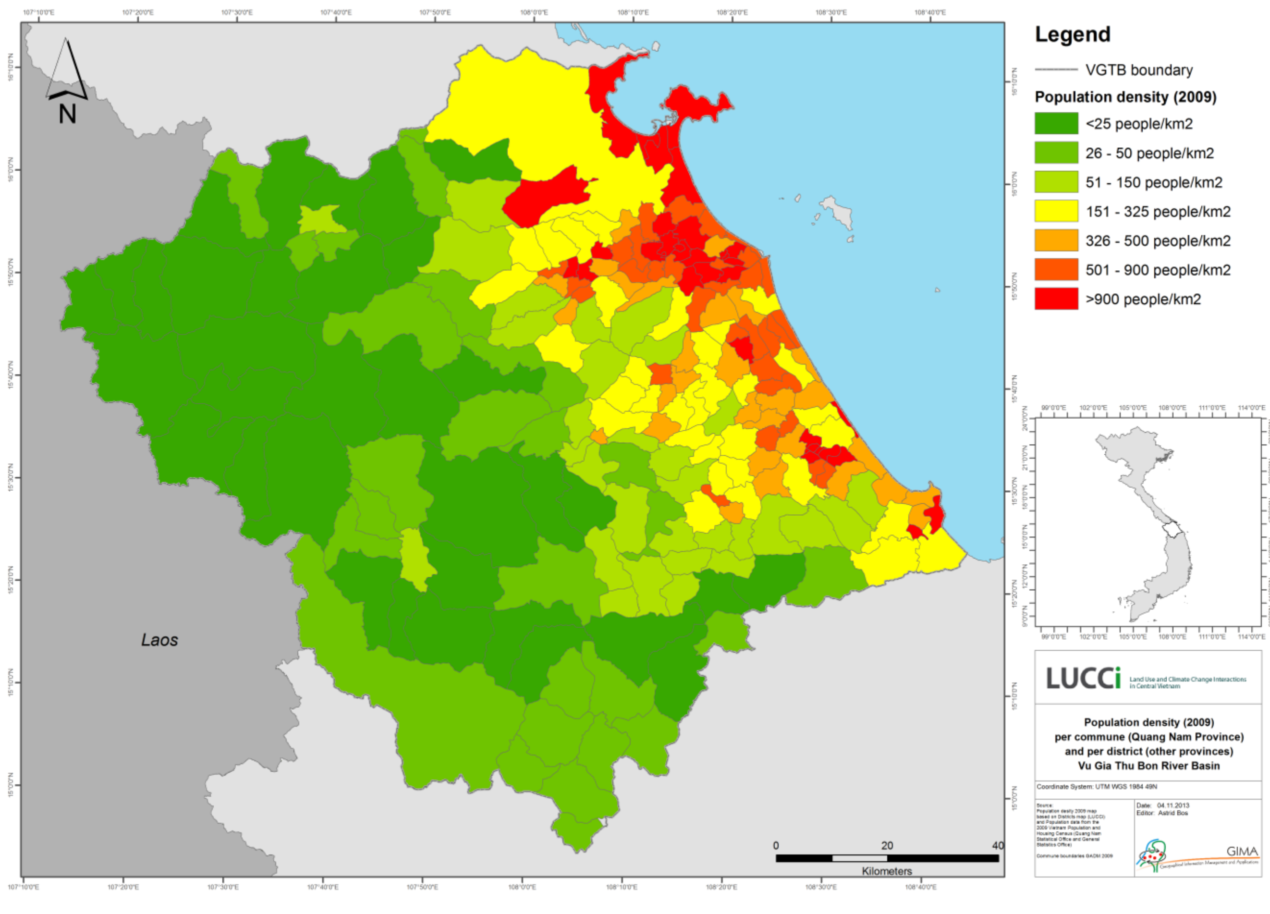


Figure 23 Population density per commune and district (2009)

DEM

The average DEM for forest change cells is notably correlated to the values of the general DEM. Still, deforestation tends to happen more in lowland areas, or in the valleys of the mountainous areas. When comparing the two eras, there is a slight tendency that the average DEM for forest change cells is increasing in the midlands of the communes in the second era (2005-2010), which could indicate that people look for cropland in areas with higher altitudes when compared to the first era (2001-2005).

Slope

Here, the same conclusion can be drawn as for the DEM. The slope of deforested areas increases slightly towards the mountainous areas, but overall, deforestation is occurring most in flat areas.

Distance to All Roads

In remote areas in the eastern part of the region, the average distance of deforested cells to the roads is much higher than in other areas, which is in line with the expectations since in the remote areas the presence of roads is limited.

Distance to Paved Roads

There are only a limited number of paved roads in the area. Deforestation cells in the upper north west and south have therefore a higher average mean distance.

Distance to Small Settlements

In general, by far the most communes have an average mean for distance of forest change cells to small settlements of less than 800 meters. This average seems to decrease slightly when comparing the two eras.

Distance to Large Settlements

As mentioned earlier, the communes in the south western part of the region marked as red (Appendix L) may be less remote in reality, since the shapefile on which the Large Settlements data was based only included Da Nang and Quang Nam province.

Barring that in mind, the average distance seems to increase slightly when comparing the two periods.

Distance to Cropland

Although the two maps representing the two different eras show large differences, it must be noted that the maximum and minimum, and even the mean, values did not differ too much. Overall, deforestation occurs relatively close to cropland, that is, mostly within 125m from cropland. Small shifts within these 125m lead to large differences in the colours on these maps.

Distance to Grassland

In the eastern part of the region, only few cells are classified as grassland, which is why the average distance to grassland of forest change cells is relatively high. When comparing the two eras there seems to be an overall tendency that the distance to grassland is decreasing.

4.3.3 Hypothesis testing part II – differences based on deforestation magnitudes

The previous two sections presented the results of the spatial analysis focusing on different topographic and thematic factors. As a follow up, the idea was to distinguish between two types of agents for the ABM, that is, to distinguish large scale (commercial) deforestation caused by firms or large scale organisations, and individual households who deforest small patches for clearing land for agricultural purposes. To justify this classification of agents, the second phase of hypothesis testing focused therefore on the size of the deforested patches.

The question to answer was therefore:

Are there significant differences in levels of correlation between on the one hand patches of deforestation?

Again, this hypothesis was analysed using several potential deforestation indicators (Table 2).

	Factor	Type	Original file raster/ vector	Source:
1	Digital Elevation Model (DEM)	Topographic	Raster	DEM LUCCi
2	Slope	Topographic	Raster	Calculated from DEM
3	Distance to Forest Plantation (distinguishing between already existing; emerging; and future plantations)	Thematic	Raster	Based on a combination of land cover data (Landsat TM) and FIPI forest map (2001, 2005, 2010) (see Figure 9 on page 17)

Table 2 Factors included in spatial analysis, hypothesis testing part II

The results of the correlation analyses can be found in Appendix M. Overall, no substantial differences were found based on the size of deforested patches. Only for DEM there was a small but clear trend that large scale deforestation only occurred on lower altitudes.

The fact that no significant differences were found based on the magnitude of deforestation does not mean there is no difference between small and large patches per se, it does however say that there is no substantial difference when taking into account these three factors (DEM, slope and distance to forest plantations). Before distinguishing between agent types based on the size of deforestation, it is therefore wise to search for other, more significant factors. As will be explained in the next chapter in more detail, for this research it was decided not to base any assumptions for agent typology on weak links regarding the size of deforestation. Therefore, the first version of the SoDRA LUCCi model as presented in this thesis only regards one general agent type, that is, the rural household.

4.4 Discussion & conclusion

Before one can draw conclusions from the results presented in the previous sections, a few issues must be discussed that may have influenced these results.

First, it must be noted that the testing of the hypotheses of explanatory factors for forest change was limited to those eight (topographic and thematic) factors presented in this chapter. The factors were chosen based on literature research and observations during past fieldwork. Drivers from outside these indicators and from other scales (indirect drivers) are difficult to reveal. Furthermore, data availability limited the scope on potential forest change driver analysis. Future research may focus on exploring other potential factors, using the same methodology as presented here.

Second, it was difficult to check the quality of the data provided for this research. The data came from a variety of sources, and in general it was impossible to recover the source of the data. Also within the data files some irregularities occurred. In these cases, it was tried to take the most out of the data by distinguishing between the data features in total and features within the file that contained proper attributes. For example, the road layer was split in two layers. One contained all features (All Roads) while the other was used to analyse only those layers classified as paved road.

One might argue that calculating Manhattan distances is more realistic than calculating Euclidean distances, as has been done in this analysis. The road network layer that was available for this research did not suffice for any proper network analysis. Furthermore, it was assumed that in rural areas, people cross forests and other non-road areas in order to reach their cropland. In those cases, working with Euclidean distances rather than Manhatten distances is preferred. Still, it must be noted that natural boundaries such as rivers and steep mountain ridges have been ignored in this research, which may have had a small influence on the distances calculated.

A final note should be made concerning the distance calculation in eCognition. eCognition calculates distances by taking the average minimum distance of a segment to the closest neighbouring segment, regardless of the class of its own segment. This means that for segments (e.g. Forest to Cropland) that were overlapped by its thematic layer (e.g. All Roads), the distance of that segment to the nearest other segment classified as All Road was calculated, rather than having a distance value of 0. Still it must be noted that this is more a conceptual flaw rather than a practical one, as the general trend will not be influenced by this. The segments generated are relatively small, which means that in the All Roads example some segments should have been given a value of 0 rather than 30m.

That being said, a few conclusions can be drawn. All analysed factors in the first hypothesis testing are significantly correlated to forest change. Although correlation does not say anything yet about cause and effect, the correlations reveal that there is a negative relationship between forest change and DEM, slope and the different distance factors.

Of all factors, distance to cropland and distance to small settlements seem to have the largest (negative) relationship with forest change. Still, since also the other factors appeared to be significantly correlated to forest change, factors with a “minor” relationship should not be ignored when making the model. This is also indicated by the binary logistic regression model which takes all variables into account.

Although the second hypothesis testing regarding the size of deforested patches did not reveal clear relationships between the size of the patches and the different potential explanatory factors, this does not mean that there are no significant differences to be found between different agent types and their deforestation behaviour. For this research however, no clear agent typology can be created based on the spatial analysis only. Thus, the link between deforestation size and different agent types should be further investigated. Agent typology is often largely based on socio-economic characteristics. Whether these differences in characteristics were found for this case or not, will be discussed in the next chapter, which regards the agent-based modelling phase.

5. Agent-based modelling

5.1 Introduction

There are different types of modelling possible when one wants to explore LUCC patterns. For this research it is chosen to use an agent-based modelling (ABM) approach, by using a NetLogo programming environment. ABM allows the user to study the relationships between micro-level individual motives and macro-level social behaviour (Gilbert, 2008). It is particular of use when one wants to understand certain processes and their consequences. "(...) Unlike most mathematical models, agent-based models can include agents that are heterogeneous in their features and abilities, can model situations that are far from equilibrium, and can deal directly with the consequences of interaction between agents" (Gilbert, 2008, p. xi). Especially the interactivity between the agents and its environment is something that distinguishes ABM from other types of models. Using ABM, the user can test which differences in options lead to certain land changes. The goal of ABMs can vary widely, ranging from providing assistance in decision-making, gaining understanding through raising awareness, facilitating communication, promoting coordination or mitigating conflicts (Le Page et al., 2013). In this case, the model gives insights in the effects of interventions such as REDD measures. The advantages of such experiments are that they "(...) allow one to be sure that it is the treatment that is causing the observed effects, because it is only the treatment that differs between the treated and the control systems and the systems are isolated from other potential causes of change" (Gilbert, 2008, p. 3).

There are many different drivers of land use change, and these drivers are most often space and time dependent. Examples of such drivers include agent behaviour and preferences; land accessibility; transportation costs; positive and negative local spatial externalities; biophysical characteristics of the land including slope, elevation and soil quality; social relationships and norms, information availability and accessibility; demographic characteristics; and external institutional factors (Geist & Lambin, 2002; Maguire et al., 2005).

ABMs represent these potential influences to a greater or lesser extent, generating a complex or rather generic model. As Maguire et al. (2005, pp. 413-414) acknowledge, such a model should be:

- an agent decision model capable of implementing optimising, boundedly rational, and rule-based decision models for a heterogeneous group of agents;
- a network model capable of representing both social and transport networks;
- a model that expresses both positive and negative local spatial influences, flexible with respect to the impact radius and the functional form of diffusion;
- the ability to input spatial layers representing institutional, socioeconomic, and biophysical data and constraints;
- the ability to input relevant global socioeconomic and biophysical parameters; and
- the ability to link to separately developed biophysical process models."

However, to what extend a model is able to include and accurately represent these factors is often dependent on data availability and knowledge about the processes involved. Therefore, rather than to increase the uncertainty of a model by including more assumptions that still need to be tested, it is often better to limit the model to including only those parameters of which the modeller is reasonably sure.

The remainder of this chapter is structured as follows. The next section explains the methodology of the modelling process. For this research, an ABM called the Simulation of Deforestation Risk Areas (SoDRA) LUCCi model was created. The third section describes the SoDRA LUCCi model in detail. Section 5.4 discusses the calibration and parameterisation, and is followed by a presentation of the results for the business-as-usual scenario and for the combined REDD scenario. Section 5.6 deals with a reflection on the verification and validation of the model, after which the issue of uncertainty is discussed. The discussion and conclusions are given at the end of the chapter.

5.2 Methodology

5.2.1 Agent definition & typology

The decision-making actors for the SoDRA LUCCi model are individual rural households, i.e. farms, as they are considered to have the decisive power regarding land use and land change.

In most ABMs for LUCC purposes, these agents are further divided into different agent types. The reasoning behind this is that modelling individual behaviour for completely heterogeneous actors is practically impossible. Generating agent types allows modelling of heterogeneous behaviour while keeping it feasible to model the different behaviour styles. Formalising groups of agents into agent types can be a harsh process and a research in itself. Typology issues for ABMs are far from standardised, although Valbuena, Verburg, and Bregt (2008) present an empirical method for typology formulation in a LUCC model, based on differences in views, farm characteristics and location.

For this model however, it was hard to distinguish between different agent types since detailed socio-economic data was missing (see Appendix B) and the spatial analysis did not reveal clear differences in deforestation based on the size of deforestation patches, which would have indicated differences between large scale and small scale deforestation actors (see Section 4.3.3). Rather than to assume differences between agents based on for example location (lowland-highland) or purpose (for generating cropland or for starting forest plantations) that are hard to verify, it was decided that for the first version of the model as presented in this thesis, only one type of agent, i.e. the rural household, was modelled. Other agent types or subtypes of the rural households may become part of later version of the model, if the availability of socio-economic data allows proper behaviour simulation.

5.2.2 Site selection for first experiments – the DEMO version of SoDRA LUCCi

The SoDRA LUCCi model aims to reveal future deforestation risk areas in the VGTB river basin[3] and to give insights in the effects of REDD measures on the projected deforestation. There are different platforms available for creating and experimenting with ABMs, under which Swarm, RePast, GAMA, MASON and NetLogo. The model was programmed in the NetLogo environment, as this platform is considered to

[3] That is, the focus area, so excluding the coastal communes. See Section 1.1.2 for more details.

provide an easy to use programming language with excellent ABM building, observing and using capabilities (Railsback & Grimm, 2012). Furthermore, it is able to use and produce GIS data easily in the form of ASCII files. For practical reasons, it was decided to first design and test the model at commune level. Although detailed calibration tests and sensitivity analyses were executed using the model at VGTB level, the first tests were to executed using the demo version, which is why this commune should be reasonably representative for the whole river basin.

Therefore, the commune was selected based on the following criteria:

- Average number of inhabitants
- Average population density
- Average deforestation in ha
- Average deforestation rate (deforestation as percentage of total forested area)

Accordingly, the selected experimental area is situated in:

- Province Quảng Nam (hereafter Quang Nam)
- District Đại Lộc (hereafter Dai Loc)
- Đại Chánh Commune (hereafter Dai Chanh)

As will be explained in Section 0, an important parameter in the model regards the number of agents, i.e. rural households. For practical reasons, it was decided not to vary the number of agents within the model, immigration and migration out of the focus area are not considered. According to the Quang Nam Statistical Office (2010), in 2009 there were 5647 inhabitants in Dai Chanh commune.

Urban-rural ratios on commune level were not available, but according to GSO Vietnam (2010) 88,9% of the inhabitants of Dai Loc are considered rural inhabitants (Table 3). Furthermore, is was found that by far, the most households in the Dai Loc district consisted of one family, and the average number of family members in a household is 4.29 (IPUMS, 2011)(See Appendix N).

Total inhabitants	Urban	Rural
145.935	16.215	129.720
100%	11,1%	88,9%

Table 3 Urban-rural ratio for Dai Loc district

Hence, the number of rural households in Dai Chanh commune was estimated as follows:

$$\frac{Number\ of\ inhabitants * rural/urban\ ratio}{Average\ number\ of\ household\ members\ per\ household} = \frac{5647 * 88.9\%}{4.29} \approx 1170\ rural\ households$$

For the demo version of the SoDRA LUCCi model, all GIS layers were prepared and clipped using the border of the Dai Chanh commune. These layers (30m resolution) were loaded into the NetLogo environment using its GIS extension. The code was programmed in a stepwise process of continued testing and bug fixing. Monitors and plots were added to the interface tab to allow visual checks (Figure 24). After several rounds of testing, the code was used as initial input for the full version of the SoDRA LUCCi model.

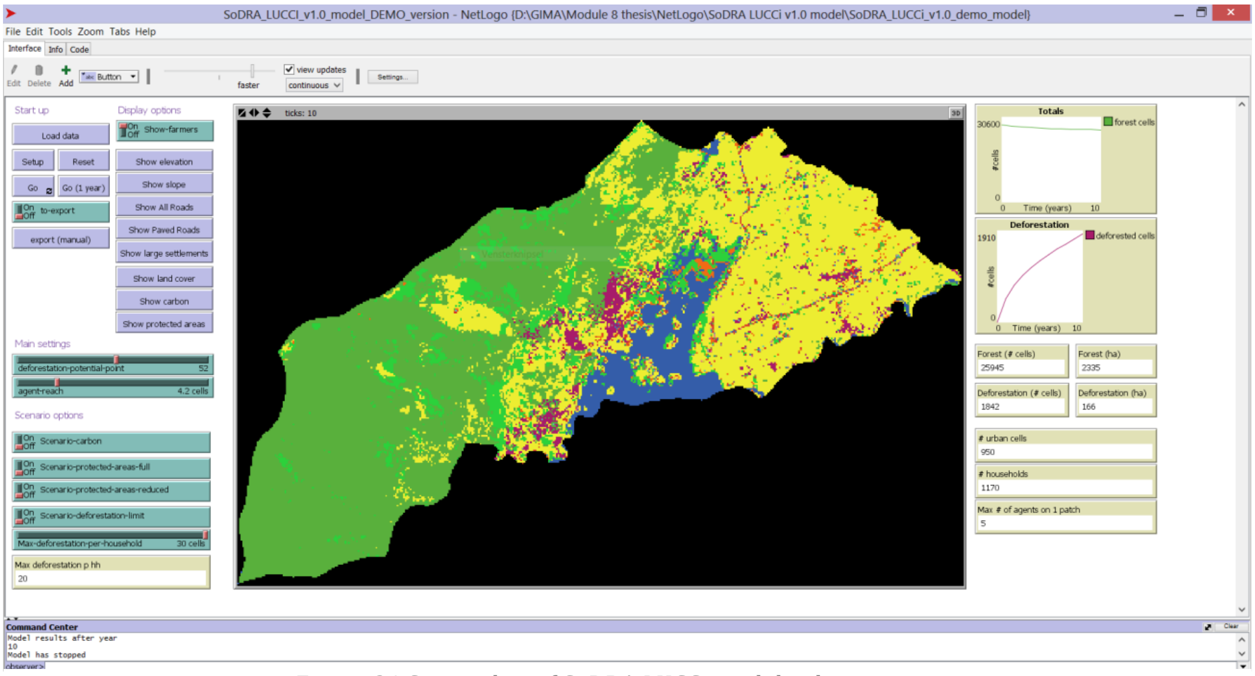


Figure 24 Screenshot of SoDRA LUCCi model – demo version

5.2.3 Expanding the model to the VGTB level

The next phase in the modelling process involved the expansion of the initial model to the VGTB level (excluding the coastal communes). First, all GIS layers were prepared and clipped using the focus area boundaries and loaded into NetLogo[4]. All GIS layers were resampled using the nearest-neighbour method or generated again at 90m resolution. The initial code needed to be adjusted to the new resolution (90m resolution) and the new number of agents for the whole area.

The number of agents for the whole focus area was calculated in similar ways as described in the previous section. However, several different data sources were used to gather the information needed for all communes and districts. An overview of the data sources and demographic values regarding inhabitants and households can be found in Appendix B. The number of rural households for the complete focus area is 107,262, which is equal to the number of agents simulated in the full SoDRA LUCCi model.

Calibration tests and sensitivity analyses were performed for this full SoDRA LUCCi model (see Section 5.4), leading to small adjustments of the parameter values for the full and revised DEMO model.

5.2.4 Implementation of REDD measures – the scenarios

The initial versions of the model represented only the Business-as-Usual (BAU) scenario, which assumes that future deforestation develops in similar ways as past deforestation. The final step in the modelling process involved the implementation of REDD measures, i.e. the scenarios. The REDD scenarios are represented by a set of options that can be implemented solely or as combination and tested for comparison with the BAU scenario. NetLogo offers three ways to vary parameters in experiments (explained in more detail in Railsback & Grimm, 2012):

[4] As Microsoft Windows does not allow JAVA to increase the RAM considerably, the SoDRA LUCCi model at focus area level was generated and tested in Ubuntu Linux. The initial resolution of the GIS layers was 30m, but this needed to be adjusted to 90m. Furthermore, the RAM of NetLogo was increased to 5500MB to allow for simulation of the large amount of agents and patches in the model.

- Sliders (to change a global variable on a continues scale)
- Switches (for Boolean variables)
- Choosers (defined options)

Only the first two options are used in SoDRA LUCCi. As explained in Section 1.1.3, the model only simulates the effects of REDD measures, rather than the means to achieve the effects. Table 4 shows the different REDD options modelled in SoDRA LUCCi model and their implications on cell level.

Carbon scenario	Switch	On Off Scenario-carbon	REDD measures may be based on those areas that have a high carbon value. If the switch is set to "on", areas with the highest carbon stock class will be fully protected. Deforestation in these areas is –effectively- prohibited. Here, the deforestation-potential value is set to 0. Areas with the second highest carbon stock class are protected in such a way, that they are less likely to be deforested. Here, the deforestation-potential value is reduced with 10.
Protected areas (full)	Switch	On Off Scenario-protected-areas-full	Protected areas may be national parks and reserves. In the protected areas –full scenario, any deforestation in the protected areas is –effectively- prohibited. The deforestation-potential value is set to 0.
Protected areas (reduced)	Switch	On Off Scenario-protected-areas-reduced	In the protected areas- reduced scenario, it is assumed that complete prohibition of deforestation in protected areas cannot be enforced. However, it is made more difficult to deforest these areas. The deforestation-potential value is therefore reduced with 10.
Deforestation limit per household	Switch/slider	On Off Scenario-deforestation-limit Max-deforestation-per-household 5 cells	Another REDD option could be to set deforestation quotas for each individual household. If the switch is turned on, the slider defines the maximum amount of cells (of 8100m^2 each for the full model) that individual households may deforested for the duration of the model run (here, 10 years). If a household reaches its limit before the end of the model run, it will be prohibited to deforest any more forest cells for the remaining years.

Table 4 The REDD scenarios and their implications for the model

5.3 Model description

In this section, the model will be described using the Overview, Design concepts and Details (ODD) framework. The ODD is a description protocol that allows communication, replication, and comprehension of simulation models (Grimm, Polhill & Touza, 2013). The ODD is different from other model descriptions since conventional differential equations and statistical modelling cannot describe how an ABM functions. The purpose of ODD is "to facilitate writing and reading of model descriptions, to better enable replication of model-based research, and to establish a set of design concepts that should be taken into account while developing an ABM" (Grimm et al., 2013). The ODD for the SoDRA LUCCi model is presented in Table 5. The full code with additional comments can be found in Appendix P. A screenshot of the interface tab of the SoDRA LUCCi model is given in Appendix O. Since this first version of SoDRA LUCCi is still a relatively simple model, not all elements of the ODD protocol are applicable.

Table 5 Elements of the SoDRA LUCCi model following the ODD protocol

Overview	1. Purpose	The purpose of the SoDRA LUCCi model is to model future deforestation risk areas by modelling deforestation behaviour of individual households. Moreover, it aims to give insights in the effects of different REDD measures on the projected deforestation.
	2. Entities, state variables, scales	The model distinguishes one agent type, i.e. rural households, and the environment, represented by patches in NetLogo. The *global variables* represent the GIS layers that are loaded in the background of the viewer. As a default, only the landcover-dataset is visible. The global variables are: landcover-dataset; elevation-dataset; slope-dataset; allroads-dataset; distance-allroads-dataset; pavedroads-dataset; distance-pavedroads-dataset; largesettlements-dataset; distance-largesettlements-dataset; carbon-dataset; protected-dataset[5]. The *agent variables* represent the characteristics of the individual agents, i.e. the rural households. The agent variables are: Start-patch (“home”); forest-harvest (number of deforested cells in model run) The *patch variables* represent the state of the individual environment cells and are often derived from the GIS layers in the background. The patch variables are: land-cover (class); elevation; slope; distance-allroads; distance-pavedroads; distance-smallsettlements; distance-largesettlements; carbon (class); protected-area; deforestation-potential (defines the "suitability" of forest cells for deforestation); occupied (counts the number of agents on their home patch) DEMO version: Spatial extent Dai Chanh commune; Spatial resolution 30m Full version: Spatial extent VGTB focus area; Spatial resolution 90m

[5] Please note that the area classified as “protected area” in the DEMO SoDRA LUCCi model is fictitious. This DEMO version should be used for demonstration purposes only.

3. Process overview, scheduling

Each round (tick) the agents search within their reach (specified in the interface slider, default 1.5 cells) for the forest cell with the highest deforestation-potential value. If the deforestation-potential is higher than the deforestation-potential on its own (current) cell, the agent will move to that cell. If the deforestation-potential is also higher than a set threshold (deforestation-potential-point, default 52), the agent will deforest that cell (see Figure 25).

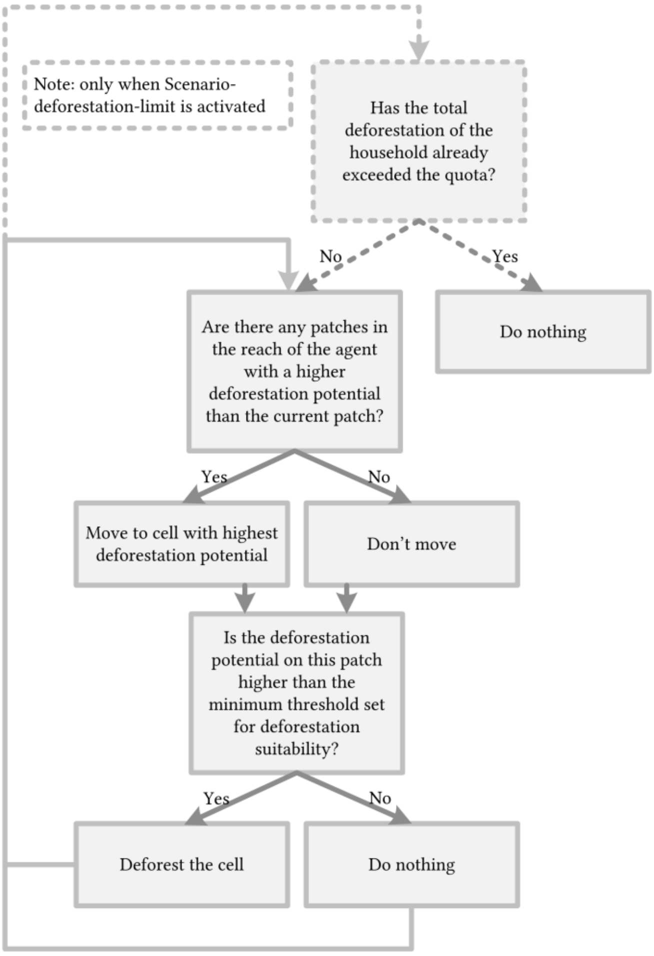


Figure 25 Process tree of agents' considerations and decisions

Design concepts	4. Design concepts	Basic principles	The model is based on the conceptual model as presented in Section 3.1. Forest change is enforced by humans actions who base their decisions on socio-economic and biophysical factors. The deforestation-potential value of forest cells is calculated using eight weighted factors that were proven to be correlated to deforestation (see Chapter 4).
		Emergence	The main model output consists of the projected deforestation cells. Both the quantities and patterns of deforestation that emerge from the simulations can be exported.
		Adaptation	Agents search for the highest deforestation-potential cells within their reach. A forest cell can only be deforested by one agent once. Other agents in the area will thus search for other cells in their reach suitable for deforestation.
		Objectives	The agents' objective is to deforest one cell per year (tick). This objective is thwarted by both environmental conditions (i.e. the deforestation-potential is too low or the reach of the agent is too small) and institutional interventions (i.e. the REDD measures)
		Learning	This version of the model does not incorporate learning processes
		Prediction	N/A
		Sensing	Agents are assumed to be able to measure the deforestation-potential of forest cells within their reach. In other words, they know when a forest cell is too far from a road, too steep, etc., and thus not suitable (enough) to take the effort to deforest the patch.
		Interaction	There is no interaction between the agents. With regards to the starting patch (home) only a limited amount of agents can be placed on one cell (i.e. 10 farmers for the full version). If the limit is reached, remaining agents need to be placed elsewhere.
		Stochasticity	The agents are placed randomly on urban cells, because information on exact locations (villages, towns) is lacking. The order of agents in procedures (placement, move, deforest) is random.
		Collectives	Agents represent households. There are no aggregations of households in the model, since all 107,262 rural households are modelled individually. If the model will be extended or become more complex in the future, aggregation of agents is advised.
		Observation	The interface tab of the model shows monitors (quantities) and plots (graphs) of all relevant outcomes of the model, that is, deforestation and forest rates (absolute, in cells, ha, and km^2) and the maximum amount of deforested cells per household during the whole model run.

Details	5. Initialisation	The initial land cover is based on the known land cover (based on the Landsat TM data) for 2010. The agents (107,262) are randomly placed at urban cells. When, after the first model run, the model is reset, the agents return home. The deforestation-potential values are recalculated, i.e. reset to original values.
	6. Input data	The input data used in the model can be derived from the global variables. It includes the initial land cover data, the two topographic inputs (elevation and slope), the six distance factors (the distance to cropland and distance to grassland are directly calculated in NetLogo using the land cover values), and the input for the scenarios (carbon data and protected areas).
	7. Sub-models	N/A

5.4 Calibration & parameterisation

Calibration involves a "process of determining appropriate values for one or more parameters that are not specified by theory or past practice" (Maguire et al., 2005, p. 14). It is often done using data from past history. As Railsback and Grimm (2012, p. 255) state, "calibration is a special kind of parameterisation in which we find good values for a few especially important parameters by seeing what parameter values cause the model to reproduce patterns observed in the real system".

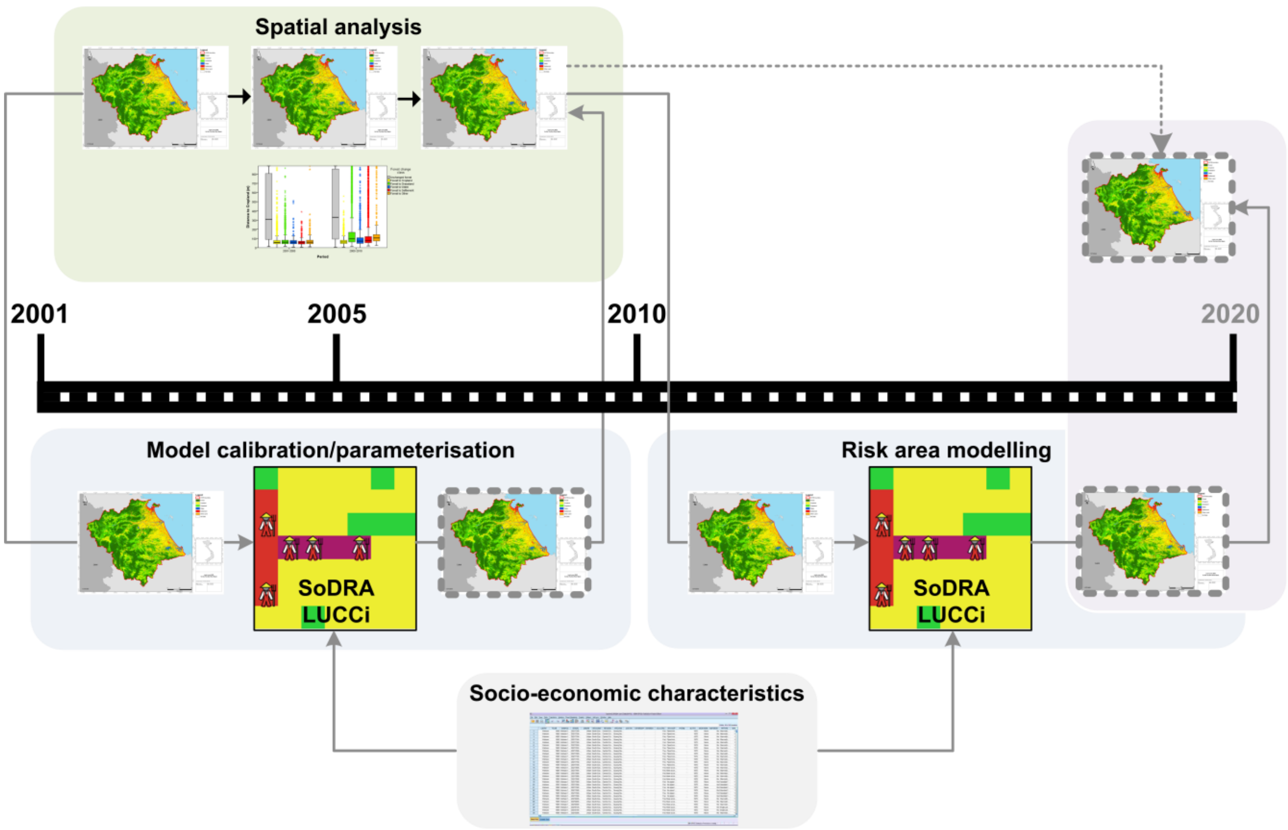


Figure 26 Visualisation of the modelling process

The purpose of model calibration is threefold (Railsback & Grimm, 2012):

1. To make the model more accurate and credible by forcing the model to match empirical observations as well as possible;
2. To estimate the value of parameters that we cannot evaluate directly (inverse modelling);
3. To test a model's structural realism, i.e. to test a model's robustness.

In this case, the calibration was based on the data of the satellite images and land cover maps. Instead of predicting the future, the model was used to model the situation of our baseline year, i.e. 2010, based on the information available from 2001 and 2005. Following the pattern-oriented modelling approach (POM), it was attempted to reproduce patterns observed in reality in the modelled system (Grimm et al., 2005; Magliocca, Brown & Ellis, 2013). The calibration results were then used to adapt and refine the models parameters. Thus, the calibration is an iterative process which was executed throughout the whole modelling phase (Figure 26).

Although this is a proven method to calibrate models, it postulates that land change processes are rather linear and that changes in the past are similar to changes in the future. However, Bakker and Veldkamp (2012) argue that relationships between land use and the environment are non-stationary in time. Still, since this research covers a timeframe of only 20 years in total, it is thought that this way of calibrating the model was the soundest choice.

5.4.1 Calibration of the parameters agent-reach and deforestation-potential-point

Two parameters in the model cannot be verified using statistical values based on past trends because they are rather fictitious. These parameters are (1) the search area in which the agent each year search for a cell to deforest (*agent-reach*) and (2) the threshold value below which the agent will not deforest a cell because it is not appropriate enough to deforest (*deforestation-potential-point*).
Generally, there are two types of calibration (Railsback & Grimm, 2012):

1. *Categorical calibration*: search for parameter values that produce results within a category or range you defined as acceptably close to the data
2. *Best-fit calibration*: search for parameter values that cause the model to best match some exact criteria (optimisation).

Here, it was decided to use a combined approach. First, categorical calibration took place based on the DEMO version of SoDRA LUCCi. Afterwards, it was further calibrated with both categorical and best-fit calibration on the full version of the model. It was decided not to perform any best-fit calibration at the DEMO version, since there is a plausible change that Dai Chanh is not perfectly representative for the whole VGTB area, generating a risk of overfitting at the sample area. To limit the time needed for calibration at VGTB level however, the categorical calibration at Dai Chanh level defined the parameter space (see alsoRailsback & Grimm, 2012, p. 259) in which the parameter values had to be found.

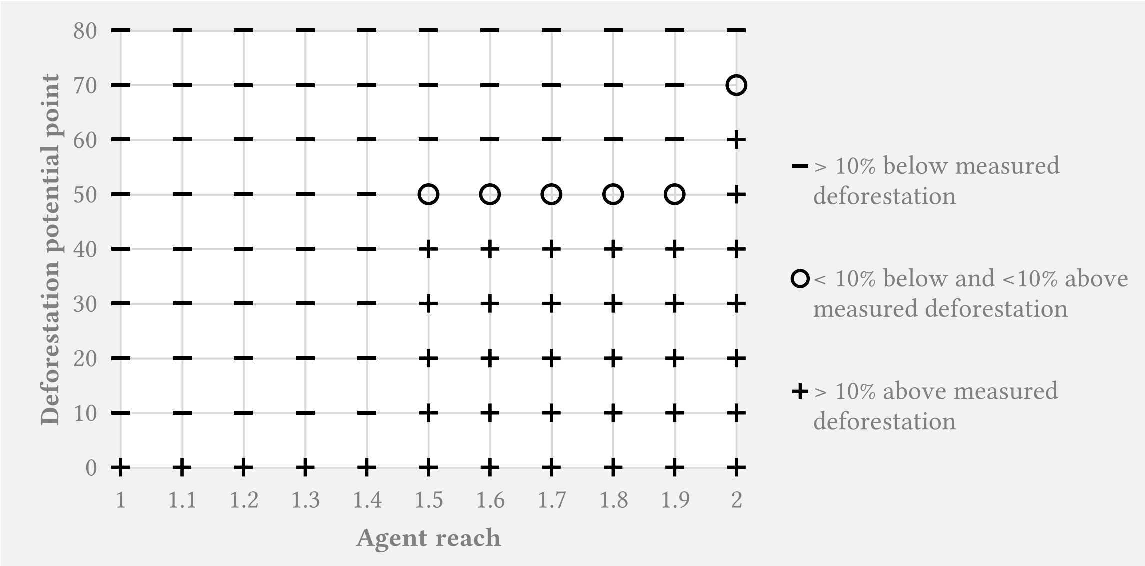


Figure 27 Results first calibration test (0-10-80;1-0.1-2;1;99)

Table 6 presents the settings of the first categorical calibration test at the full version of SoDRA LUCCi. The outcomes of each model run, in terms of modelled deforestation amount, were compared with the actual measured deforestation rates of the land cover maps. The results of this test are presented in Figure 27. The circles represent values that are the most in accordance with the measured rates. These values were analysed further in the second calibration test.

Total runs	99		
Repetitions	1		
	Min	**Increment**	**Max**
Deforestation potential point	0	10	80
Agent reach	1	0.1	2

Table 6 Settings for first calibration test

The settings for the second calibration test are given in Table 7, with the corresponding results in Figure 28. Here, each model run was repeated five times. Consequently, the ⊕ symbol in the graph shows that with the same parameter values, the model output occasionally matched the measured deforestation and in another run showed a too high deforestation rate compared to the measured deforestation. A ⊖ symbol would indicate the presence of both matching and too low deforestation model outputs, but this did not occur in this calibration test.

Total runs	200		
Repetitions	5		
	Min	**Increment**	**Max**
Deforestation potential point	50	5	70
Agent reach	1.4	0.1	2.1

Table 7 Settings for second calibration test

Since the deforestation-potential-point of 70 -meaning that only forest cells with a score of 70 or higher on a scale from to 100 are appropriate to deforest- was considered to be not very realistic, it was decided to narrow the third test down using the remaining parameter combinations with a "○" score.

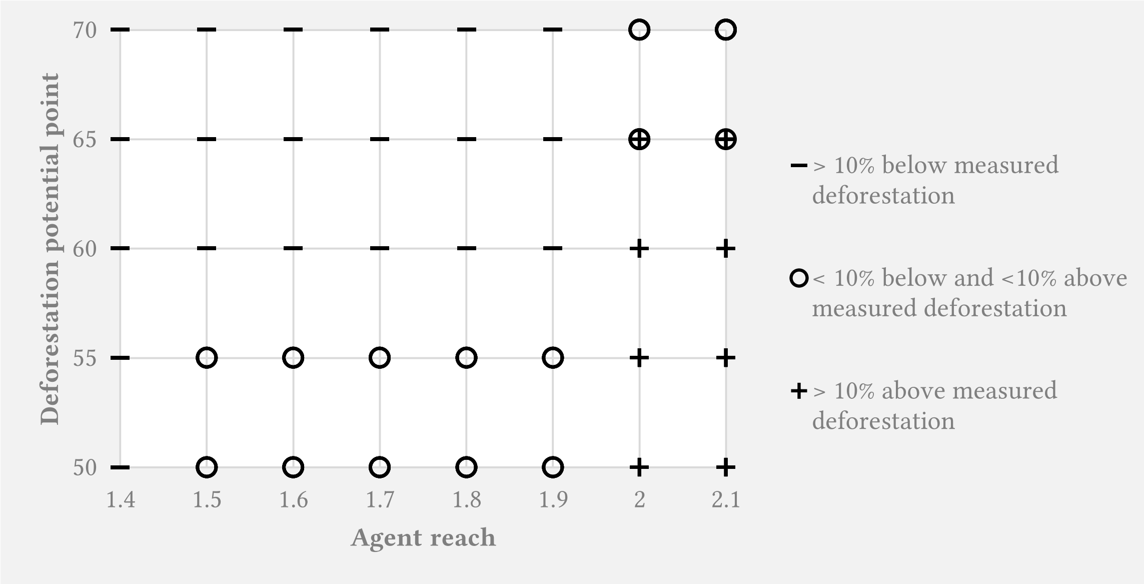


Figure 28 Results second calibration test (50-5-70;1.4-0.1-2.1;5)

A third and final calibration test aimed at best-fit calibration. The settings are presented in Table 8 and the corresponding results in Figure 29. This time, only model output values within 5% above and below the measured deforestation were considered good matches. When looking at the 5 repetitions, the parameter combination of 52 (deforestation-potential-point) and 1.5 (agent-reach) showed the best-fit. From these five model runs, the average amount of cells that were deforested after 10 years was 29,223.8, against 29,224 cells[6] in the land cover data. These best-fit parameter values were used in the final version of the model, including the sensitivity analysis runs.

Total runs	490		
Repetitions	5		
	Min	**Increment**	**Max**
Deforestation potential point	45	1	58
Agent reach	1.4	0.1	2.0

Table 8 Settings for third calibration test

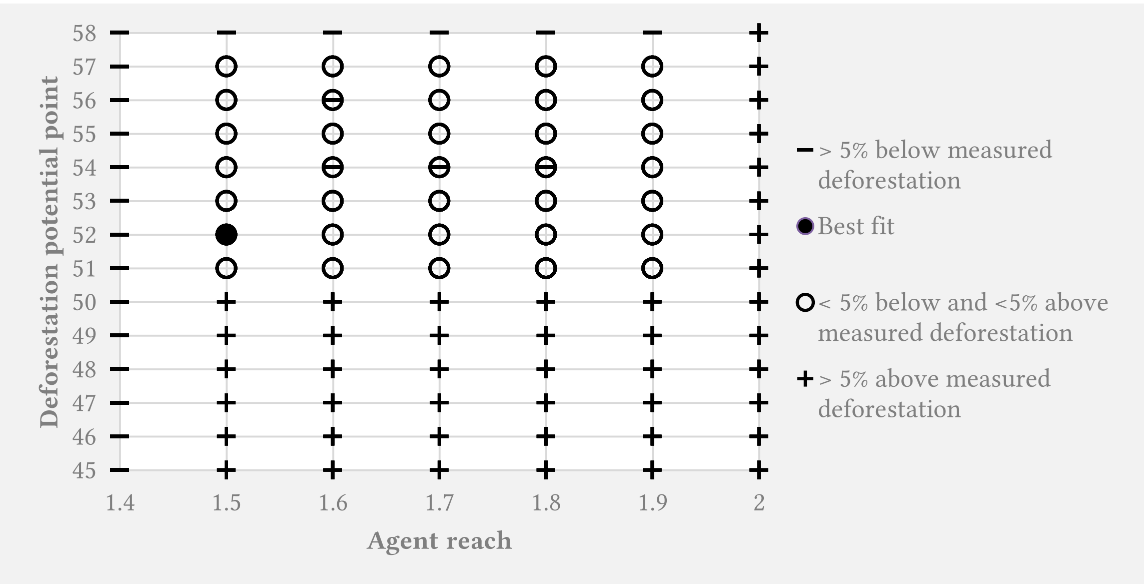


Figure 29 Results third calibration test (45-1-58;1.4-0.1-2.0;5)

6 23,671 ha derived from 30m resolution land cover data. Recalculated to 90m resolution results in 29,224 cells

5.4.2 Parameterisation of the factors influencing the deforestation-potential variable

The deforestation-potential value defines the "suitability" or likelihood of a forest cell to be deforested. This likelihood is influenced by the factors examined during the spatial analyses (Chapter 4) and their individual weights. This will be explained by the code part for influencing the deforestation-potential value based on elevation.

```
if (land-cover = 1) and (elevation <= 618)
  [set deforestation-potential deforestation-potential + 12]
```

For each factor, a threshold was taken based on 75% of the forest changed cells in the spatial analysis. For the elevation factor, it means that 75% of the deforested cells had a DEM value of 618m or lower. The other 75% values are presented in Table 9.

	75% (real)	75% (absolute)
DEM	618.00	618
Slope	21.11	21
Distance to All Roads (m)	634.36	634
Distance to Paved Roads (m)	6900.94	6901
Distance to Small Settlements (m)	702.92	703
Distance to Large Settlements (m)	1845.47	1845
Distance to Cropland (m)	87.72	88
Distance to Grassland (m)	85.15	85

Table 9 Parameter definition for the individual factors influencing the deforestation-potential value

Next, all factor were weighted based on their correlation factors (Table 10; see also Appendix H). For elevation, a combined and normalised correlation factor of 0.117 was found, i.e. 12 when scaled to 100. All weights added up to a maximum of 100, although implementing the different REDD scenarios can lower some or all deforestation-potential values.

	Era	DEM (m)	Slope (degrees)	Distance to All Roads (m)	Distance to Paved Roads (m)	Distance to Small Settlements (m)	Distance to Large Settlements (m)	Distance to Cropland (m)	Distance to Grassland (m)
Initial correlation factor (real value)	01-'05	-0.483	-0.271	-0.483	-0.389	-0.568	-0.576	-0.595	-0.256
Initial correlation factor (absolute value)	01-'05	0.483	0.271	0.483	0.389	0.568	0.576	0.595	0.256
Normalised correlation factor	01-'05	0.133	0.075	0.133	0.107	0.157	0.159	0.164	0.071
Initial correlation factor (real value)	05-'10	-0.297	-0.174	-0.433	-0.091	-0.568	-0.432	-0.509	-0.462
Initial correlation factor (absolute value)	05-'10	0.297	0.174	0.433	0.091	0.568	0.432	0.509	0.462
Normalised correlation factor	05-'10	0.100	0.059	0.146	0.031	0.192	0.146	0.172	0.156
Combined normalised correlation factor	01-'10	0.117	0.067	0.140	0.069	0.174	0.152	0.168	0.113
Weight factor (scaled to 100)	01-'10	**12**	**7**	**14**	**7**	**17**	**15**	**17**	**11**

Table 10 Definition of weights of factors influencing the deforestation-potential value

5.5 Results

5.5.1 Business-as-usual scenario

The business-as-usual (BAU) scenario models the deforestation for 2010-2020[7] in the focus area of VGTB river basin[8]. The default values for the main parameters are based on the calibration tests and parameterisation as presented in the previous section.

To lower the effects of coincidences due to the (small) degree of randomness in the model, the SoDRA LUCCi model following the BAU scenario was run 50 times. The results are presented in Figure 30 (1). Most deforestation occurred in all model runs (dark red) and only little deforestation (light red) was due to the inherent randomness of the model. The projected deforestation shows patterns in which there is a high risk of deforestation to the northwest and central part of the focus area. The communes with the highest projected deforestation rates are presented in Table 11. More scattered deforestation occurs in the south of the focus area. Little to no deforestation is found in the remote areas closest to the border with Laos.

Rank	Commune	District	Province	Deforested cells	Ha deforestation
1	Xã Tư	Huyện Đông Giang	Quảng Nam	997	807.57
2	Xã Mà Cooi	Huyện Đông Giang	Quảng Nam	714	578.34
3	Thị trấn Thạnh Mỹ	Huyện Nam Giang	Quảng Nam	629	509.49

Table 11 Top 3 communes with highest modelled deforestation rate (BAU) 2010-2020

5.5.2 REDD scenarios

As explained in Section 5.2.4, the switches and sliders in the scenario options section on the interface tab of the SoDRA LUCCi model allow the user to experiment with REDD measures. This section presents the model results for the 2010-2020 era.

First, the REDD scenarios were tested individually (*ceteris paribus*). The quantitative results are presented in Table 12 and the corresponding maps in Figure 31. As expected, setting a quota of deforestation to a maximum of 3 cells per household per decade, leads to the largest reduction in deforestation when compared to the BAU scenario. Also the other quota scenarios have a substantial effect on the projected deforestation. The first three scenarios however, have only limited effect on the projected deforestation. The reason why prohibiting or discouraging deforestation in protected areas has such a minor influence is because in the VGTB area, only some areas in the east are defined as protected area. According to the model, only limited deforestation occurs in this area under the BAU scenario. Consequently, reducing the deforestation in these areas has only a minor impact.

As the maps in Figure 31 show, the patterns of deforestation do not differ between the scenarios, the magnitude, however, does vary considerably.

---

[7] Or, to be more precise, to the end of year 2019

[8] The results for the BAU 2001-2010 scenario were compared to the measured deforestation in Section 5.6.2

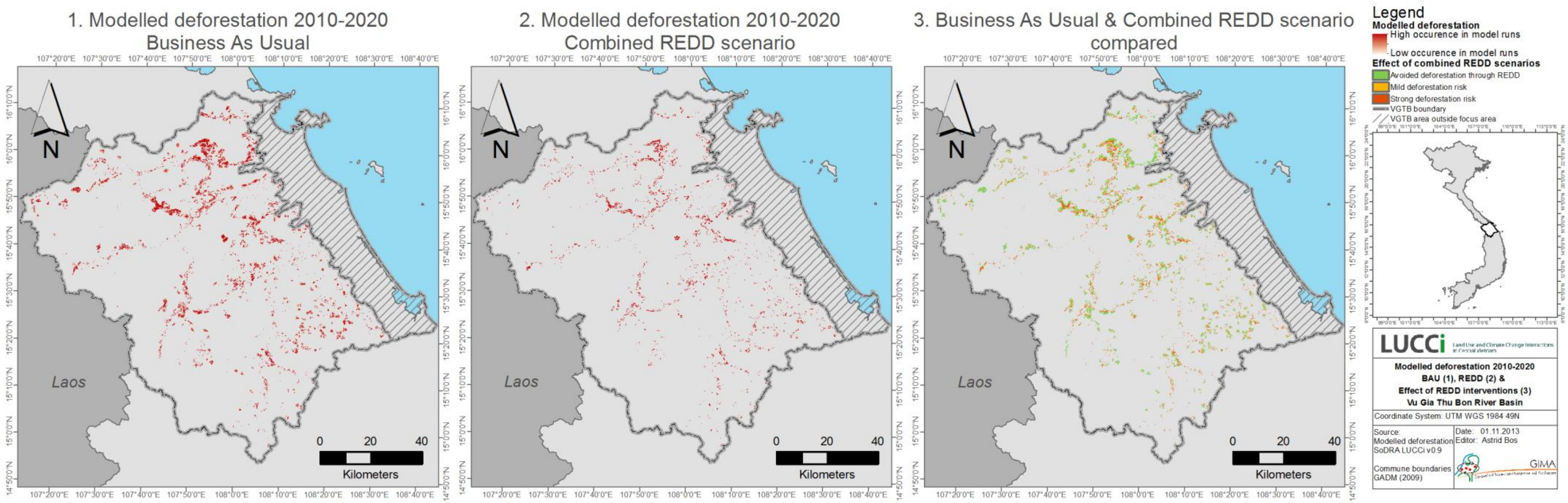


Figure 30 Modelled deforestation for 2010-2020 (1) BAU Scenario, (2) Combined REDD scenario and (3) BAU and combined REDD compared

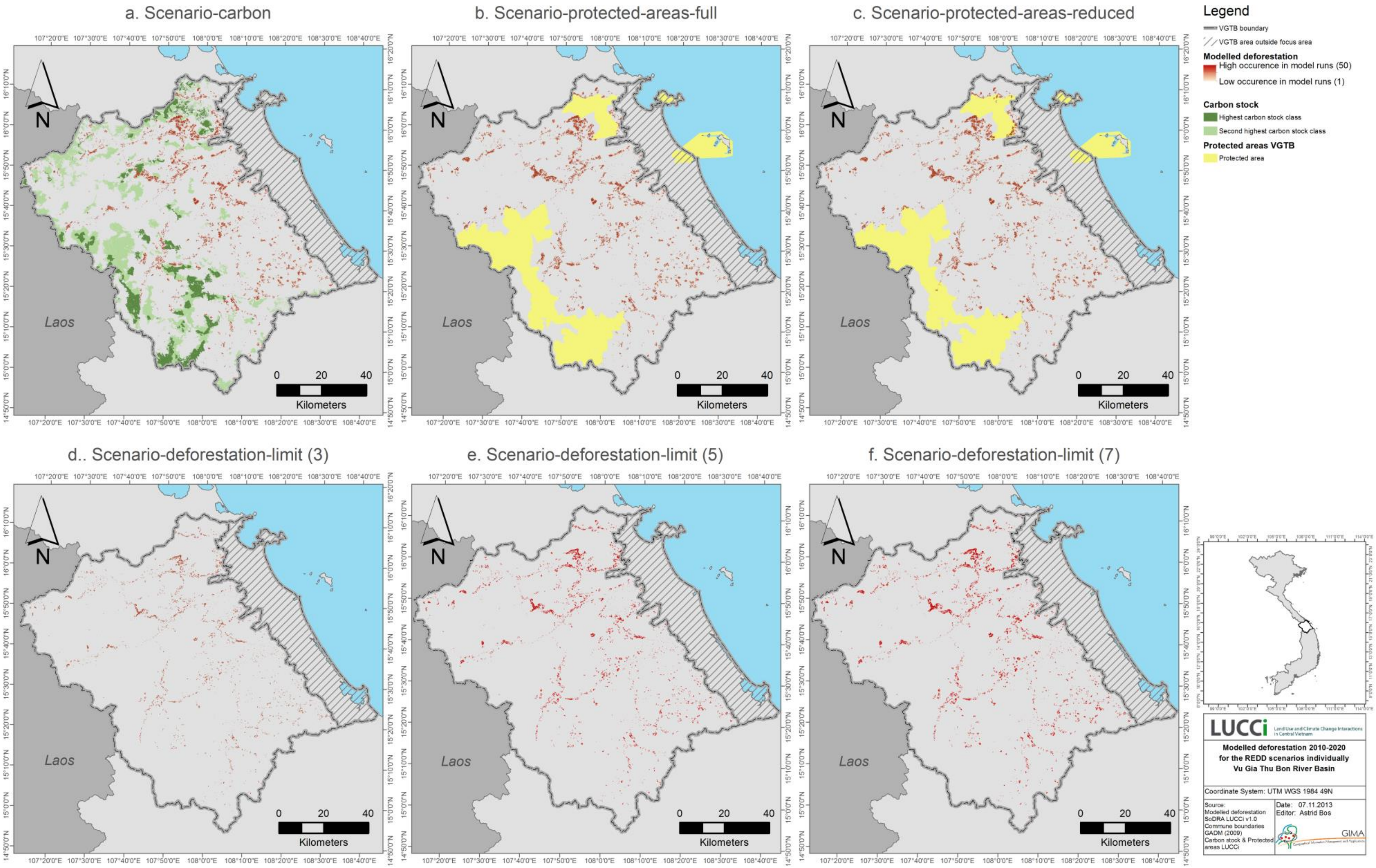


Figure 31 Modelled deforestation n2010-2020 – Individual REDD Scenarios

REDD Scenario	Average number of deforested cells (50 runs)	Average deforestation (ha)	Change compared to BAU
1. Scenario-carbon = true	27846.48	22555.65	-8.02%
2. Scenario-protected-areas-full = true	28060.14	22728.71	-7.20%
3. Scenario-protected-areas-reduced = true	29324.22	23752.62	-2.58%
4. Scenario-deforestation-limit = true AND Max-deforestation-per-household = 3	13508.00	10941.48	-122.68%
5. Scenario-deforestation-limit = true AND Max-deforestation-per-household = 5	18265.00	14794.65	-64.68%
6. Scenario-deforestation-limit = true AND Max-deforestation-per-household = 7	21774.26	17637.15	-38.14%
Scenario Business-as-usual	*30079.42*	*24364.33*	

Table 12 Quantitative results of individual REDD scenario model runs 2010-2020

Next, the results of a combined REDD scenario are presented in Figure 30 (2). The settings of this scenario are as follows (*ceteris paribus*):

Scenario-carbon	True
Scenario-protected-areas-full	True
Scenario-protected-areas-reduced	False
Scenario-deforestation-limit	True
Max-deforestation-per-household	5 cells

Again, the model using these settings was run 50 times. The results show a considerable decrease in deforestation. The overall patterns do not change much, but the amount of deforestation in the areas which were under particular deforestation risk under the BAU scenario, reduces significantly. In absolute numbers the deforestation reduces from an average of 30079.42 cells under the BAU to an average of 15996.42 cells under the combined REDD scenario Table 13. This entails a decrease in deforestation of over 46%. Nevertheless, it must be noted that in practise, it may be very hard to achieve this scenario. It entails a full protection and thus complete prohibition of deforestation in the national parks. Furthermore, in the remaining areas, households may still deforest, but for some of them this is half of the amount of deforestation when compared to the BAU scenario.

	cells	ha	km^2
Average deforestation under BAU scenario	30079.42	24364.33	243.64
Average deforestation under Combined REDD scenario	15996.42	12957.1	129.571
Change			-46.82%

Table 13 Deforestation rates under BAU and combined REDD scenarios compared 2010-2020

Still, the map in Figure 30 (2) in combination with Figure 30 (3) may give a good insight into *where* certain REDD measures may have the most significant impact. Figure 30 (3) presents an overlay of the BAU and combined REDD scenarios results. The legend items in represent the following classes:

Green - Avoided deforestation through REDD	According to the BAU model runs, deforestation occurred here, but in 100% of the model runs of the combined REDD scenarios, deforestation did *not* occur here.

Orange - Mild deforestation risk	According to the BAU model runs, deforestation occurred here, and in <50% of the model runs of the combined REDD scenarios, deforestation did *still* occur here.
Red - Strong deforestation risk	According to the BAU model runs, deforestation occurred here, and in >50% of the model runs of the combined REDD scenarios, deforestation did *still* occur here. Thus deforestation occurred in those areas, regardless of the implementation of the REDD measures.

Again, some areas in the northwest of the region show highest deforestation risk, since the most deforestation occurred here, regardless of the effects of the REDD implementation. Thus, these areas should get particular attention by the Vietnamese authorities.

5.5.3 Remarks regarding the results

Once again, it must be stressed that the model only incorporates forest change in the sense of deforestation. In Vietnam, and particularly in the province of Quang Nam, the government stimulates the construction of forest plantations (DONRE Tam Ky, 2010). In reality, there may be a forest net increase expected in the coming decade. However, since this study focuses on deforestation risk areas and afforestation is thus not considered in the SODRA-LUCCi model, the output from the model may only be considered and interpreted in a qualitative way. That is, it only provides insight in deforestation patterns, deforestation rates and deforestation risk areas; it does not however provide any insight in *net* forest change quantities.

5.6 Verification & validation

"We can never conclusively demonstrate that a simulation is fault free, but we can increase our confidence in a program within the range of the parameter space tested by adopting good programming and testing practices in order to minimise faults" (David, 2013, p. 141)

In order to put the results presented in the previous section into perspective, one should pay attention to the verification & validation of the model. This includes a sensitivity analysis on the most important parameters and, possibly, adjustment of these parameters for future versions of the model.

Model *verification* involves checking whether the model is built in a correct way. It entails the "the evaluation of the implementation of the model in terms of the researchers' intentions" (David, 2013, p. 136). A large part of model verification involves debugging (Gilbert, 2008) and changing the model codes accordingly. This is an iterative process, and took place during each phase of the modelling process.

Model *validation* tests whether the *right* model is built (Gilbert, 2008). Is the model an accurate representation of reality? In other words, it evaluates "the credibility of the model as a representation of the subject modelled" (David, 2013, p. 136). Verification and validation is often not linked to one particular stage in modelling, but rather to the modelling process in general in a way to add credibility to using the simulation for its specific purpose (David, 2013)(see Figure 32). Therefore, verification and validation should always be put in the context of the simulation's particular objectives.

When comparing the conceptual model as presented in Section 3.1 to the computer representation of reality, it should be noted that, in general, this first version of the SoDRA LUCCi model should still be valued as a basic abstract model rather than a model that represents reality to a high degree. The socio-economic drivers of agent's behaviour are underrepresented in the model due to the lack of (accessible) socio-economic data.

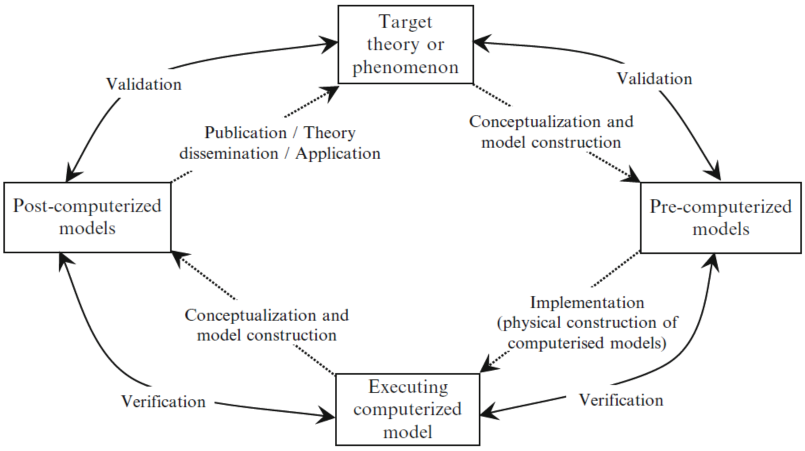


Figure 32 Verification and validation in relation to the model development process, source: David (2013)

5.6.1 Sensitivity analysis

By systematically varying the different parameters and inputs, their impact on the model's output can be observed, allowing the user to find out how "sensitive" these parameters and inputs are to changes in the parameter values (Maguire et al., 2005). Parameters with a high sensitivity needed to be checked an extra time in order to make sure that their values are appropriate and correct.

Testing and documenting the sensitivity of model output to changes in parameter values is important because (Railsback & Grimm, 2012):

1. It can show how strongly the model represents the real world phenomena
2. It helps us to understand the relative importance of model processes.

Here, the local sensitivity method is using, varying the most important parameters in the model with +/-5% (0.05*P*) (Railsback & Grimm, 2012). The lowest parameter value that is analysed is therefore *P-dP* = 0.95P and the highest value is *P+dP* = 1.05*P*. The sensitivity *S* is calculated as an approximation of the partial derivative of the currency with respect to the parameter (Railsback & Grimm, 2012). The sensitivity *S* is calculated by:

$$S+= \ (C^+ - C)/(dP/P)$$
$$S-= \ (C - C^-)/(dP/P)$$

With *P* reference parameter value, *C* currency value at *P* (here, average number of deforestation cells after 10 years), and C^+ and C^- the currency values for respectively *P+dP* and *P-dP*. For each P, *P+dP* and *P-dP*, the model was run 50 times. C, C^+ and C^- represent the average number of deforestation cells of these 50 runs for the given parameter values. The exact results can be found in tabular form in Appendix Q, the sensitivity values of S^- and S^+ per parameter are shown in Figure 33.

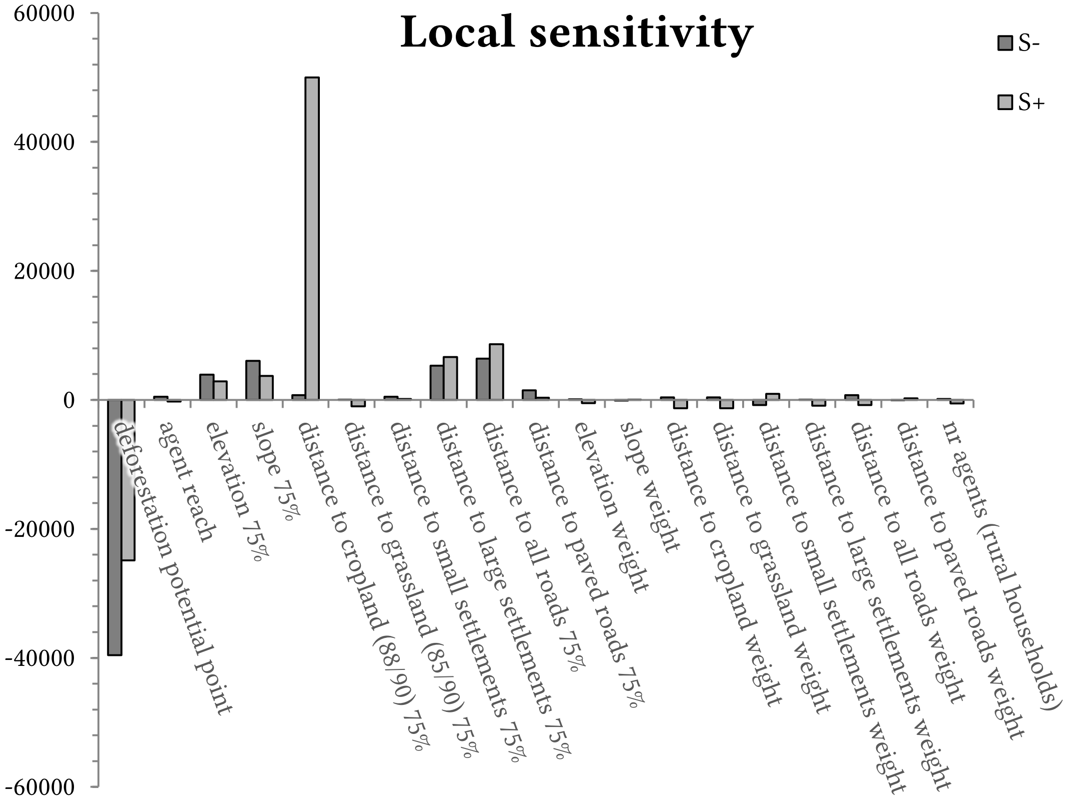


Figure 33 Results local sensitivity analysis

The results show a high sensitivity for the parameters deforestation-potential-point (75% value) and distance-to-cropland (75% value). An average sensitivity is found for the elevation (75% value), slope (75% value), distance-to-large-settlements (75% value) and distance-to-all-roads (75% value). Varying the weights of the factors with plus and minus 5% did not lead to significant sensitivity values. The same applies for changing the number of agents in the model with plus and minus 5%.

The distance-to-cropland parameter is highly sensitive to changes in the parameter value. This can be explained by the fact that this parameter had also the highest correlation with deforestation (see Chapter 4) and because many cells near forest are classified as cropland. Increasing the 75% value of this parameter would therefore considerably increase the number of forest cells in the model that receive a 17 point increase in the deforestation-potential value, thus leading to a substantial increase of cells “suitable” for deforestation.

The deforestation-potential-point is also highly sensitive. Lowering or increasing this parameter value (default 52) would have large effects on the projected deforestation rates. This value cannot be derived from empirical data directly. Furthermore, the third calibration test (Section 5.4) shows a range of values for this parameter that lead to a relatively good match with measured deforestation rates. Further analysis is therefore needed to determine the correct value of this parameter.

5.6.2 Qualitative deforestation pattern monitoring

As can be derived from the third calibration analysis (Section 5.4.1), the parameter values for agent-reach (1.5) and deforestation-potential-point (52) led to a modelled deforestation rate in the test runs that matched the amount of measured deforestation for 2001-2010 (based on the Landsat TM satellite data) almost perfectly. However, this check does not say much about the model's representation of the deforestation patterns, i.e. the distribution of the deforested patches over the area.

Therefore, a second step in the validation process looks at the criterion that the model simulation should generate results that are *qualitatively* comparable to those observed in the real world (Gilbert, 2008). This is done by comparing the measured deforestation of 2001-2010, based on the Landsat TM satellite data with the modelled deforestation for the same era.

The results of the overlay are presented in Figure 34. The definition of the legend items are as follows:

False positive – mild:	Up to 50% of the model runs showed deforestation, where there was no deforestation found in the land cover data.
False positive – strong:	More than 50% of the model runs showed deforestation, where there was no deforestation found in the land cover data.
False negative:	The model runs showed no deforestation, but there was deforestation found in the land cover data.
True positive - light:	Up to 50% of the model runs showed deforestation, where there was also actual deforestation according to the land cover data.
True positive – strong:	More than 50% of the model runs showed deforestation, where there was also actual deforestation according to the land cover data.

As can be derived from the map, only few modelled deforestation cells match the "actual" deforestation. Especially in the southern highlands, there are quite a number of false negatives and false positives. Only on micro level (see the zoom insets in Figure 34) the true positives become visible.

However, it must be noted that cells are already considered to be "false" (positive or negative) if the deforested cells are adjacent to each other, and thus not overlapping. In reality, if one is interested in general deforestation patterns at river basin level, and if a modelled deforestation cell then has a deviation of only a few hundred meters to the actual deforestation, then that can be considered a pretty good result. The results of Figure 34 may therefore show a too negative outcome of the comparison.

Hence, another test was done, calculating the distance of the modelled forest cells to the nearest measured deforestation cell. The results are shown in Figure 35. This method has a few downsides. False negatives will not be discovered which is why this method can only be presented in combination with the previous test. Furthermore, this method does not compare if the *number* of modelled deforestation cells on micro level are correct compared to the measured deforestation. If for example, ten modelled deforestation cells are relatively close to only one measured deforestation cell, all these ten modelled deforestation cells will be shown as a green patch on the map.

In general, this second pattern test shows that most of the cells are reasonably close to modelled deforestation. The largest XY deviation between modelled and measured deforestation was 2800 metres. The red spots in the North and North East of the VGTB area show the areas with the biggest deviation.

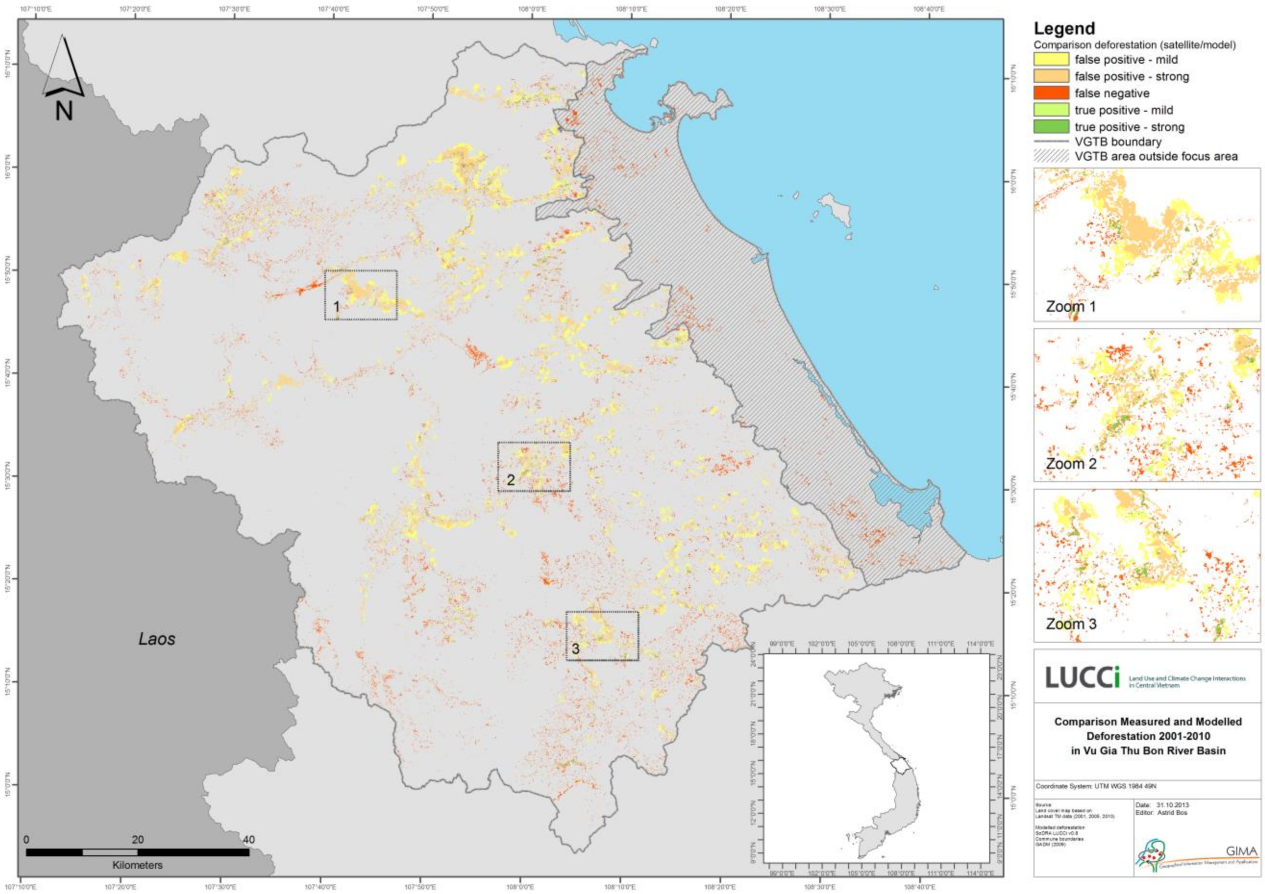


Figure 34 Comparison measured and modelled deforestation 2001-2010

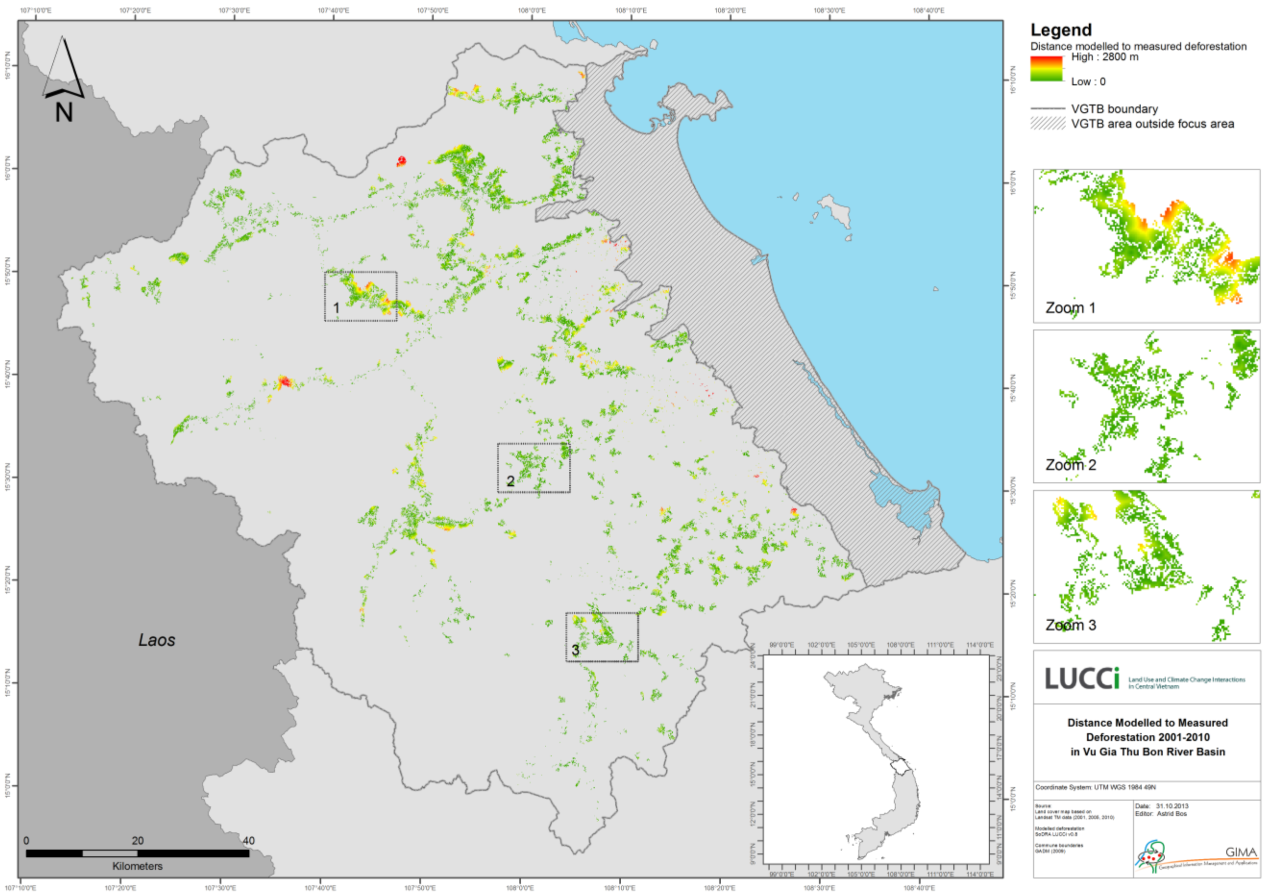


Figure 35 Distance modelled to measured deforestation

5.7 Uncertainty

Before moving to the discussion and conclusion, a few remarks regarding uncertainty are necessary. First, the uncertainty of a model is influenced by the parameters which are hard to verify. In particular, the agent-reach and the deforestation-potential-point parameter cannot be derived from empirical data. They have a close relationship to each other. A high agent-reach and high deforestation-potential-point may show similar (quantitative) results as model runs with a low agent-reach and low deforestation-potential-point (see also the results of the calibration tests in Section 5.4). The parameter deforestation-potential-point is especially uncertain, since it is based on assumptions that are hard to check. It assumes that all agents are fully aware of the degree of suitability for deforestation of forest patches in its neighbourhood. Furthermore it assumes that every agent has the same standards when defining the suitability of a forest patch for deforestation. The parameter is therefore rather fictitious. Consequently, it could be calibrated only in such a way that the general model outputs matched the measured deforestation rates of the land cover data. Because the deforestation-potential-point has a high uncertainty and is highly sensitive (see Section 5.6.1), close attention on this parameter value is needed.

Second, it is very much conceivable that the eight topographic and factors used in this research are not an exclusive list of factors influencing deforestation. Any missing factors increase the uncertainty of this model. Further research should not focus solely on the factors discussed in this research, but should have a wider scope instead. An example of a factor to be included in further analysis is the distance to existing plantation. Then, it can be analysed whether original forest area is cleared to start a forest plantation near existing plantations.

A third factor influencing the uncertainty regards the start location of households. There is uncertainty regarding the location of the houses and regarding the reach of agents. Here, the agents are randomly placed over the cells classified as urban, leading to approximately nine to ten agents per urban cell as start location. The false negatives from Figure 34 mostly represent deforested cells far from urban cells. Because in ten years, agents have a maximum reach of 10*1.5= 15 cells=1350m, in this model forest cells outside a radius of 1350m from urban cells cannot be deforested. Thus, to simulate the patterns of deforestation in remote areas, either the agent-reach should be increased, or the starting location of agents should not be limited to urban cells.

5.8 Discussion & conclusion

"In general, it is better to regard a model as a basis for reducing uncertainty about the future from some prior state of complete ignorance to one of more limited uncertainty, rather than to think of a model as failing if its predictions are not perfectly accurate" (Maguire et al., 2005, p. 14).

This chapter showed that it is possible to model future deforestation risk areas based on past land cover change data. However, there are some issues to consider when discussing the reliability and usability of the model.

NetLogo is a JAVA based modelling environment. Microsoft Windows limits the RAM that JAVA may use considerably. Therefore, the full SoDRA LUCCi model was coded, tested and run for demonstrations in the operating system Ubuntu Linux. However, processing model output in GIS software packages such as ArcGIS is limited in Ubuntu and was therefore done in Microsoft Windows. Switching between these two operating systems limits the user-friendliness of the model.

Moreover, since the focus area covers an area over 10,500 km^2, it was not feasible to model deforestation at 30m resolution. Therefore, the full SoDRA LUCCi model has a 90m resolution. Still, all spatial analysis was executed at 30m resolution, making it difficult to compare the results from the model with the measured deforestation for the 2001-2010 era.

Inherent to modelling is the simplification of factors and complex processes experienced in reality. In the SoDRA LUCCi model, some simplifications may have a considerable impact on the model output and should therefore be acknowledged. First, in and out migration and demographic developments is not considered, the number of agents in the model is stable. Second, the impact of socio-economic drivers of human behaviour and thus land change are underrepresented in the model. This is due to the fact that proper (accessible) socio-economic data to define agent behaviour was lacking. Rather than to add more ungrounded assumptions to the agent typology and their behaviour, it was decided to keep the agent behaviour very simple. Third, in this first version there are no loops or learning processes in the model. This means that if households do not find any forest in their range, they will not deforest at all. This may not be very realistic since it can be assumed that people who *want* to deforest a patch, first move to the (closest) forest border, and will deforest if the conditions permit so (not too steep etc.). Future versions of the model should pay special attention to the agent behaviour.

The parameterisation of the deforestation-potential parameter is based on the bivariate correlation analysis (weight factor) and binary 75% observed values threshold as presented in the previous chapter. For the factor elevation, the 75% threshold means that cells on or below 618m are suitable (based on that factor only) and all cells that have a value higher than 618 are not suitable. In practice, this may is most likely not realistic. Rather than dichotomous thresholds (suitable/unsuitable), it would be more realistic to include mathematical functions that match the measured values of the spatial analysis. For elevation, this would mean that areas with low elevations are most suitable. The suitability would gradually decrease with rising elevation.

A major assumption in the model is that future deforestation develops in similar ways as past deforestation. It assumes that past drivers and processes are the same drivers and processes involved in future deforestation. This is quite logical, since it makes more sense to base your model on things observed in reality rather than on rough estimates. Still, when using the SoDRA LUCCi model it is important to be aware of this assumption, in order to appreciate the model for its goal: providing insights in future deforestation risk areas and examining the effects of REDD measures.

In this research, special attention was paid to calibration and parameterisation methods. Since there is no data on past effects of REDD measures, the REDD scenario options cannot be calibrated. Consequently, they only represent what would happen *if* authorities successfully implement REDD measures that achieve the given goals (full protection etc.).

The calibration tests for the 2001-2010 era showed that, with a combination of the values 1.5 for agent-reach and 52 for deforestation-potential-point, the modelled deforestation rates are very similar compared to the measured deforestation based on the land cover data. The qualitative deforestation pattern monitoring showed however, that the representation of the deforestation patterns can still be improved considerably. Especially the small deforestation patches in the remote areas are underrepresented in the SoDRA LUCCi model, leading to false negatives. Still, the modelled deforestation was most of the time relatively close to the measured deforestation, with some exceptions in the north and northwest of the region.

For the 2010-2020 era, the results of the combined REDD scenario showed that deforestation rates can be reduced with over 45% when compared to the BAU scenario. This can be achieved under the following conditions:

- Authorities succeed in enforcing measures that lead to full protection of national parks;
- Full protection of high carbon stock areas and restricted deforestation in middle carbon stock areas is enforced;
- Implementation of a deforestation quota at a maximum of 5 cells (i.e. 4.05 ha) per 10 years per household.

Although these conditions may be impractical to pursue, the REDD scenarios do provide insights in their relative effectiveness. This may facilitate decision-making processes regarding the selection and implementation of particular REDD policies.

6. Discussion

This chapter aims to critically reflect upon the results presented in the previous chapters. Before moving on to the conclusions, which will be given in the next chapter, it is necessary to discuss those issues regarding the research context and research process that may have influenced the outcomes of the research.

6.1 Data availability, data quality and quality control

As discussed earlier in this thesis, the lack of available and/or accessible socio-economic data limited the validation of agent behaviour in the SoDRA LUCCi model. Data availability also played a role in the spatial analysis phase, since the factors chosen for correlation analysis were largely based on what data were available. Data on tenure systems, existing REDD measures, and the location of markets would already have broadened the scope of this research. Because of lacking information, the issue of hydropower dams was ignored in this research. Although the development of dams are not directly linked to the individual livelihood strategies, as they are implemented at the governmental level, hydropower dams play an important role in Vietnam and should therefore not be ignored. In Vietnam, and in particular in the VGTB, there are many dams or dam development plans (DONRE Tam Ky, 2010). For the construction of these dams and to make room for the dam reservoirs, often large areas of forest are cleared. Because data on the location of existing and planned dams was only available at district level, and thus the exact location was unknown, spatial analysis on the effects of these dams on deforestation was not possible. For this reason, the issue of dams was also not incorporated in the model.

Most of the vector data was provided by sources outside the LUCCi project. The quality of most vector data was poor. For example, network analysis on the road data was practically impossible because of lacking attributes, wrongly places roads (two lines representing the same road), and connection problems. In general, metadata was missing, making it difficult to control the quality of the data. With regards to the data on the administrative zones, boundaries were not always up to data since rearrangements of commune and district boundaries are very common in Vietnam. This hampered data interoperability, which will be discussed in more detail later in this chapter.

6.2 Spatial analysis related issues

Section 4.4 already dealt with some spatial analysis specific considerations. In addition, it must be noted that the spatial analysis over land cover change was based on satellite data, with a temporal resolution of five years. This makes is difficult to value land cover change and to reveal temporal deforestation. For example, the harvest of acacia plantations[9] may be detected as forest change from forest to grassland, but this deforestation may be just one part of a cycle, after which replanting starts. These kinds of cycles are not detected, and may have influenced the results of the land cover maps, which formed the backbone for the spatial analysis part of this thesis. Ideally, bottom-up field work observations or secondary empirical

[9] Which in general occurs every 7 years.

evidence should support the main assumptions which were based on top-down land cover satellite data analysis.

6.3 Modelling related issues

6.3.1 Limitations of the SoDRA LUCCi model

A model is inherently a simplified representation of reality. The SODRA LUCCI model can be considered a simple model, as it only takes into account a limited number of factors. Besides the additional factors proposed in Section 6.1, examples of factors that *may* influence land use decisions and deforestation, but are not included in the model are tenure systems, t patterns in agriculture, demographic factors such as migration, and the influence of people and forces from outside the modelled region. Regarding the latter, illegal deforestation from outside ones focus area is always hard to examine and thus hard to model. Still, this does not justify implicit exclusion from the list of potential deforestation drivers.

A final remark regarding the model encompasses the spatial resolution. For the full SoDRA LUCCi model to run on a standard desktop computer, a reduction of the cell resolution was needed. The aggregation of the data was based on the bilinear (for DEM) and nearest-neighbour method (for remaining layers). Every downsampling involves a simplification of the data and this may have affected the values used for analysis and modelling. The initial spatial analysis as presented in Chapter 4 was based on 30m resolution, whereas the full SoDRA LUCCi model as presented in Chapter 5 uses and operates at 30m resolution data. This may have caused some of the mismatches between modelled and measured deforestation (see Section 5.6.2) as the full SoDRA LUCCi model is unable to simulate single cell deforestation at 30m resolution.

6.3.2 Room for improvement

As the previous section already revealed, there is room for improvement of the model. Section 5.8 already discussed that the model can be improved by basing the deforestation-potential parameter on different mathematical functions per factor that fit the empirical observations on forest change as presented in the Chapter 4. Currently, the deforestation-potential parameter is calculated per factor on binary thresholds, i.e. given a certain value for factor *x*, the forest cell is suitable or not for deforestation.

At this stage, the SoDRA LUCCi model output consist of a land cover map with an extra class, i.e. deforested land. The model does not specify the future land cover class of the changed forest cell. However, by comparing the overlapping deforestation cells of measured and modelled deforestation for the 2001-2010 era, it can be examined which forest conversion type is represent best and worst by the SoDRA LUCCi model. Subsequently, these insights may be used to fine-tune the model.

6.4 Language issues & database interoperability

A frequently occurring, to overcome, but time-consuming problem regards the inability of many database processing software such as ArcGIS to deal with all characters involved in the Vietnamese language (Figure 36). Linking for example government documents with (spatial) databases based on a written definition of the spatial location (place names, districts, provinces etc.) is therefore a troublesome process.

Administrative codes often do not accord to each other when dealing with different data sources. Furthermore, continuing changes in commune and district boundaries and regularly occurring displacement of villages while lacking an openly accessible, up to date GI baseline system lead by the Vietnamese government makes it difficult for researchers to link data from different sources and different eras into a single GIS.

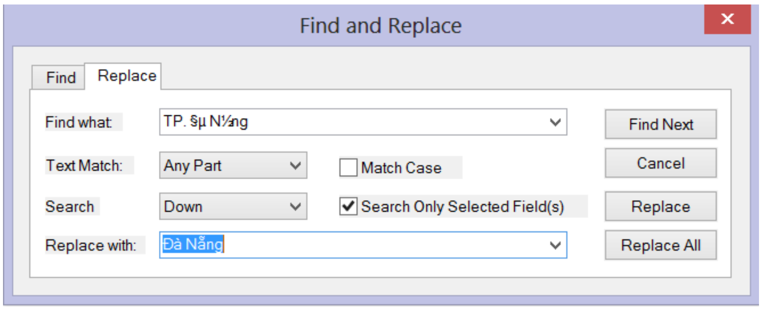


Figure 36 A common problem when working with Western and Vietnamese databases interchangeably

Furthermore, a general problem regards the unavailability of official documents such as laws and regulations in English. For many of the relevant documents, we were depending on the translation through fellow LUCCi project members, which is naturally a time-consuming process.

6.5 Final considerations

Some final remarks should be made before moving to the conclusions of this research in the next chapter.

First, a basic but crucial assumption that has been made is that the decision-making of human beings is always based on rational consideration. Modelling human behaviour is implicitly difficult, because decisions are often based on social constructs such as power and network relations, and humans barely base decisions on rational grounds solely. One should be aware of the possible implications of this when valuing model outputs of any ABM that considers human behaviour.

Second, as its name suggests, the LUCCi project focuses on climate change. In this research however, the focus on climate change is limited to the extent that it only considers climate change mitigation through avoiding deforestation. Vietnam and especially this central region are thought to be highly affected by climate change in the next coming decades (Ribbe, 2010). The model has not considered these changes. One might argue that people's behaviour may change drastically as a result of climate change, especially with regards to farming practices. These adaptive behaviour patterns should be considered in further research.

Third, ideally, this model facilitates decision-making processes with regards to REDD(+) and other deforestation prevention frameworks. However, as Matthews et al. (2007, p. 1447) acknowledge, past models have proven that agent-based land-use models are "probably more useful as research tools to develop an underlying knowledge base which can then be developed together with end-users into simple rules-of-thumb, rather than as operational decision support tools". As discussed earlier in this thesis, the scale of the study area in combination with the limitations of standard desktop computers, hampers an easy use of the SoDRA LUCCi model for decision-support purposes. Furthermore, if used in demonstrations, one should be very honest about the scope, particular focus and limitations of the SoDRA LUCCi model, since misinterpretation of the model output forms a considerable risk.

7. Conclusions

In this chapter the conclusions are presented which are based on the results presented in the previous chapters. Conclusions related to the specific spatial analysis and modelling phases can be found in respectively Section 4.4 and 5.8. If directly derived from a particular section in this thesis, to the right of the paragraph a cross-reference to the corresponding section is given.

Land use / land cover change in relation to forest change 4.2.1

An increase in the amount of agricultural land has put a pressure on existing forest land in Vietnam and also in the VGTB region in particular. Especially forest areas close to existing cropland are under particular deforestation risk. Distance to cropland was one of the main factors with a high negative correlation factor. This may indicate that most of the cleared forest areas will be converted to cropland in the future. Deforestation in the two eras 2001-2005 and 2005-2010 was further highly correlated to the distance to small settlements. This may indicate that areas with a higher population density suffer from greater deforestation risk compared to remote areas that are more difficult to access.

All eight topographic and thematic factors that were considered in this research showed, to a greater or lesser extent, a significant correlation with deforestation. These findings do not reveal cause-effect connections, but do give an insight in which factors are more or less related to forest change.

Areas under deforestation risk

The model results under the business-as-usual scenario show that deforestation for the 2010-2020 era is expected to be scattered all over the VGTB area. In particular, areas close to existing cropland and settlements in the north-east and central part of the VGTB region have a high deforestation risk. 5.5

Policy interventions

The SoDRA LUCCi experiments showed that the REDD measures with the largest effect on preventing projected deforestation are the introduction of a quota system in which households may only deforested a limited amount of forest patches per decade. Enforcing complete or partial prohibition of deforestation in existing protected areas or protecting areas with a relatively high carbon stock are expected however to have only limited effect on the projected deforestation. The patterns of deforestation do not differ considerably when compared to the business-as-usual scenario. The magnitude of deforestation however does change to a greater or lesser degree, depending on the scenario implemented. 5.5.2

As explained earlier, this research does not consider the means to implement the REDD scenarios. Still, it is important to make some notes about feasibility to achieve implement. Law enforcement in remote forest areas may be difficult, making full prohibition of deforestation or successfully implementing tree quota challenging. Furthermore, focussing REDD measures on the most important forest areas, i.e. the areas containing the largest carbon stock, may be a sound choice in theory, but it hard to put in practice as these rates may vary considerably from place to place.

When thinking about how to implement REDD measures that achieve the reductions in deforestation as presented in this thesis, one should not be focussing on restrictive measures exclusively. These reductions may well be achieved by other commonly used REDD measures that focus on provision of (financial) incentives, capacity building and technology transfer for stimulating (alternative) sustainable livelihood activities and strategies.

Limitations of the SoDRA LUCCi model

In addition to the limitations discussion in the corresponding chapter, some remarks regarding the limitations of the SoDRA LUCCi model should be made. 5.8

Since detailed socio-economic data about agent behaviour was missing, the model does not fully embody the complexity of the socio-ecological feedbacks that shape the land use/ land cover 5.2.1
change interactions.

In general, the SoDRA LUCCi model succeeds in representing the magnitude of deforestation when comparing the model results of the business-as-usual scenario of 2001-2010 to the 5.4.1
corresponding measured deforestation using the land cover data. The model can be improved with regards to simulating the deforestation patterns. Especially the small-scale deforestation patterns in the remote areas are underrepresented in the model results. 5.6.2

One of the factors with a high sensitivity also contains high uncertainty. This parameter, deforestation-potential-point, represents the point below which the location factors of a forest 5.7
patch are unsuitable for deforestation. Extra focus on this parameter is therefore needed.

To see the forest for the trees

Magliocca et al. (2013, p. 10) state*: "Certainly, over-simplifying the context in which land-use decision-making is embedded can lead to incomplete and/or incorrect understanding of the forces that shape land-use choices. On the other hand, representing the full complexity of social interactions that influence land-use choices runs counter to the aim of understanding larger-scale trends in land change; the impracticality of acquiring such detailed data across sites, coupled with the limitations of human cognition to navigate such complexity, is prohibitive".*

The SoDRA LUCCi model does not claim to epitomise the full complexity of reality on both micro and macro level. Still, in the author's opinion, studying large-trends in land change is implicitly linked to efforts focussing on understanding social interactions. The interconnectedness of spatial levels, meaning that micro-level actions trigger macro-level land change, is already visible from this first version of the LUCCi model. Next, it is key to comprehend these prevalent interconnections even better.

Agent-based modelling provides a tool for revealing large-scale patterns that are induced by micro-level actions. Rather than getting lost in a forest of details, it offers an instrument for greater understanding of the bigger picture while recognising that those details form the backbone of the system. To see the forest for the trees…

8. Recommendations

8.1 Recommendations for LUCCi and the Vietnamese authorities

This research only focussed on modelling the projected effects of certain REDD policies. How to successfully implement REDD measures is a different study. Lessons can be learned from other case studies, even from outside Vietnam, but it is advised to specify REDD measures at local circumstances. What works in one area may not work in other areas. For successful implementation of REDD measures, it is key that the measures address the drivers of deforestation in that particular area, otherwise, projected deforestation will not be reduced.

8.2 Recommendations for further research

One of the values of NetLogo is that it enables the user to study the bidirectional linkages between human behaviour and land cover change. In the SoDRA LUCCi model, for simplicity reasons, the only change a human can enforce is forest to non-forest change. These cells will then be classified as deforested. However, the new land cover class of that cell (e.g. cropland) is not defined in the model. The new status may however change the behaviour of the agents again. Further versions of the model may want to focus on these bidirectional linkages in more detail.

The influence of social networks has not been studied in this model. However, field work has shown that the influence of institutions such as villages and farmer cooperatives should not be underestimated. In communes just outside the VGTB area it was found that the behaviour of one or few promoters of rubber plantations had a significant influence on the behaviour of other farmers in the village. In other words, these promoters could be considered trendsetters in their village. These social processes need to be studied in more detail before they can be incorporated in the model.

References

Adley, J. & Grant, A. (2006). The Environmental Consequences of War Retrieved 16 May, 2013, from http://www.sierraclub.ca/national/postings/war-and-environment.html

Agarwal, C., Green, G. M., Grove, J. M., Evans, T. P. & Schweik, C. M. (2002). A review and assessment of land-use change models: Dynamics of space, time, and human choice. *A Review and Assessment of Land-use Change Models: Dynamics of Space, Time, and Human Choice.*

Axelrod, R. (1997). *The complexity of cooperation : agent-based models of competition and collaboration*. Princeton: Princeton University Press.

Baatz, M. & Schäpe, A. (2000). Multiresolution segmentation - An optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII*, 12-23.

Bakker, M. & Veldkamp, A. (2012). Changing relationships between land use and environmental characteristics and their consequences for spatially explicit land-use change prediction. *Journal of Land Use Science, 7*(4), 407-424.

Bakker, M. M. & Veldkamp, A. (2008). Modelling land change: The issue of use and cover in wide-scale applications. *Journal of Land Use Science, 3*(4), 203-213.

Berger, T. (2001). Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics, 25*(2-3), 245-260.

Bryman, A. (2008). *Social Research Methods* (Vol. 3). Oxford: Oxford University Press.

Carloni, A. S. & Crowley, E. (2005). Rapid guide for missions - Analysing local institutions and livelihoods *Institutions for Rural Development*. Rome: Food and Agriculture Organization of the United Nations (FAO).

Chambers, R. & Conway, G. R. (1991). Sustainable Rural Livelihoods: Practical Concepts for the 21st Century. *IDS Discussion Paper*(296).

David, N. (2013). Validating Simulations. In B. Edmonds, S. Moss & R. Meyer (Eds.), *Simulating Social Complexity: A Handbook* (pp. 135-172). Berlin: Springer.

Deadman, P., Robinson, D., Moran, E. & Brondizio, E. (2004). Colonist household decisionmaking and land-use change in the Amazon Rainforest: An agent-based simulation. *Environment and Planning B: Planning and Design, 31*(5), 693-709.

DONRE Tam Ky. (2010). Land Use Master Plan Quang Nam 2011-2020. Tam Ky: Department of Natural Resources and Environment.

FAO. (2009). State of the world's forests. Rome: Food and Agriculture Organization of the United Nations.

Field, A. (2009). *Discovering Statistics Using SPSS*. London: SAGE Publications.

Geist, H. J. & Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation. *BioScience, 52*(2), 143-150.

Gilbert, N. (2008). *Agent-Based Models*. Los Angeles, US: Sage.

Grimm, V., Polhill, G. & Touza, J. (2013). Documenting Social Simulation Models: The ODD Protocol as a Standard. In B. Edmonds, S. Moss & R. Meyer (Eds.), *Simulating Social Complexity: A Handbook* (pp. 117-134). Berlin: Springer.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., et al. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science, 310*(5750), 987-991.

Gros, C. (2013). *Complex and Adaptive Dynamical Systems: A Primer*: Springer.

Gross, D. & Strand, R. (2000). Can agent-based models assist decisions on large-scale practical problems? A philosophical analysis. *Complexity, 5*(6), 26-33.

Holland, J. H. (2006). Studying complex adaptive systems. *Journal of Systems Science and Complexity, 19*(1), 1-8.

Huong, H. L. & Viet, T. Q. (2009). Vu Gia Thu Bon River Basin. Cologne.

IPUMS. (2011). Microdata Vietnam. In GSO Vietnam (Ed.), *Integrated Public Use Microdata Series* (Vol. 6.1). Minneapolis: Minnesota Population Center, University of Minnesota.

Lambin, E. F., Geist, H. & Rindfuss, R. R. (2006). Introduction: Local Processes with Global Impacts. In E. F. Lambin & H. Geist (Eds.), *Land-Use and Land-Cover Change* (pp. 1-8). Berlin Heidelberg: Springer.

Lambin, E. F. & Meyfroidt, P. (2010). Land use transitions: Socio-ecological feedback versus socio-economic change. *Land Use Policy, 27*(2), 108-118.

Le Page, C., Bazile, D., Becu, N., Bommel, P., Bousquet, F., Etienne, M., et al. (2013). Agent-Based Modelling and Simulation Applied to Environmental Management. In B. Edmonds, S. Moss & R. Meyer (Eds.), *Simulating Social Complexity: A Handbook* (pp. 499-540). Berlin: Springer.

Le, Q. B., Park, S. J., Vlek, P. L. G. & Cremers, A. B. (2008). Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. I. Structure and theoretical specification. *Ecological Informatics, 3*(2), 135-153.

Ligtenberg, A. (2006). *Exploring the use of Multi-Agent Systems for Interactive Multi-Actor Spatial Planning.* (Ph.D. thesis), Wageningen University, Wageningen.

Magliocca, N. R., Brown, D. G. & Ellis, E. C. (2013). Exploring Agricultural Livelihood Transitions with an Agent-Based Virtual Laboratory: Global Forces to Local Decision-Making. *PLoS ONE, 8*(9).

Maguire, D. J., Batty, M. & Goodchild, M. F. (2005). *GIS, spatial analysis, and modeling*: ESRI Press.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G. & Gotts, N. M. (2007). Agent-based land-use models: A review of applications. *Landscape Ecology, 22*(10), 1447-1459.

McElwee, P. D. (2008). Forest environmental income in Vietnam: Household socioeconomic factors influencing forest use. *Environmental Conservation, 35*(2), 147-159.

Mertens, B. & Lambin, E. F. (1997). Spatial modelling of deforestation in southern Cameroon: Spatial disaggregation of diverse deforestation processes. *Applied Geography, 17*(2), 143-162.

Meth, P. & Williams, G. (2006). Literature Reviews and Bibliographic Searches. In V. Desai & R. B. Potter (Eds.), *Doing Development Research* (pp. 209-221). London: SAGE.

Minh, V. H. & Warfvinge, H. (2002). Issues in Management of Natural Forests by Households and Local Communities of Three Provinces in Vietnam: Hoa Binh, Nghe An, and Thua Thien-Hue *Working Paper Series* (Vol. 5). Santa Barbara, California USA: Asia Forest Network.

Munthali, K. G. (2012). Agent-Based Simulation in Geospatial Analysis. In Y. Murayama (Ed.), *Progress in Geospatial Analysis* (pp. 143-156). Tokyo: Springer Japan.

Parker, C., Mitchell, A., Trivedi, M., Mardas, N. & Sosis, K. T. (2009). The Little REDD+ Book. Oxford, UK: Global Canopy Programme.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J. & Deadman, P. (2003). Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. *Annals of the Association of American Geographers, 93*(2), 314-337.

Parker, D. C. & Meretsky, V. (2004). Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment, 101*(2-3), 233-250.

Pesticide Action Network North America. (1998). Agent Orange and Dioxin in Vietnam: New Findings. *Global Pesticide Campaigner, 08*(4). Retrieved from http://www.panna.org/legacy/gpc/gpc_199812.08.4.10.dv.html

Railsback, S. F. & Grimm, V. (2012). *Agent - Based and Individual - Based Modeling : A Practical Introduction.* Princeton: Princeton University Press.

Ribbe, L. (2010). LUCCi - Land Use and Climate Change Interactions in the Vu Gia Thu Bon River Basin, Central Vietnam: Research Proposal. Cologne: Institute for Technology and Resources Management in the Tropics and Subtropics, Cologne University of Applied Sciences.

Schultz, M. & Avitabile, V. (2012). 2001-2005-2010 Land Cover Change Analysis of the Vu Gia Thu Bon River Basin, Central Vietnam. Technical Report – Concepts, Methods and Results. Wageningen: LUCCi.

Scoones, I. (1998). Sustainable Rural Livelihoods: A Framework For Analysis *IDS Working Paper 72.* Brighton: University of Sussex.

Theobald, D. M., Stevens Jr, D. L., White, D., Urquhart, N. S., Olsen, A. R. & Norman, J. B. (2007). Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environmental Management, 40*(1), 134-146.

Valbuena, D., Verburg, P. H. & Bregt, A. K. (2008). A method to define a typology for agent-based analysis in regional land-use research. *Agriculture, Ecosystems and Environment, 128*(1-2), 27-36.

Valbuena, D., Verburg, P. H., Bregt, A. K. & Ligtenberg, A. (2010). An agent-based approach to model land-use change at a regional scale. *Landscape Ecology, 25*(2), 185-199.

Veldkamp, A. & Fresco, L. O. (1996). CLUE: A conceptual model to study the conversion of land use and its effects. *Ecological Modelling, 85*(2-3), 253-270.

Verburg, P. H. (2006). Simulating feedbacks in land use and land cover change models. *Landscape Ecology, 21*(8), 1171-1183.

Verburg, P. H., Kok, K., Pontius Jr., R. G. & Veldkamp, A. (2006). Modeling Land-Use and Land-Cover Change. In E. F. Lambin & H. Geist (Eds.), *Land-Use and Land-Cover Change* (pp. 117-135). Berlin Heidelberg: Springer.

Wood, C. H. & Porro, R. (2002). *Deforestation and land use in the Amazon.* Gainesville: University Press of Florida.

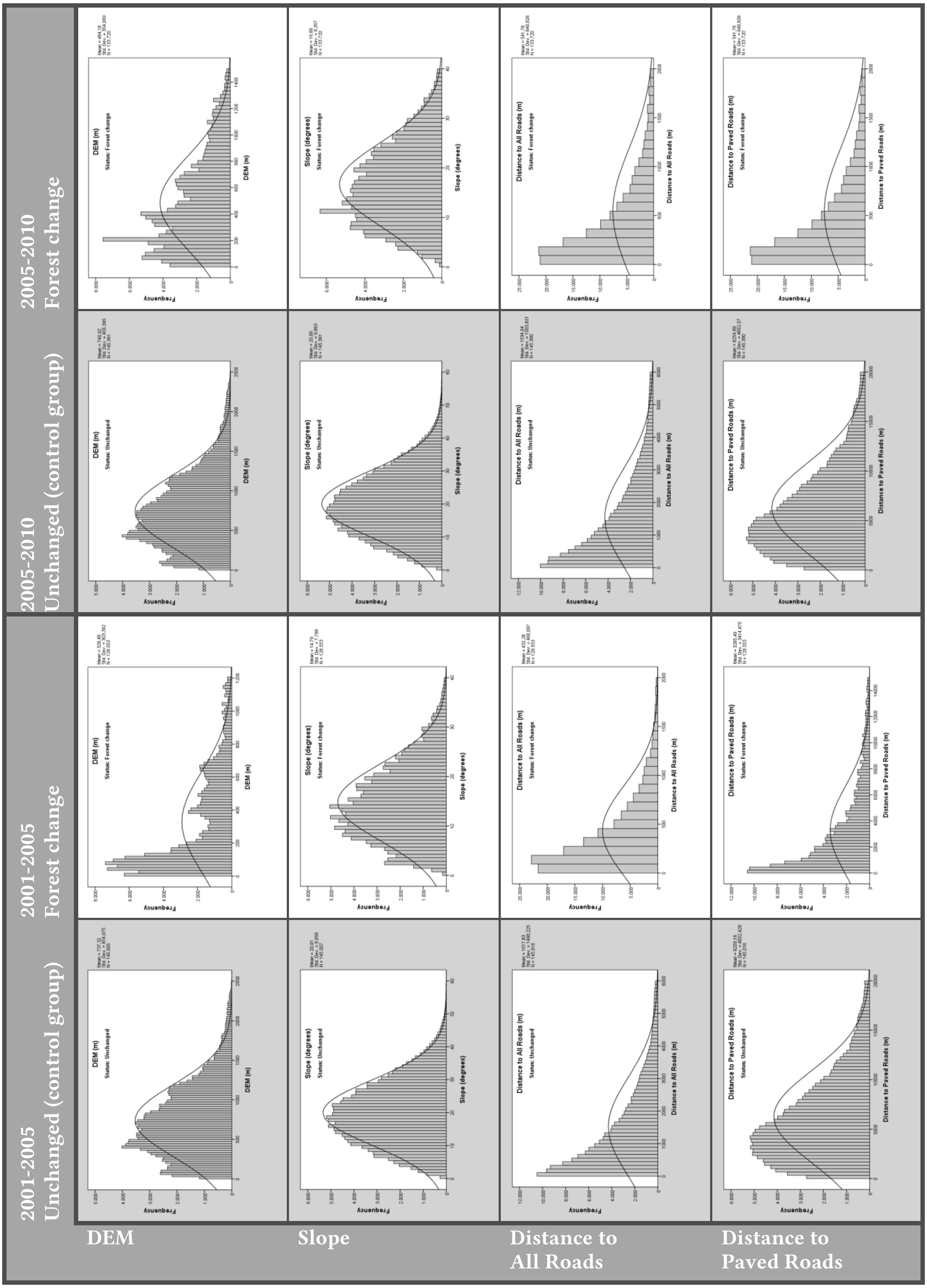
Appendices

Province	District code	District name	Commune	Commune code	#inhabitants total	# inhabitants urban	# inhabitants rural	rural ratio	average hh members in rural	# rural households
Đà Nẵng	497	Hoà Vang	Hòa Bắc	5011103	3486[1]			1.000[2]	4.38[2]	796
Đà Nẵng	497	Hoà Vang	Hòa Ninh	5011105	4188[1]			1.000[2]	4.38[2]	956
Đà Nẵng	497	Hoà Vang	Hòa Phú	5011127	4229[1]			1.000[2]	4.38[2]	966
Kon Tum	610	Đắk Glei	Đắk Choong	6010311	1433[1]			0.839[2]	4.92[2]	244
Kon Tum	610	Đắk Glei	Đắk Man	6010305	917[1]			0.839[2]	4.92[2]	156
Kon Tum	610	Đắk Glei	Đăk Plô	6010303	1003[1]			0.839[2]	4.92[2]	171
Kon Tum	610	Đắk Glei	Mường Hoong	6010313	2280[1]			0.839[2]	4.92[2]	389
Kon Tum	610	Đắk Glei	Ngọk Linh	6010315	2010[1]			0.839[2]	4.92[2]	343
Kon Tum	612	Tu Mơ Rông	Ngọk Yêu	6011115	1189[1]			0.720[2]	5.38[2]	159
Quảng Nam	502	Phú Ninh	Tam Đàn	5030131	9051[3]			0.743[2]	4.17[2]	1613
Quảng Nam	502	Phú Ninh	Tam Lãnh	5030139	6419[3]			0.743[2]	4.17[2]	1144
Quảng Nam	502	Phú Ninh	Tam Lộc	5030117	7506[3]			0.743[2]	4.17[2]	1337
Quảng Nam	502	Phú Ninh	Tam Vinh	5030129	8393[3]			0.743[2]	4.17[2]	1495
Quảng Nam	506	Đại Lộc	All	N/A	145935[4]	16215[4]	129720[4]	0.889	4.3[2]	30167
Quảng Nam	507	Điện Bàn	Điện Hồng	5030911	12261[3]			0.953[2]	4.33[2]	2699
Quảng Nam	508	Duy Xuyên	Duy Châu	5031111	6763[3]			0.812[2]	4.32[2]	1271
Quảng Nam	508	Duy Xuyên	Duy Hòa	5031109	8960[3]			0.812[2]	4.32[2]	1684
Quảng Nam	508	Duy Xuyên	Duy Phú	5031105	4151[3]			0.812[2]	4.32[2]	780
Quảng Nam	508	Duy Xuyên	Duy Sơn	5031115	9995[3]			0.812[2]	4.32[2]	1879
Quảng Nam	508	Duy Xuyên	Duy Tân	5031107	5146[3]			0.812[2]	4.32[2]	967
Quảng Nam	508	Duy Xuyên	Duy Thu	5031103	4272[3]			0.812[2]	4.32[2]	803
Quảng Nam	509	Quế Sơn	Quế An	5031723	4907[3]			0.894[2]	3.94[2]	1113

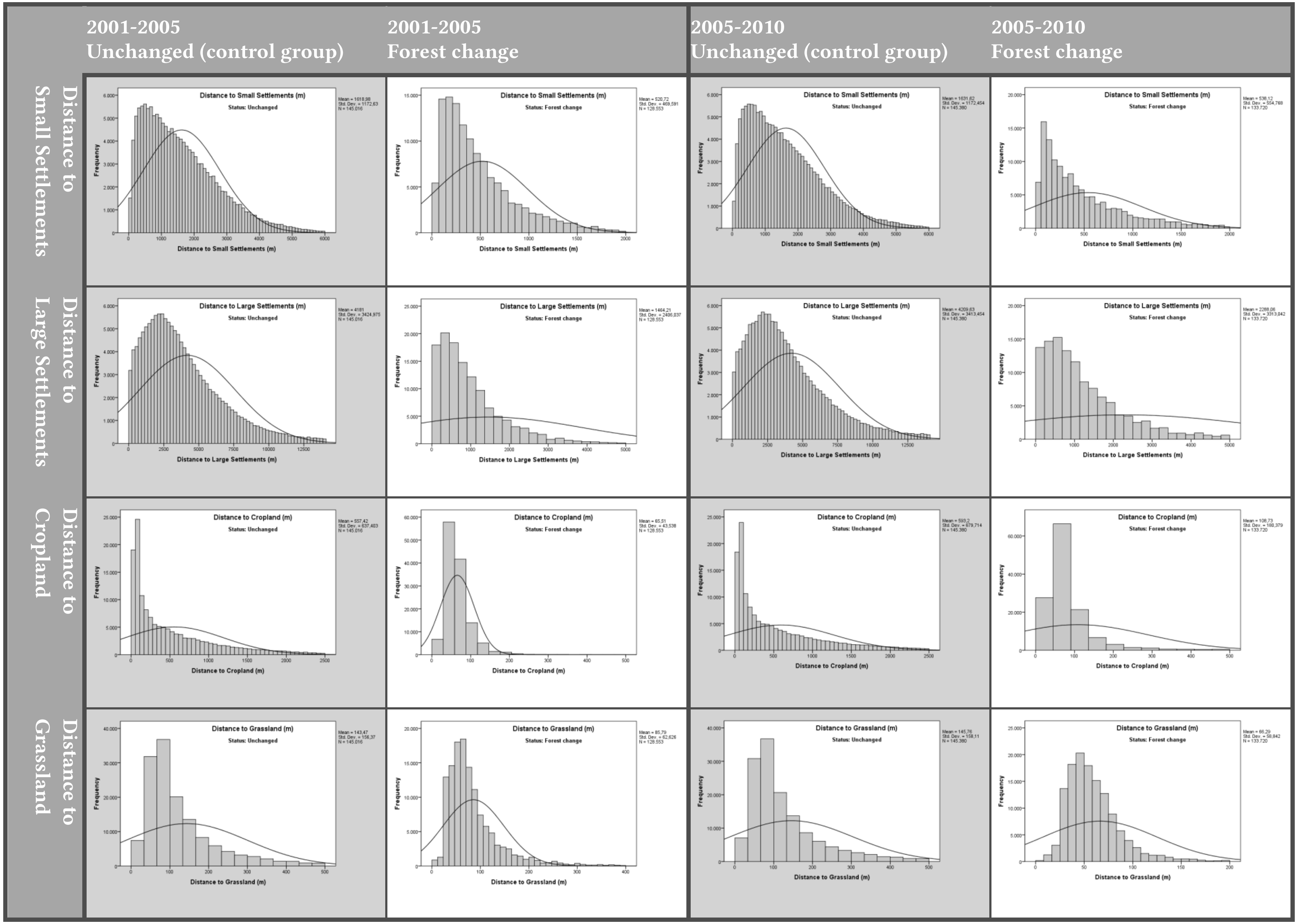
[1] Source: Socialist Republic of Viet Nam Government Portal http://gis.chinhphu.vn/
[2] Source: IPUMS Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.1. Minneapolis: University of Minnesota, 2011
[3] Source: Quang Nam Provincial General Statistical Office, 2009
[4] Source: Viet Nam General Statistical Office, 2009

Province	District code	District name	Commune	Commune code	#inhabitants total	# inhabitants urban	# inhabitants rural	rural ratio	average hh members in rural	# rural households
Quảng Nam	509	Quế Sơn	Quế Châu	5031727	6788[3]			0.894[2]	3.94[2]	1540
Quảng Nam	509	Quế Sơn	Quế Hiệp	5031717	3367[3]			0.894[2]	3.94[2]	764
Quảng Nam	519	Nông Sơn	Quế Lâm	5031705	3911[3]			1.000[2]	4.85[2]	806
Quảng Nam	519	Nông Sơn	Quế Lộc	5031711	5682[3]			1.000[2]	4.85[2]	1172
Quảng Nam	509	Quế Sơn	Quế Long	5031715	3892[3]			0.894[2]	3.94[2]	883
Quảng Nam	509	Quế Sơn	Quế Minh	5031725	4882[3]			0.894[2]	3.94[2]	1108
Quảng Nam	519	Nông Sơn	Quế Ninh	5031707	3606[3]			1.000[2]	4.85[2]	744
Quảng Nam	519	Nông Sơn	Quế Phước	5031703	2282[3]			1.000[2]	4.85[2]	471
Quảng Nam	509	Quế Sơn	Quế Phong	5031713	5511[3]			0.894[2]	3.94[2]	1250
Quảng Nam	509	Quế Sơn	Quế Thuận	5031729	6379[3]			0.894[2]	3.94[2]	1447
Quảng Nam	519	Nông Sơn	Quế Trung	5031709	10125[3]			1.000[2]	4.85[2]	2088
Quảng Nam	509	Quế Sơn	Thị Trấn Đông Phú	5031701	7933[3]			0.894[2]	3.94[2]	1800
Quảng Nam	510	Nam Giang	All	N/A	22417[4]	6822[4]	15595[4]	0.696	4.85[2]	3215
Quảng Nam	511	Phước Sơn	All	N/A	22586[4]	6220[4]	16366[4]	0.725	4.54[2]	3605
Quảng Nam	512	Hiệp Đức	All	N/A	38001[4]	3111[4]	34890[4]	0.918	4.7[2]	7423
Quảng Nam	513	Thăng Bình	Bình Định Bắc	5031521	4518[3]			0.910[2]	4.39[2]	937
Quảng Nam	513	Thăng Bình	Bình Định Nam	5031521	4444[3]			0.910[2]	4.39[2]	921
Quảng Nam	513	Thăng Bình	Bình Lãnh	5031517	6072[3]			0.910[2]	4.39[2]	1259
Quảng Nam	513	Thăng Bình	Bình Phú	5031525	3845[3]			0.910[2]	4.39[2]	797
Quảng Nam	513	Thăng Bình	Bình Trị	5031519	6432[3]			0.910[2]	4.39[2]	1333
Quảng Nam	514	Tiên Phước	All	N/A	68877[4]	6953[4]	61924[4]	0.899	4.76[2]	13009
Quảng Nam	517	Núi Thành	Tam Sơn	5032509	4017[3]			0.941[2]	4.02[2]	940
Quảng Nam	517	Núi Thành	Tam Trà	5032529	2772[3]			0.941[2]	4.02[2]	649
Quảng Nam	504	Tây Giang	All	N/A	16534[4]	0[4]	16534[4]	1.000	4.98[2]	3320
Quảng Nam	505	Đông Giang	All	N/A	23428[4]	4075[4]	19353[4]	0.826	4.39[2]	4408
Quảng Ngãi	525	Trà Bồng	Trà Tân	5050723	1471[1]			0.744[2]	4.59[2]	238
									TOTAL	**107262**

Source	Year(s)	Advantages	Disadvantages
Own survey Bach Ma National Park	2010	✓ Detailed on land use	✘ small sample ✘ (just) outside study area
Household Living Standard Survey (VARHS 2008)	2008	✓ Nation wide ✓ Info on land use and land ownership	✘ 91 respondents in 12 different communes
Vietnamese household census (access through IPUMS)	1989, 1999, 2009	✓ Nation wide ✓ Large sample in study area (≈860.000 hh)	✘ General survey on hh data (not on land use)
Rural, agricultural and fishery census	2001, 2006, 2011	✓ Nation wide ✓ (Expected) large sample in study area ✓ Variables available are very relevant for this study	✘ Currently, no access has been provided by GSO ✘ Only statistics at national level are currently available



(continues on next page)


2001-2005
Unchanged (control group)
2001-2005
Forest change
2005-2010
Unchanged (control group)
2005-2010
Forest change
Distance to
Small Settlements
Distance to
Large Settlements
Distance to
Cropland
Distance to
Grassland
Distance to Small Settlements (m)
Status: Unchanged
Status: Forest change
Distance to Large Settlements (m)
Distance to Cropland (m)
Distance to Grassland (m)
Frequency

Statistics

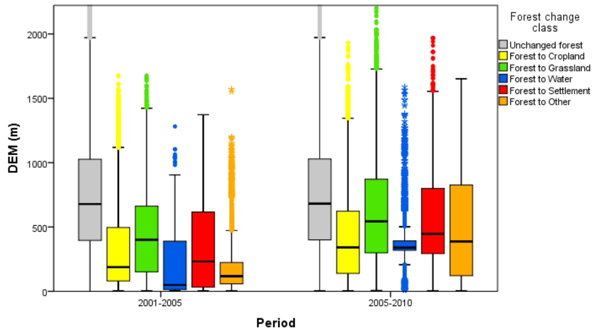
Status			DEM (m)	Slope (degrees)	Distance to All Roads (m)	Distance to Paved Roads (m)	Distance to Small Settlements (m)	Distance to Large Settlements (m)	Distance to Cropland (m)	Distance to Grassland (m)
Unchanged	N	Valid	145005	145007	145016	145016	145016	145016	145016	145016
		Missing	11	9	0	0	0	0	0	0
	Mean		737,32	20,01	1517,83	6259,16	1618,98	4181,00	557,42	143,47
	Std. Error of Mean		1,195	,026	3,934	12,217	3,079	8,994	1,674	,411
	Median		678,00	19,32	1056,83	5220,52	1380,00	3327,51	308,87	94,87
	Mode		417	11	54	54	150	75	30	67
	Std. Deviation		454,975	9,858	1498,225	4652,429	1172,630	3424,975	637,403	156,370
	Variance		207002,550	97,171	2244679,535	21645097,71	1375061,281	11730454,05	406282,061	24451,699
	Skewness		,723	,480	1,795	1,313	1,103	1,704	1,844	4,780
	Std. Error of Skewness		,006	,006	,006	,006	,006	,006	,006	,006
	Kurtosis		,347	,409	3,979	2,137	1,362	3,729	3,913	38,476
	Std. Error of Kurtosis		,013	,013	,013	,013	,013	,013	,013	,013
	Range		2554	88	10140	28694	8403	22102	4512	2778
	Minimum		0	0	18	15	25	21	12	12
	Maximum		2554	88	10158	28709	8429	22123	4524	2790
	Percentiles	25	396,00	12,65	424,26	2834,00	697,78	1799,56	90,00	61,85
		50	678,00	19,32	1056,83	5220,52	1380,00	3327,51	308,87	94,87
		75	1026,00	26,57	2130,69	8571,27	2280,20	5553,55	806,78	161,55
Forest change	N	Valid	128553	128553	128553	128553	128553	128553	128553	128553
		Missing	0	0	0	0	0	0	0	0
	Mean		328,48	14,79	432,28	3285,49	520,72	1464,21	65,51	85,79
	Std. Error of Mean		,846	,022	1,300	9,523	1,310	6,711	,121	,175
	Median		207,00	13,88	286,57	2120,71	382,51	795,37	58,52	69,01
	Mode		29	8	30	1368	67	30	30	30
	Std. Deviation		303,392	7,799	466,097	3414,475	469,591	2406,037	43,538	62,626
	Variance		92046,670	60,819	217246,412	11658640,04	220515,540	5789014,878	1895,526	3922,060
	Skewness		1,100	,565	3,445	1,612	2,196	3,946	13,917	3,176
	Std. Error of Skewness		,007	,007	,007	,007	,007	,007	,007	,007
	Kurtosis		,592	,021	23,472	2,460	7,142	16,258	527,216	17,988
	Std. Error of Kurtosis		,014	,014	,014	,014	,014	,014	,014	,014
	Range		1677	62	7593	20703	5950	18357	2640	1001
	Minimum		3	0	15	24	7	0	0	0
	Maximum		1680	62	7608	20727	5958	18357	2640	1001
	Percentiles	25	84,50	8,92	127,72	757,58	201,25	387,47	42,43	47,85
		50	207,00	13,88	286,57	2120,71	382,51	795,37	58,52	69,01
		75	518,88	19,77	579,59	4570,95	686,33	1480,79	78,82	100,50

Appendix E Frequency statistics 2005-2010

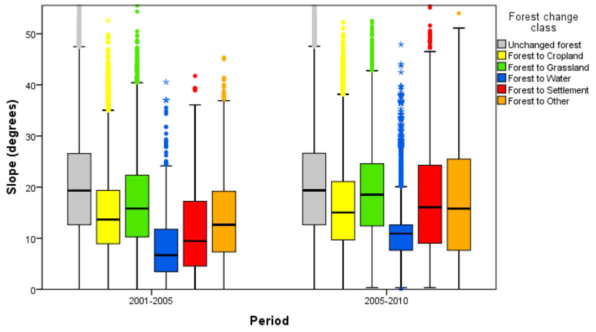
Statistics

Status			DEM (m)	Slope (degrees)	Distance to All Roads (m)	Distance to Paved Roads (m)	Distance to Small Settlements (m)	Distance to Large Settlements (m)	Distance to Cropland (m)	Distance to Grassland (m)
Unchanged	N	Valid	145361	145361	145380	145380	145380	145380	145380	145380
		Missing	19	19	0	0	0	0	0	0
	Mean		740,92	20,06	1534,04	6254,89	1631,62	4209,63	593,20	145,76
	Std. Error of Mean		1,194	,026	3,944	12,202	3,075	8,952	1,783	,415
	Median		682,00	19,36	1075,51	5209,67	1388,13	3359,61	330,00	96,05
	Mode		428	11	54	54	150	67	30	67
	Std. Deviation		455,395	9,863	1503,821	4652,570	1172,454	3413,454	679,714	158,110
	Variance		207384,314	97,283	2261478,511	21646406,64	1374648,094	11651671,05	462010,716	24998,803
	Skewness		,722	,484	1,788	1,327	1,103	1,694	1,800	4,743
	Std. Error of Skewness		,006	,006	,006	,006	,006	,006	,006	,006
	Kurtosis		,343	,431	3,944	2,188	1,357	3,689	3,503	37,863
	Std. Error of Kurtosis		,013	,013	,013	,013	,013	,013	,013	,013
	Range		2573	88	10210	28770	8421	22067	4512	2812
	Minimum		4	0	12	18	18	21	6	11
	Maximum		2577	88	10222	28788	8439	22088	4518	2823
	Percentiles	25	400,00	12,66	435,39	2850,04	711,20	1840,87	93,34	64,03
		50	682,00	19,36	1075,51	5209,67	1388,13	3359,61	330,00	96,05
		75	1029,00	26,62	2153,17	8545,47	2292,60	5584,70	855,42	167,71
Forest change	N	Valid	133720	133720	133720	133720	133720	133720	133720	133720
		Missing	0	0	0	0	0	0	0	0
	Mean		484,18	16,66	541,78	5675,52	538,12	2288,06	108,73	66,29
	Std. Error of Mean		,971	,023	1,752	13,207	1,517	9,060	,493	,161
	Median		400,00	15,74	330,34	4318,83	355,87	1099,33	68,72	54,08
	Mode		341	11	30	3041	30	2854	30	30
	Std. Deviation		354,959	8,357	640,826	4829,525	554,768	3313,042	180,379	58,842
	Variance		125995,789	69,840	410658,554	23324311,45	307767,207	10976244,58	32536,584	3462,348
	Skewness		,881	,535	3,019	1,031	2,212	2,838	7,747	5,343
	Std. Error of Skewness		,007	,007	,007	,007	,007	,007	,007	,007
	Kurtosis		,268	-,072	14,331	,795	6,697	8,739	82,731	35,951
	Std. Error of Kurtosis		,013	,013	,013	,013	,013	,013	,013	,013
	Range		2355	67	8975	28388	5476	22148	3670	658
	Minimum		4	0	18	21	8	15	0	0
	Maximum		2359	67	8993	28409	5484	22164	3670	658
	Percentiles	25	204,80	10,19	141,51	1661,94	152,97	523,83	49,84	40,31
		50	400,00	15,74	330,34	4318,83	355,87	1099,33	68,72	54,08
		75	693,00	22,25	693,16	8964,84	720,22	2284,43	98,62	73,18

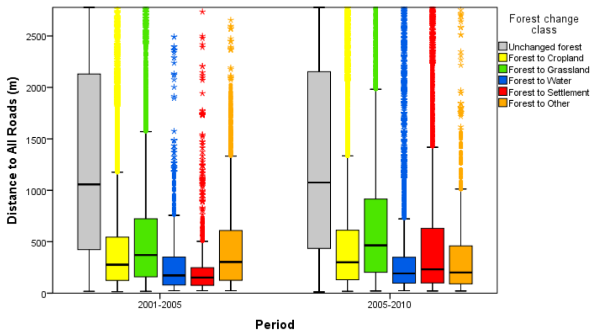
DEM



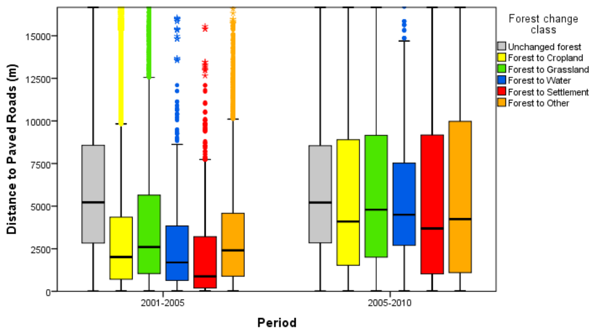
Slope



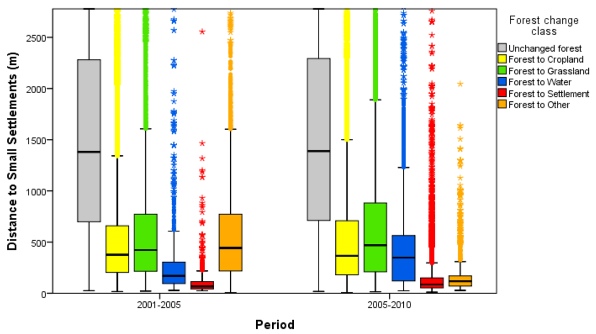
Distance to All Roads



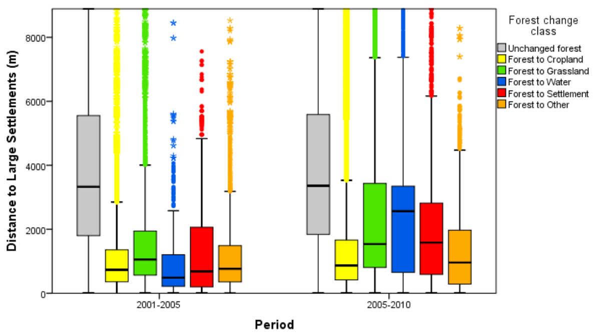
Distance to Paved Roads



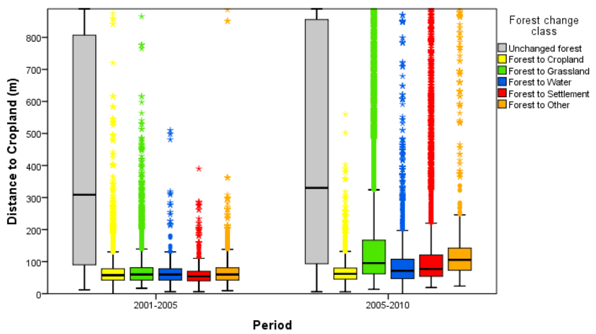
Distance to Small Settlements



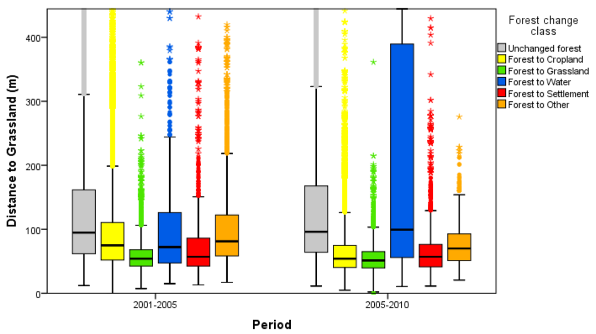
Distance to Large settlements



Distance to Cropland



Distance to Grassland



Group Statistics

Period		Status	N	Mean	Std. Deviation	Std. Error Mean
2001-2005	DEM (m)	Unchanged	145005	737,32	454,975	1,195
		Forest change	128553	328,48	303,392	,846
	Slope (degrees)	Unchanged	145007	20,01	9,858	,026
		Forest change	128553	14,79	7,799	,022
	Distance to All Roads (m)	Unchanged	145016	1517,83	1498,225	3,934
		Forest change	128553	432,28	466,097	1,300
	Distance to Paved Roads (m)	Unchanged	145016	6259,16	4652,429	12,217
		Forest change	128553	3285,49	3414,475	9,523
	Distance to Small Settlements (m)	Unchanged	145016	1618,98	1172,630	3,079
		Forest change	128553	520,72	469,591	1,310
	Distance to Large Settlements (m)	Unchanged	145016	4181,00	3424,975	8,994
		Forest change	128553	1464,21	2406,037	6,711
	Distance to Cropland (m)	Unchanged	145016	557,42	637,403	1,674
		Forest change	128553	65,51	43,538	,121
	Distance to Grassland (m)	Unchanged	145016	143,47	156,370	,411
		Forest change	128553	85,79	62,626	,175
2005-2010	DEM (m)	Unchanged	145361	740,92	455,395	1,194
		Forest change	133720	484,18	354,959	,971
	Slope (degrees)	Unchanged	145361	20,06	9,863	,026
		Forest change	133720	16,66	8,357	,023
	Distance to All Roads (m)	Unchanged	145380	1534,04	1503,821	3,944
		Forest change	133720	541,78	640,826	1,752
	Distance to Paved Roads (m)	Unchanged	145380	6254,89	4652,570	12,202
		Forest change	133720	5675,52	4829,525	13,207
	Distance to Small Settlements (m)	Unchanged	145380	1631,62	1172,454	3,075
		Forest change	133720	538,12	554,768	1,517
	Distance to Large Settlements (m)	Unchanged	145380	4209,63	3413,454	8,952
		Forest change	133720	2288,06	3313,042	9,060
	Distance to Cropland (m)	Unchanged	145380	593,20	679,714	1,783
		Forest change	133720	108,73	180,379	,493
	Distance to Grassland (m)	Unchanged	145380	145,76	158,110	,415
		Forest change	133720	66,29	58,842	,161

(continues on next page)

Independent Samples Test

Period			Levene's Test for Equality of Variances		t-test for Equality of Means						
										95% Confidence Interval of the Difference	
			F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
2001-2005	DEM (m)	Equal variances assumed	16635,909	,000	272,861	273556	,000	408,840	1,498	405,903	411,776
		Equal variances not assumed			279,244	254674,100	,000	408,840	1,464	405,970	411,709
	Slope (degrees)	Equal variances assumed	6750,853	,000	152,265	273558	,000	5,220	,034	5,153	5,287
		Equal variances not assumed			154,387	270145,985	,000	5,220	,034	5,154	5,286
	Distance to All Roads (m)	Equal variances assumed	75096,327	,000	249,309	273567	,000	1085,546	4,354	1077,012	1094,080
		Equal variances not assumed			261,986	176041,119	,000	1085,546	4,144	1077,425	1093,667
	Distance to Paved Roads (m)	Equal variances assumed	9454,709	,000	188,536	273567	,000	2973,671	15,772	2942,757	3004,585
		Equal variances not assumed			191,969	264585,723	,000	2973,671	15,490	2943,310	3004,032
	Distance to Small Settlements (m)	Equal variances assumed	74101,114	,000	314,211	273567	,000	1098,261	3,495	1091,410	1105,112
		Equal variances not assumed			328,205	195028,736	,000	1098,261	3,346	1091,702	1104,819
	Distance to Large Settlements (m)	Equal variances assumed	22645,464	,000	237,213	273567	,000	2716,786	11,453	2694,339	2739,234
		Equal variances not assumed			242,105	260385,529	,000	2716,786	11,222	2694,792	2738,780
	Distance to Cropland (m)	Equal variances assumed	150993,748	,000	276,134	273567	,000	491,913	1,781	488,422	495,405
		Equal variances not assumed			293,118	146540,874	,000	491,913	1,678	488,624	495,203
	Distance to Grassland (m)	Equal variances assumed	18821,808	,000	123,741	273567	,000	57,676	,466	56,763	58,590
		Equal variances not assumed			129,252	195038,032	,000	57,676	,446	56,802	58,551
2005-2010	DEM (m)	Equal variances assumed	6845,048	,000	165,121	279079	,000	256,744	1,555	253,697	259,792
		Equal variances not assumed			166,811	271861,969	,000	256,744	1,539	253,728	259,761
	Slope (degrees)	Equal variances assumed	3158,255	,000	97,751	279079	,000	3,397	,035	3,329	3,466
		Equal variances not assumed			98,423	277228,487	,000	3,397	,035	3,330	3,465
	Distance to All Roads (m)	Equal variances assumed	53061,369	,000	223,351	279098	,000	992,263	4,443	983,556	1000,971
		Equal variances not assumed			229,911	199973,687	,000	992,263	4,316	983,804	1000,722
	Distance to Paved Roads (m)	Equal variances assumed	1385,776	,000	32,271	279098	,000	579,371	17,953	544,183	614,559
		Equal variances not assumed			32,221	275080,653	,000	579,371	17,981	544,129	614,614
	Distance to Small Settlements (m)	Equal variances assumed	57842,833	,000	310,570	279098	,000	1093,500	3,521	1086,599	1100,401
		Equal variances not assumed			318,910	211163,906	,000	1093,500	3,429	1086,780	1100,221
	Distance to Large Settlements (m)	Equal variances assumed	2032,606	,000	150,678	279098	,000	1921,568	12,753	1896,573	1946,564
		Equal variances not assumed			150,866	278293,479	,000	1921,568	12,737	1896,604	1946,533
	Distance to Cropland (m)	Equal variances assumed	115555,975	,000	252,587	279098	,000	484,472	1,918	480,712	488,231
		Equal variances not assumed			261,924	167425,955	,000	484,472	1,850	480,846	488,097
	Distance to Grassland (m)	Equal variances assumed	29503,744	,000	173,096	279098	,000	79,467	,459	78,567	80,367
		Equal variances not assumed			178,657	187826,798	,000	79,467	,445	78,595	80,339

Appendix H Bivariate correlation test (Spearman's rho)

Correlations

Period				Status	DEM (m)	Slope (degrees)	Distance to All Roads (m)	Distance to Paved Roads (m)	Distance to Small Settlements (m)	Distance to Large Settlements (m)	Distance to Cropland (m)	Distance to Grassland (m)
Spearman's rho	2001-2005	Status	Correlation Coefficient	1,000	-,483**	-,271**	-,483**	-,389**	-,568**	-,576**	-,595**	-,256**
			Sig. (1-tailed)		0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
			N	273569	273558	273560	273569	273569	273569	273569	273569	273569
	2005-2010	Status	Correlation Coefficient	1,000	-,297**	-,174**	-,433**	-,091**	-,568**	-,432**	-,509**	-,462**
			Sig. (1-tailed)		0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
			N	279100	279081	279081	279100	279100	279100	279100	279100	279100

**. Correlation is significant at the 0.01 level (1-tailed).

Classification Table[a,b]

Period		Observed		Predicted: Status: Unchanged	Predicted: Status: Forest change	Predicted: Percentage Correct
2001-2005	Step 0	Status	Unchanged	145005	0	100,0
			Forest change	128553	0	,0
		Overall Percentage				53,0
2005-2010	Step 0	Status	Unchanged	145361	0	100,0
			Forest change	133720	0	,0
		Overall Percentage				52,1

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

Period			B	S.E.	Wald	df	Sig.	Exp(B)
2001-2005	Step 0[a]	Constant	-,120	,004	988,242	1	,000	,887
2005-2010	Step 0[a]	Constant	-,083	,004	485,286	1	,000	,920

a. Variable(s) entered on step 1: dem, Slope, RoadsAll, RoadsPav, SetSmall, SetLarg, Crop, Grass.

Model Summary

Period	Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
2001-2005	1	195201,103[a]	,488	,651
2005-2010	1	234723,315[b]	,419	,559

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than ,001 for split file Period = 2001-2005.

b. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001 for split file Period = 2005-2010.

Hosmer and Lemeshow Test

Period	Step	Chi-square	df	Sig.
2001-2005	1	267770,745	8	,000
2005-2010	1	8645,490	8	,000

Classification Table[a]

Period		Observed		Predicted: Status: Unchanged	Predicted: Status: Forest change	Predicted: Percentage Correct
2001-2005	Step 1	Status	Unchanged	112877	32128	77,8
			Forest change	10444	118109	91,9
		Overall Percentage				84,4
2005-2010	Step 1	Status	Unchanged	109909	35452	75,6
			Forest change	16322	117398	87,8
		Overall Percentage				81,4

a. The cut value is ,500

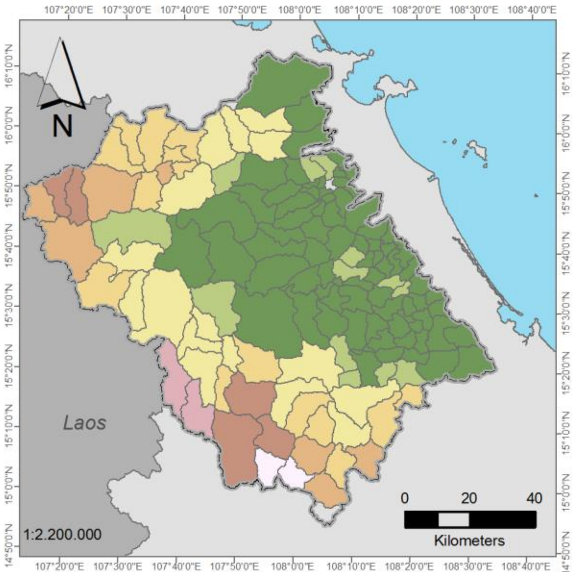
(continues on next page)

Variables in the Equation

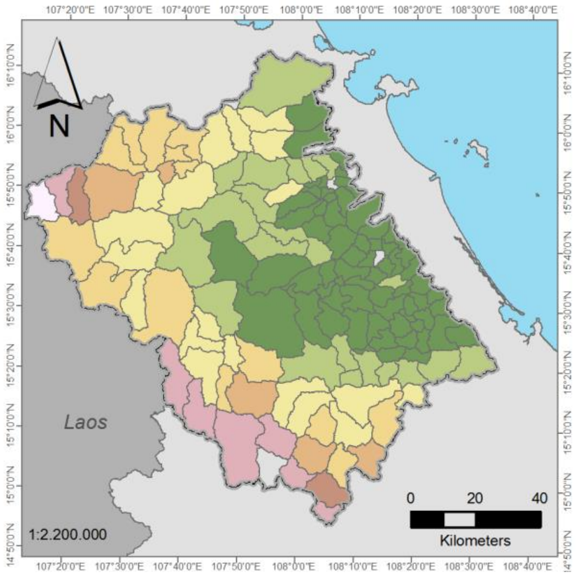
Period			B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	95% C.I.for EXP(B) Upper
2001-2005	Step 1[a]	dem	,000	,000	173,910	1	,000	1,000	1,000	1,000
		Slope	-,026	,001	1354,142	1	,000	,974	,973	,976
		RoadsAll	,000	,000	2170,463	1	,000	1,000	1,000	1,000
		RoadsPav	,000	,000	1471,102	1	,000	1,000	1,000	1,000
		SetSmall	-,001	,000	8210,190	1	,000	,999	,999	,999
		SetLarg	,000	,000	294,915	1	,000	1,000	1,000	1,000
		Crop	-,012	,000	14248,234	1	,000	,988	,988	,988
		Grass	-,004	,000	2684,453	1	,000	,996	,995	,996
		Constant	3,868	,020	39248,026	1	,000	47,845		
2005-2010	Step 1[a]	dem	,000	,000	216,202	1	,000	1,000	1,000	1,000
		Slope	-,024	,001	1560,724	1	,000	,976	,975	,977
		RoadsAll	,000	,000	3474,634	1	,000	1,000	1,000	1,000
		RoadsPav	,000	,000	2604,229	1	,000	1,000	1,000	1,000
		SetSmall	-,001	,000	22661,525	1	,000	,999	,999	,999
		SetLarg	,000	,000	610,215	1	,000	1,000	1,000	1,000
		Crop	-,003	,000	7372,372	1	,000	,997	,997	,997
		Grass	-,010	,000	10242,582	1	,000	,990	,990	,990
		Constant	2,956	,016	32161,083	1	,000	19,223		

a. Variable(s) entered on step 1: dem, Slope, RoadsAll, RoadsPav, SetSmall, SetLarg, Crop, Grass.

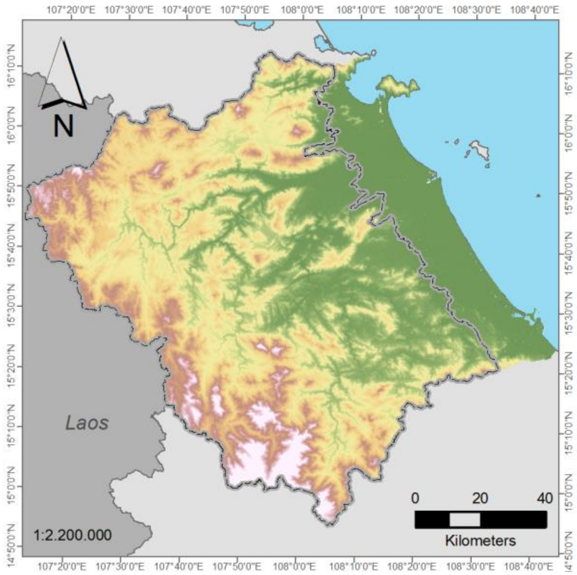
Average DEM for forest change cells 2001-2005

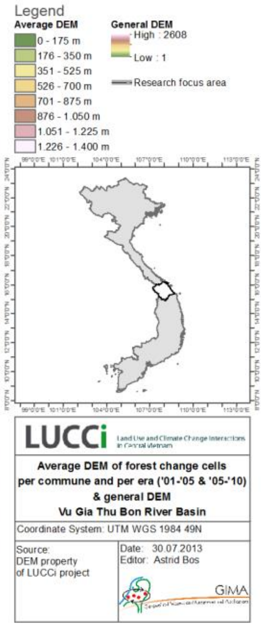


Average DEM for forest change cells 2005-2010

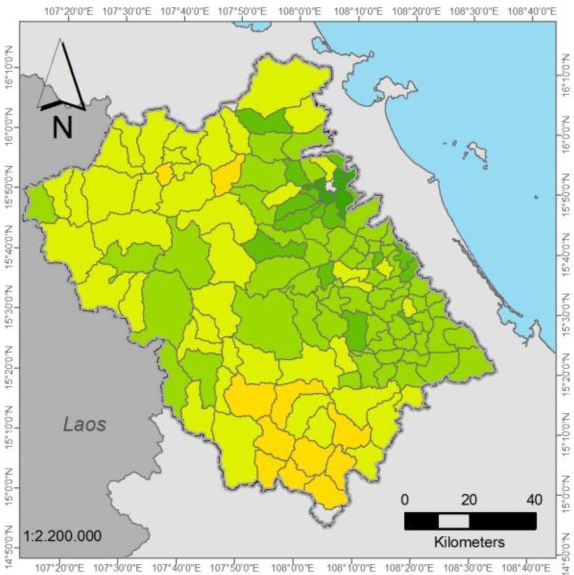


DEM

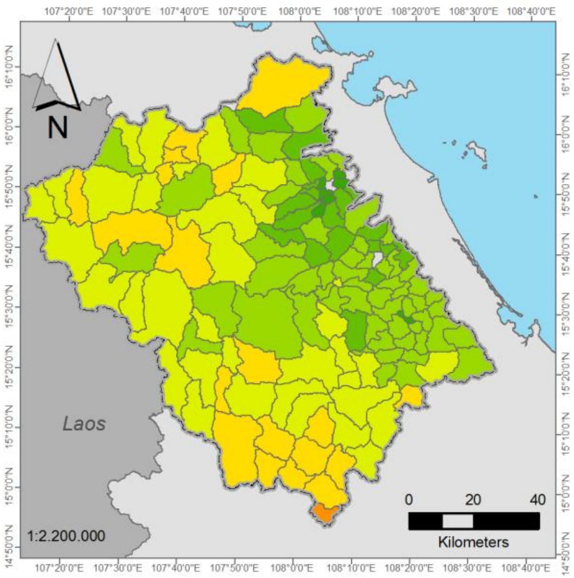




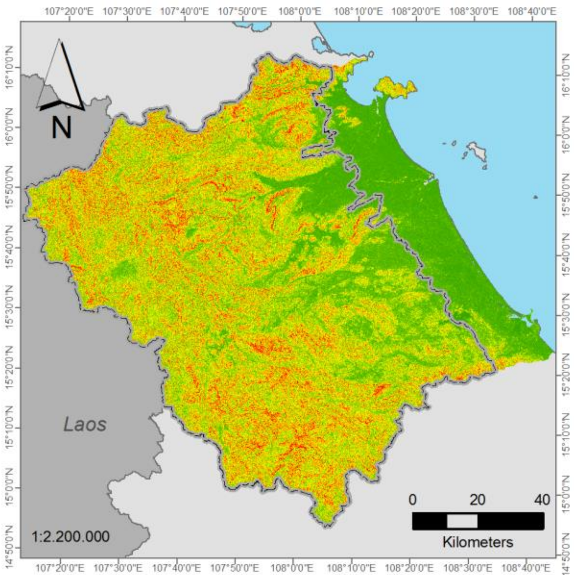
Average Slope for forest change cells 2001-2005

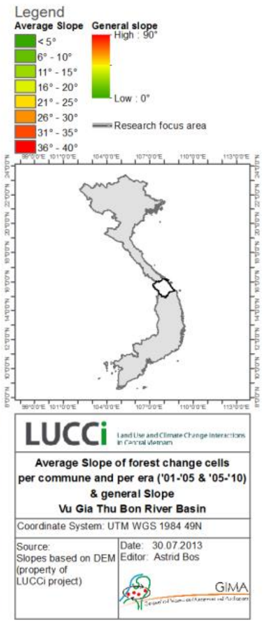


Average Slope for forest change cells 2005-2010

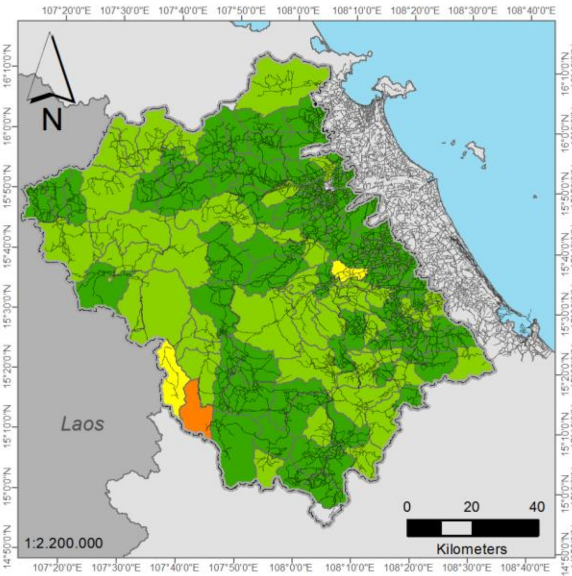


Slope

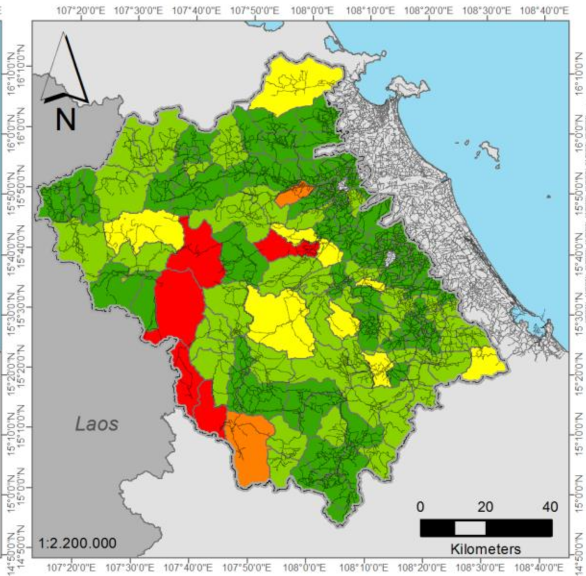


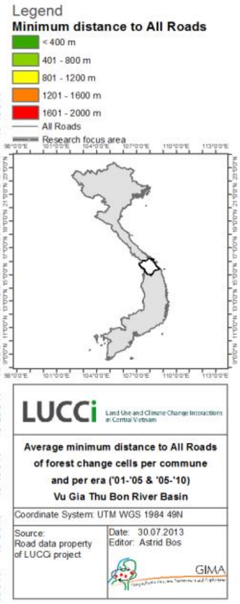


Distance to All Roads 2001-2005

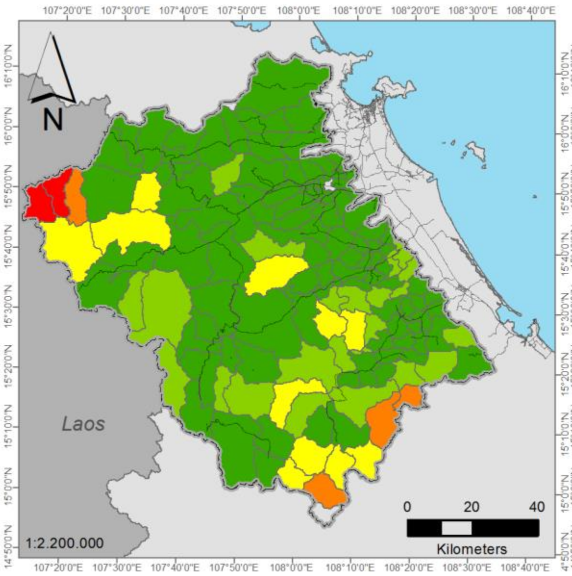


Distance to All Roads 2005-2010

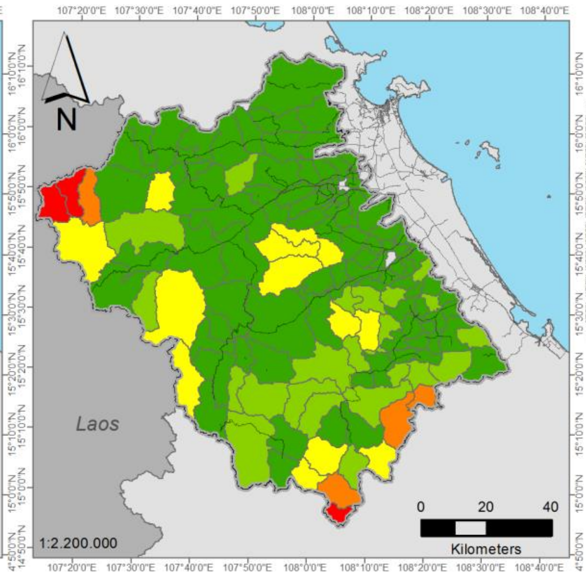


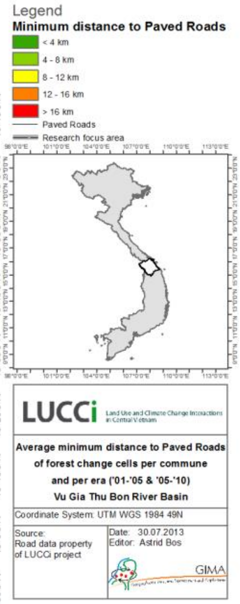


Distance to Paved Roads 2001-2005

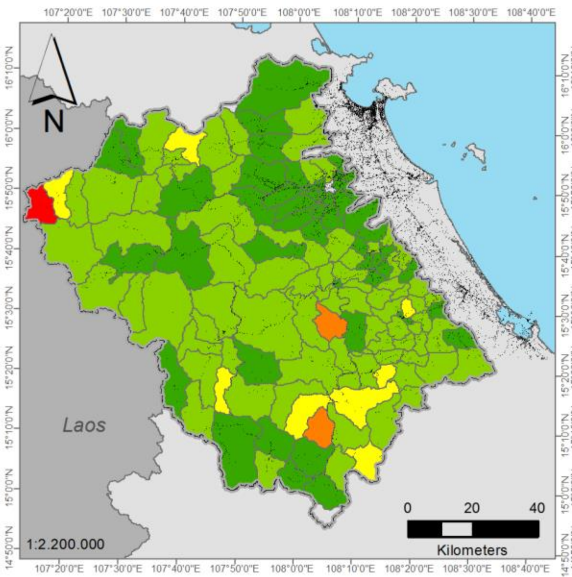


Distance to Paved Roads 2005-2010

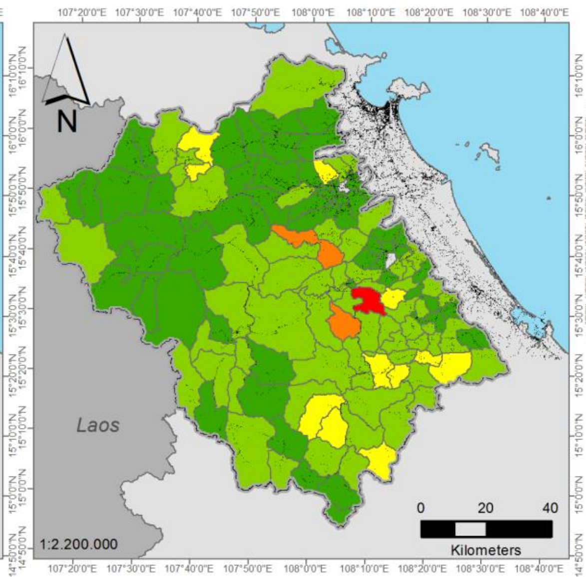


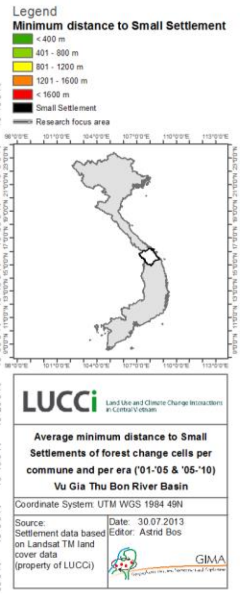


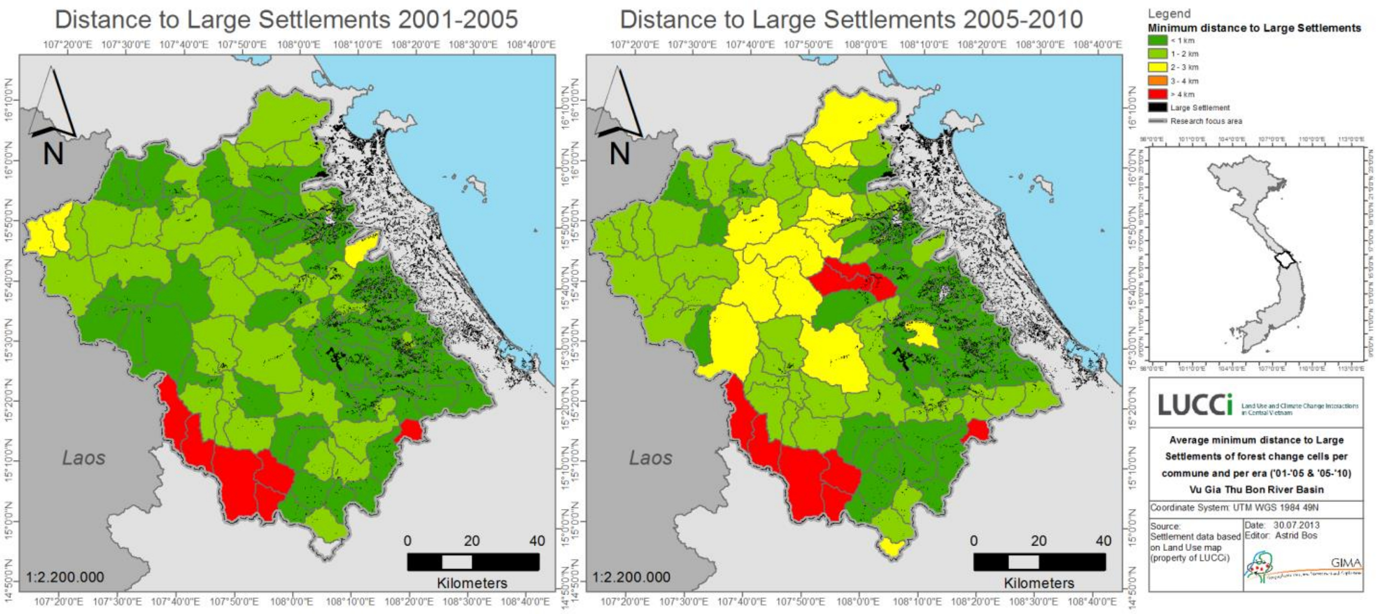
Distance to Small Settlements 2001-2005

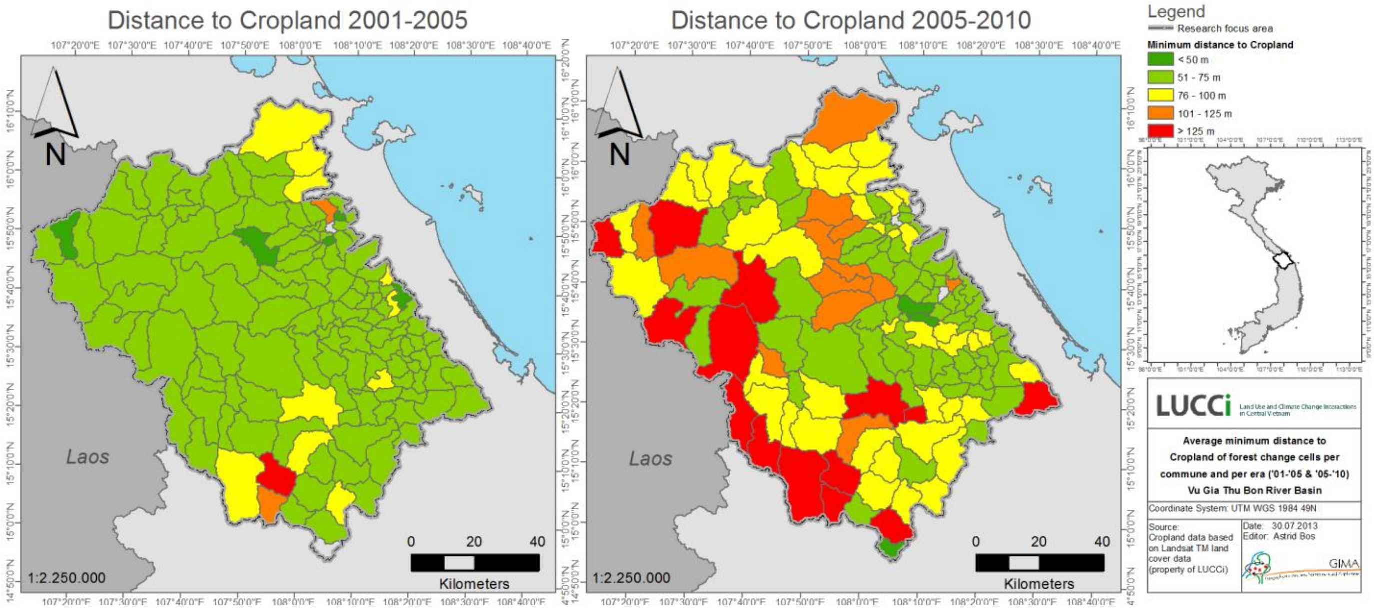


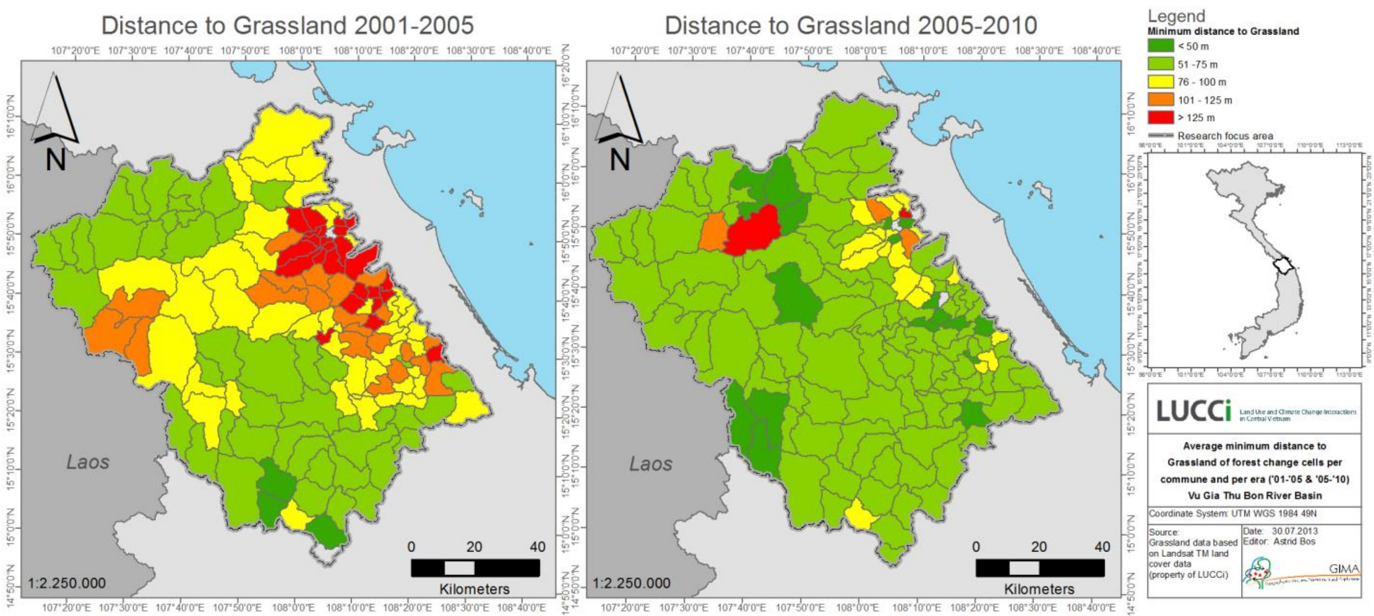
Distance to Small Settlements 2005-2010






Distance to Large Settlements 2001-2005
Distance to Large Settlements 2005-2010
Laos
1:2.200.000
Kilometers
Legend
Minimum distance to Large Settlements
< 1 km
1 - 2 km
2 - 3 km
3 - 4 km
> 4 km
Large Settlement
Research focus area
LUCCi
Average minimum distance to Large Settlements of forest change cells per commune and per era ('01-'05 & '05-'10)
Vu Gia Thu Bon River Basin
Coordinate System: UTM WGS 1984 49N
Source: Settlement data based on Land Use map (property of LUCCi)
Date: 30.07.2013
Editor: Astrid Bos
GIMA


Distance to Cropland 2001-2005
Distance to Cropland 2005-2010
Laos
1:2.250.000
Kilometers
Legend
Research focus area
Minimum distance to Cropland
< 50 m
51 - 75 m
76 - 100 m
101 - 125 m
> 125 m
LUCCi
Average minimum distance to Cropland of forest change cells per commune and per era ('01-'05 & '05-'10)
Vu Gia Thu Bon River Basin
Coordinate System: UTM WGS 1984 49N
Source: Cropland data based on Landsat TM land cover data (property of LUCCi)
Date: 30.07.2013
Editor: Astrid Bos
GIMA


Distance to Grassland 2001-2005
Distance to Grassland 2005-2010
Laos
1:2.250.000
Kilometers
Legend
Minimum distance to Grassland
< 50 m
51 -75 m
76 - 100 m
101 - 125 m
> 125 m
Research focus area
LUCCi
Average minimum distance to Grassland of forest change cells per commune and per era ('01-'05 & '05-'10)
Vu Gia Thu Bon River Basin
Coordinate System: UTM WGS 1984 49N
Source: Grassland data based on Landsat TM land cover data (property of LUCCi)
Date: 30.07.2013
Editor: Astrid Bos
GIMA

	Forest change 2001-2005	**Forest change 2005-2010**

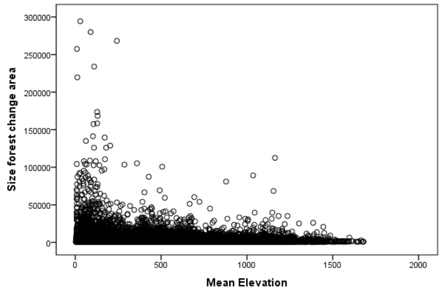
Elevation (DEM)

Forest change 2001-2005

Correlations

		Size forest change area	Mean Elevation
Size forest change area	Pearson Correlation	1	-,042**
	Sig. (2-tailed)		,000
	N	45233	45233
Mean Elevation	Pearson Correlation	-,042**	1
	Sig. (2-tailed)	,000	
	N	45233	45233

**. Correlation is significant at the 0.01 level (2-tailed).

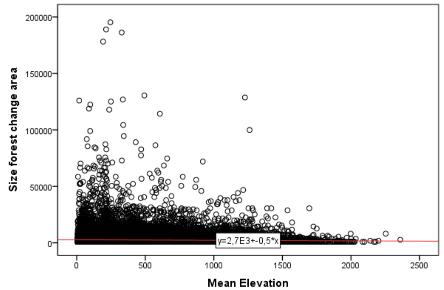


Forest change 2005-2010

Correlations

		Size forest change area	Mean Elevation
Size forest change area	Pearson Correlation	1	-,036**
	Sig. (2-tailed)		,000
	N	48578	48578
Mean Elevation	Pearson Correlation	-,036**	1
	Sig. (2-tailed)	,000	
	N	48578	48578

**. Correlation is significant at the 0.01 level (2-tailed).



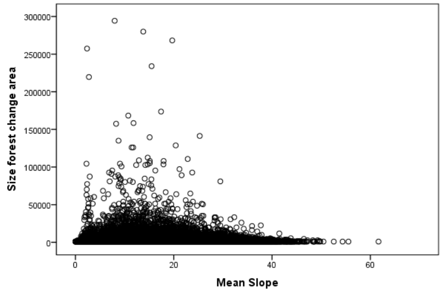
Slope

Forest change 2001-2005

Correlations

		Size forest change area	Mean Slope
Size forest change area	Pearson Correlation	1	-,011*
	Sig. (2-tailed)		,022
	N	45233	45233
Mean Slope	Pearson Correlation	-,011*	1
	Sig. (2-tailed)	,022	
	N	45233	45233

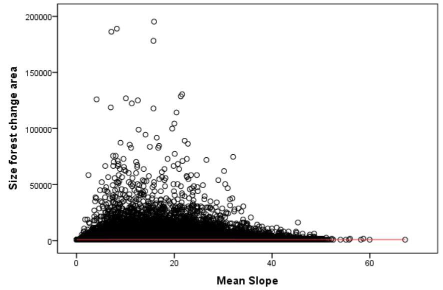
*. Correlation is significant at the 0.05 level (2-tailed).



Forest change 2005-2010

Correlations

		Size forest change area	Mean Slope
Size forest change area	Pearson Correlation	1	-,006
	Sig. (2-tailed)		,194
	N	48578	48578
Mean Slope	Pearson Correlation	-,006	1
	Sig. (2-tailed)	,194	
	N	48578	48578



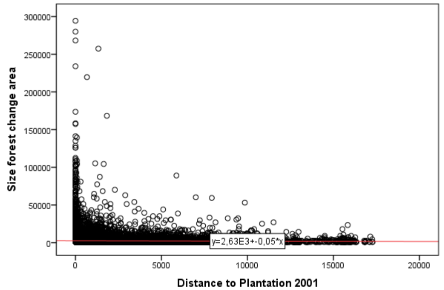
Distance to plantation 2001

Forest change 2001-2005

Correlations

		Size forest change area	Distance to Plantation 2001
Size forest change area	Pearson Correlation	1	-,020**
	Sig. (2-tailed)		,000
	N	45233	45233
Distance to Plantation 2001	Pearson Correlation	-,020**	1
	Sig. (2-tailed)	,000	
	N	45233	45233

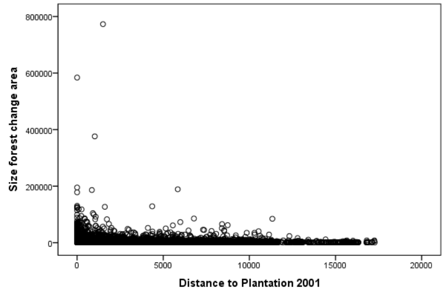
**. Correlation is significant at the 0.01 level (2-tailed).



Forest change 2005-2010

Correlations

		Size forest change area	Distance to Plantation 2001
Size forest change area	Pearson Correlation	1	-,002
	Sig. (2-tailed)		,637
	N	48581	48581
Distance to Plantation 2001	Pearson Correlation	-,002	1
	Sig. (2-tailed)	,637	
	N	48581	48581



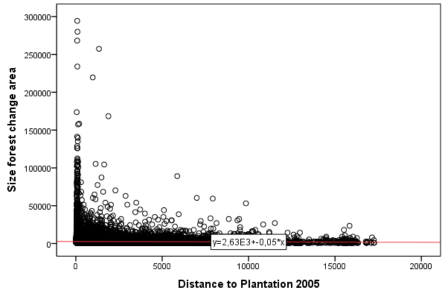
Forest change 2001-2005

Distance to plantation 2005

Correlations

		Size forest change area	Distance to Plantation 2005
Size forest change area	Pearson Correlation	1	-,019**
	Sig. (2-tailed)		,000
	N	45233	45233
Distance to Plantation 2005	Pearson Correlation	-,019**	1
	Sig. (2-tailed)	,000	
	N	45233	45233

**. Correlation is significant at the 0.01 level (2-tailed).

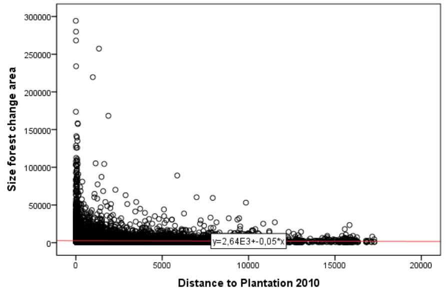


Distance to plantation2010

Correlations

		Size forest change area	Distance to Plantation 2010
Size forest change area	Pearson Correlation	1	-,020**
	Sig. (2-tailed)		,000
	N	45233	45233
Distance to Plantation 2010	Pearson Correlation	-,020**	1
	Sig. (2-tailed)	,000	
	N	45233	45233

**. Correlation is significant at the 0.01 level (2-tailed).

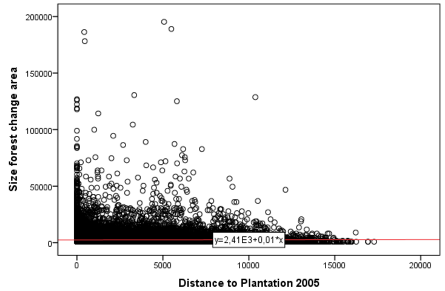


Forest change 2005-2010

Distance to plantation 2005

Correlations

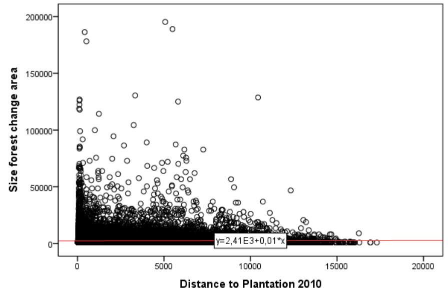
		Size forest change area	Distance to Plantation 2005
Size forest change area	Pearson Correlation	1	,008
	Sig. (2-tailed)		,065
	N	48578	48578
Distance to Plantation 2005	Pearson Correlation	,008	1
	Sig. (2-tailed)	,065	
	N	48578	48578



Distance to plantation 2010

Correlations

		Size forest change area	Distance to Plantation 2010
Size forest change area	Pearson Correlation	1	,009
	Sig. (2-tailed)		,051
	N	48578	48578
Distance to Plantation 2010	Pearson Correlation	,009	1
	Sig. (2-tailed)	,051	
	N	48578	48578

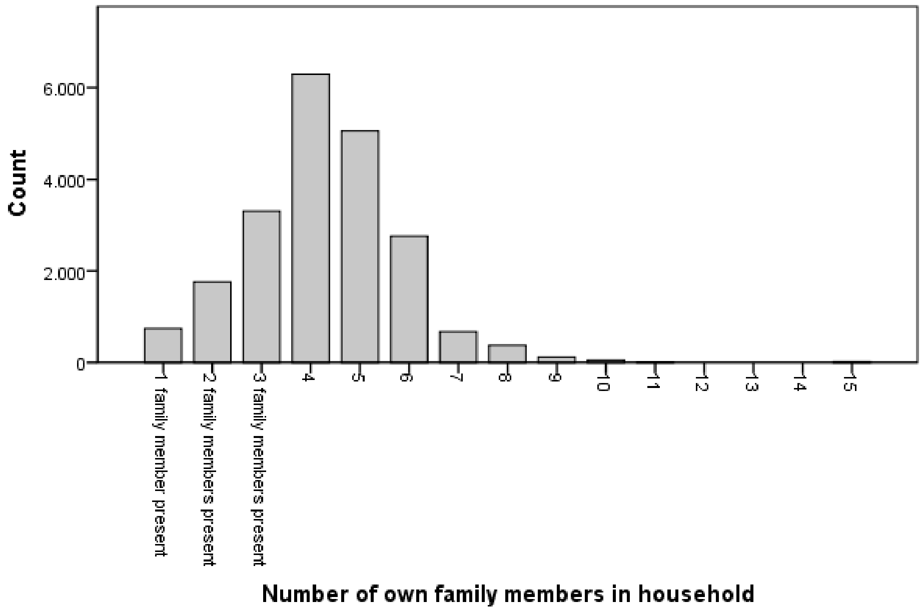


Number of families in household

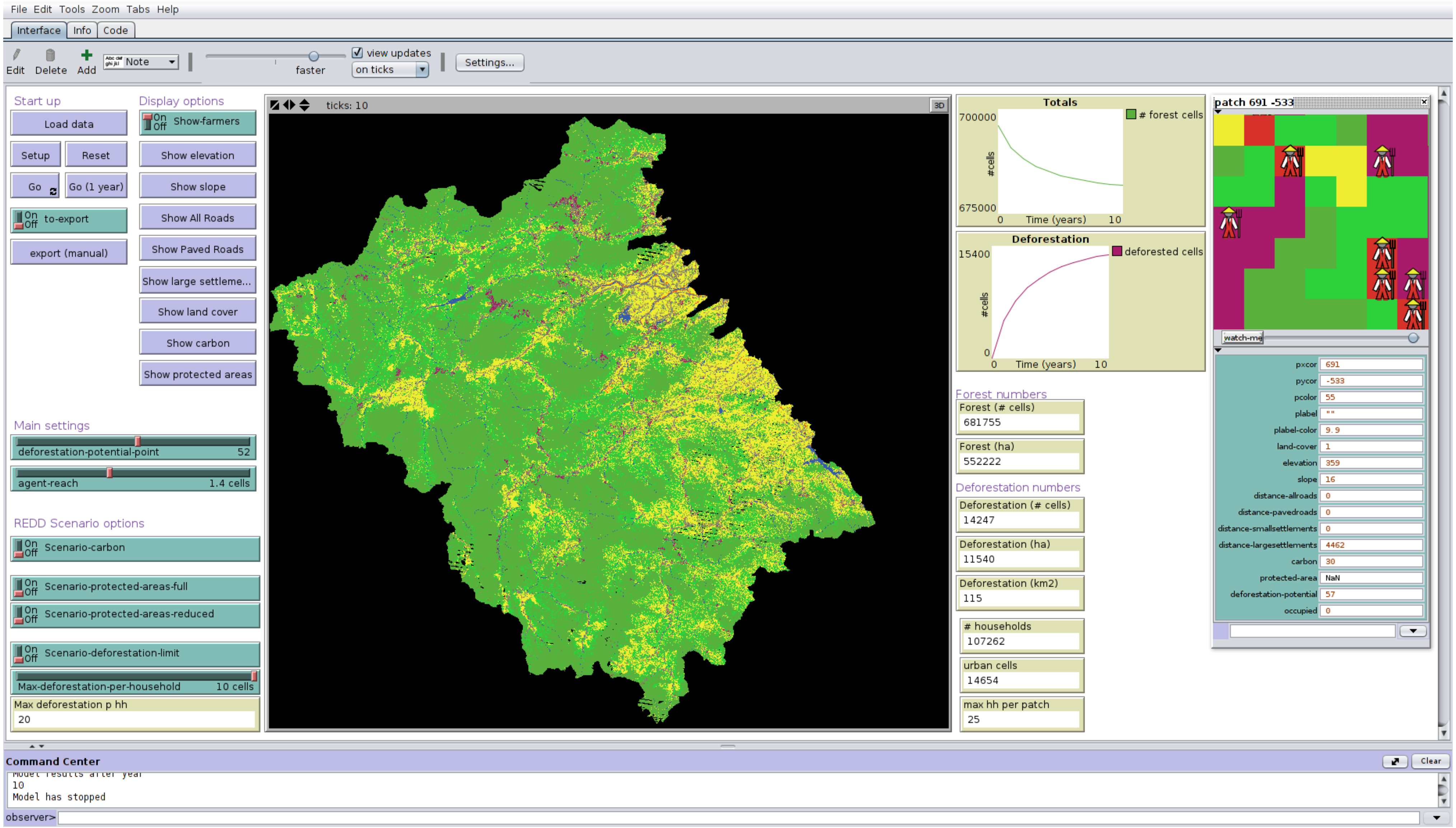
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 family	20404	96,5	96,5	96,5
	2 families	703	3,3	3,3	99,8
	3 families	32	,2	,2	99,9
	5 families	5	,0	,0	100,0
	6 families	9	,0	,0	100,0
	Total	21153	100,0	100,0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Number of own family members in household	21153	1	15	4,29	1,545
Valid N (listwise)	21153				



Appendix O Screenshot of SoDRA LUCCi v1.0 (full version)



```
extensions [gis]

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;                                                                   ;  Vararaible definition  ;                                                                ;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

globals                                ;global variables
[
  landcover-dataset
  elevation-dataset
  slope-dataset
;  allroads-dataset                    ;adviced to only load in DEMO version (i.e. not in large scale models)
  distance-allroads-dataset
;  pavedroads-dataset                  ;adviced to only load in DEMO version (i.e. not in large scale models)
  distance-pavedroads-dataset
;  largesettlements-dataset            ;adviced to only load in DEMO version (i.e. not in large scale models)
  distance-largesettlements-dataset
  carbon-dataset
  protected-dataset
]

turtles-own                            ;household related variables
[
  start-patch
  forest-harvest
]

patches-own                            ;environment related variables
[
  land-cover
  elevation
  slope
  distance-allroads
  distance-pavedroads
  distance-smallsettlements
  distance-largesettlements
  carbon
  protected-area
  deforestation-potential             ;defines the "suitability" of forest cells for deforestation
  occupied                            ;count the number of agents on their home patch
]

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;                                                                   ;  Loading Model Layers   ;                                                                ;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to load-data                                                                     ;loads all layers. needs to be executed only once

   clear-all
```

```
;Setup of land cover pixels
  print "loading land cover..."
  set landcover-dataset gis:load-dataset "./importfiles/vgtb_90_lc_2010.asc"
  gis:set-world-envelope (gis:raster-world-envelope landcover-dataset 0 0)
  gis:apply-raster landcover-dataset land-cover
  show-landcover

;===========

  print "loading DEM..."
  set elevation-dataset gis:load-dataset "./importfiles/dem_90m_utm49_m_clip.asc"
  gis:apply-raster elevation-dataset elevation

  print "loading DEM Slope..."
  set slope-dataset gis:load-dataset "./importfiles/dem_90m_utm49_m_clip_slope.asc"
  gis:apply-raster slope-dataset slope

;===========

  print "loading All Roads..."
;  set allroads-dataset gis:load-dataset "./importfiles/vgtb_dc_roadsall.shp"

  set distance-allroads-dataset gis:load-dataset "./importfiles/vgtb_90_roadsall_eucdis.asc"
  gis:apply-raster distance-allroads-dataset distance-allroads

;===========
  print "loading Paved Roads..."
;  set pavedroads-dataset gis:load-dataset "./importfiles/vgtb_dc_roadspav.shp"

  set distance-pavedroads-dataset gis:load-dataset "./importfiles/vgtb_90_roadspav_eucdis.asc"
  gis:apply-raster distance-pavedroads-dataset distance-pavedroads

;===========
  print "loading Large Settlements..."
;  set largesettlements-dataset gis:load-dataset "./importfiles/VGTB_DC_largeset.shp"

  set distance-largesettlements-dataset gis:load-dataset "./importfiles/vgtb_90_largesett_eucdis.asc"
  gis:apply-raster distance-largesettlements-dataset distance-largesettlements

;===========
  print "loading Carbon Stock..."
  set carbon-dataset gis:load-dataset "./importfiles/vgtb_90_carbon.asc"
  gis:apply-raster carbon-dataset carbon

;===========
  print "loading Protected Areas..."
  set protected-dataset gis:load-dataset "./importfiles/vgtb_90_protected.asc"
  gis:apply-raster protected-dataset protected-area

;===========
;===========

print "Loading data completed, press Setup"
end
```

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                              ;
;                                                                    ;  Model Initialisation   ;                                                              ;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                              ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;


to setup                                                            ;needs to be executed AFTER load data. can be repeated after model run
  clear-ticks                                                       ;has finished, which enforces placement of new agents, and resets deforestation
  reset-ticks
  clear-turtles
  clear-drawing
  clear-all-plots
  clear-output
  gis:apply-raster landcover-dataset land-cover
  show-landcover

;===========

;define deforestation-potential
calculate-deforestation-potential

;===========

;Setup of turtles, i.e. farmers
print date-and-time
print "Settling agents"
set-default-shape turtles "person vietnam"
ask patches [if pcolor = red [set occupied 10]]
  crt 107262                                                        ;Number of rural households in the focus area of VGTB
  [
    set size 1
    move-to one-of patches with [occupied >= 10 and occupied < 20]   ;Max 10 agents can be set on one urban patch (classified as "home")
    set occupied occupied + 1
    set start-patch patch-here
  ]
print date-and-time
print "Settling agents completed"

;===========

;Show/hide farmers
show-or-hide-farmers

;===========

;Scenario settings
set-scenarios

;===========

  reset-ticks
  print "Setup is ready, press Go"                                 ;End of Setup procedure, model is ready to go
  set-current-plot "Totals"

end
```

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                     ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;                                                                     ;          Reset          ;                                                                    ;
;                                                                     ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;


to reset
  ;movie-cancel                                                       ;can be executed AFTER model run has finished, agents return home,
  clear-ticks                                                         ;their forest-harvest rate will be reset, and deforestation will be reset
  reset-ticks
  clear-drawing
  clear-all-plots
  clear-output

  gis:apply-raster landcover-dataset land-cover
  show-landcover

;===========

;define deforestation-potential                                      ;deforestation-potential will be recalculated (original values)
calculate-deforestation-potential

;===========

;Reset turtles, i.e. farmers
print date-and-time
print "Agents return home"

ask turtles
[
  move-to start-patch
  set forest-harvest 0
]

print date-and-time
print "Agents repositioned"
;===========

;Show/hide farmers
show-or-hide-farmers

;===========

;Scenario settings
set-scenarios

;===========

  reset-ticks
  print "Reset is done, press Go"                                      ;End of Setup procedure, model is ready to go
  set-current-plot "Totals"

end
```

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;                                                                    ;        Model Run        ;                                                                    ;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to go
;  make-movie

  if (ticks >= 10)                                                              ;model represents period 2010-2020
  [
    if (to-export = true) [do-export]                                           ;if swith "to-export" is switched on, export of result is executed
    print "Model has stopped"                                                   ;the model stops after 10 yrs
    stop
  ]

  ask turtles
  [
    if (Scenario-deforestation-limit = false)                                   ;if switch Scenario-deforestation-limit is switched off,
      [search]                                                                  ;all agents start procedure "Search"

    ifelse (Scenario-deforestation-limit = true)                                ;if Scenario-deforestation-limit is switched on, only households that have not
    AND (forest-harvest >= Max-deforestation-per-household)                     ;reached their deforestation limit will start procedure "Search", others return
      [move-to start-patch]                                                     ;home.
      [search]
  ]
  tick

print "Model results after year " print ticks
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;                                                                    ;    Action definition    ;                                                                    ;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                    ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to calculate-deforestation-potential                                                       ;Is calculated when Setup or Reset is activated

  print "Calculating deforestation potential rates..."
  ask patches
  [
    set deforestation-potential 0
    if (land-cover = 1) and (elevation <= 618)
      [set deforestation-potential deforestation-potential + 12]                           ;influenced by DEM

    if (land-cover = 1) and (slope <= 21.11)
      [set deforestation-potential deforestation-potential + 7]                            ;influenced by Slope

    if (land-cover = 1) and (any? patches in-radius (0.978) with [land-cover = 2])
      [set deforestation-potential deforestation-potential + 17]                           ;influenced by distance to cropland

    if (land-cover = 1) and (any? patches in-radius (0.944) with [land-cover = 3])
      [set deforestation-potential deforestation-potential + 11]                           ;influenced by distance to grassland

    if (land-cover = 1) and (distance-smallsettlements <= 703)
```

```
      [set deforestation-potential deforestation-potential + 17]                          ;influenced by distance to small settlements

    if (land-cover = 1) and (distance-largesettlements <= 1845)
      [set deforestation-potential deforestation-potential + 15]                          ;influenced by distance to large settlements

    if (land-cover = 1) and (distance-allroads <= 634)
      [set deforestation-potential deforestation-potential + 14]                          ;influenced by distance to all roads

    if (land-cover = 1) and (distance-pavedroads <= 6901)
      [set deforestation-potential deforestation-potential + 7]                           ;influenced by distance to paved roads

    if (land-cover != 1) [set deforestation-potential 0]                                  ;non-forest cannot be deforested
  ]
end

;================================================================================================================================================================

to search                                                                         ;Will be run when Go is clicked
                                                                                  ;procedure representing search + deforestation of potential areas
;  pen-down

  let q max-one-of patches in-radius agent-reach [deforestation-potential]        ;agents search in their reach area for cells with highest
  ifelse[deforestation-potential] of q > deforestation-potential                  ;deforestation-potential. If the deforestation-potential is higher than the
    [face q move-to q]                                                            ;deforestation-potential on its own (current) cell, the agent will move to
    [move-to patch-here]                                                          ;that cell.


  if deforestation-potential >= deforestation-potential-point                     ;if the deforestation-potential on which the agent stands, is higher than or
  [                                                                               ;equal to the minimum level ("suitability") for deforestation, the agent will
    set pcolor magenta set land-cover 7 set deforestation-potential 0             ;deforest that cell.
    set forest-harvest forest-harvest + 1
  ]


  if pcolor = blue                                                               ;if farmers move into cells classified as "water",
    [move-to one-of neighbors with [(pcolor != blue) AND (pcolor != black) ]]    ;go to random neighbor that is neither water nor outside area

  if pcolor = black                                                              ;if farmers move outside study area,
    [move-to one-of neighbors with [(pcolor != blue) AND (pcolor != black) ]]    ;go to random neighbor inside study area and not water

end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                  ;
;                                                                    ;    Scenario Options     ;                                                                  ;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                  ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to set-scenarios
ask patches
[
  if (Scenario-carbon = true) AND carbon = 110                        ;If switch "Scenario-carbon" is switched on, highest carbon class will be fully protected
    [set deforestation-potential 0]

  if (Scenario-carbon = true) AND carbon = 56                         ;If switch "Scenario-carbon" is switched on, deforestation potential of 2nd highest class
    [set deforestation-potential deforestation-potential - 10]        ;will be reduced with 10
```

```
  if (Scenario-protected-areas-full = true) AND protected-area = 1     ;If switch "Scenario-protected-areas-full" is switched on,
    [set deforestation-potential 0]                                     ;deforestation is completely prohibited in protected areas


  if (Scenario-protected-areas-reduced = true) AND protected-area = 1  ;If switch "Scenario-protected-areas-reduced" is switched on,
    [set deforestation-potential deforestation-potential - 10]          ;deforestation potential in protected areas will be reduced with 10
]
end


;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;                                                                    ;     Display Options     ;                                                                ;
;                                                                    ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to show-elevation                                              ;Shows DEM
  gis:paint elevation-dataset 50
end
;===============

to show-slope                                                  ;Shows slope
  gis:paint slope-dataset 50
end
;===============

to show-allroads                                               ;Shows all roads
  gis:paint distance-allroads-dataset 1
  gis:set-drawing-color gray
;  gis:draw allroads-dataset 1
end
;===============

to show-pavedroads                                             ;Shows paved roads
  gis:paint distance-largesettlements-dataset 1
  gis:set-drawing-color 3
;  gis:draw pavedroads-dataset 2
end
;===============

to show-largesettlements                                       ;Shows large settlements
  gis:paint distance-largesettlements-dataset 1
;  gis:set-drawing-color 14
;  gis:fill largesettlements-dataset 1
end
;===============

to show-landcover                                             ;Shows original land cover classes

    ask patches
  [
    if land-cover = 1 [set pcolor green]                      ;representing forest land
    if land-cover = 2 [set pcolor yellow]                     ;representing cropland
    if land-cover = 3 [set pcolor lime]                       ;representing grassland
```

```
    if land-cover = 4 [set pcolor blue]                      ;representing water
    if land-cover = 5 [set pcolor red]                       ;representing urban/rural settlements
    if land-cover = 6 [set pcolor orange]                    ;representing other land
  ]
end
;===============
to show-carbon                                               ;Shows carbon classes in viewer
  ask patches
  [
    if (carbon > -9999)
    [set pcolor scale-color green carbon 110 0]
  ]
end
;===============

to show-protected-areas                                      ;Shows protected areas in viewer
  ask patches
  [
    if (protected-area > -9999)
      [set pcolor red]
  ]
end
;===============

to show-or-hide-farmers
  ask turtles
  [
    If (Show-farmers = true) [show-turtle]                  ;If switch "Show-farmers" in interface is switched on, the farmers will be visible
    If (Show-farmers = false) [hide-turtle]                 ;If switch "Show-farmers" in interface is switched off, the farmers will be invisible
  ]
end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                   ;
;                                                                   ;       Model Report      ;                                                                   ;
;                                                                   ;;;;;;;;;;;;;;;;;;;;;;;;;;;                                                                   ;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to do-export
  print "Export busy...."
  let export-path "./modeloutput/"

  let datetime remove ":" date-and-time
; gis:store-dataset (gis:patch-dataset land-cover) word(word export-path "SoDRA_output_") (word behaviorspace-run-number)   ;Output during BehaviorSpace runs
  gis:store-dataset (gis:patch-dataset land-cover) word(word export-path "SoDRA_output_") datetime                          ;Output during regular runs

  print "Export done."
end
```

	parameter	reference value	"-5%"	"+5%"	C	C- proc	C-	C+ proc	C+	S-	S+
1	deforestation potential point	52.000	49.400	54.600	29287.66	-5.00%	31266.38	5.00%	28043.78	-39574.4	-24877.6
2	agent reach	1.500	1.425	1.575	29287.66	-5.00%	29262.78	5.00%	29275.3	497.6	-247.2
3	elevation 75%	618.000	587.100	648.900	29287.66	-5.00%	29091.46	5.00%	29431.96	3924	2886
4	slope 75%	21.110	20.055	22.166	29287.66	-5.00%	28985.32	5.00%	29473.84	6046.8	3723.6
5	distance to cropland (88/90) 75%	0.978	0.929	1.027	29287.66	-5.00%	29250.62	5.00%	31786.86	740.8	49984
6	distance to grassland (85/90) 75%	0.944	0.897	0.992	29287.66	-5.00%	29284.4	5.00%	29239.14	65.2	-970.4
7	distance to small settlements 75%	703	668	738	29287.66	-5.00%	29263.26	5.00%	29294.18	488	130.4
8	distance to large settlements 75%	1845	1753	1937	29287.66	-5.00%	29021.64	5.00%	29619.22	5320.4	6631.2
9	distance to all roads 75%	634	602	666	29287.66	-5.00%	28966.88	5.00%	29719.14	6415.6	8629.6
10	distance to paved roads 75%	6901	6556	7246	29287.66	-5.00%	29212.94	5.00%	29304.4	1494.4	334.8
11	elevation weight	12	11.4	12.6	29287.66	-5.00%	29282.64	5.00%	29263.62	100.4	-480.8
12	slope weight	7	6.65	7.35	29287.66	-5.00%	29291.96	5.00%	29290.24	-86	51.6
13	distance to cropland weight	17	16.15	17.85	29287.66	-5.00%	29268.48	5.00%	29223.52	383.6	-1282.8
14	distance to grassland weight	11	10.45	11.55	29287.66	-5.00%	29266.9	5.00%	29222.22	415.2	-1308.8
15	distance to small settlements weight	17	16.15	17.85	29287.66	-5.00%	29327.62	5.00%	29335.52	-799.2	957.2
16	distance to large settlements weight	15	14.25	15.75	29287.66	-5.00%	29285.48	5.00%	29243.18	43.6	-889.6
17	distance to all roads weight	14	13.3	14.7	29287.66	-5.00%	29249.64	5.00%	29248.3	760.4	-787.2
18	distance to paved roads weight	7	6.65	7.35	29287.66	-5.00%	29290.82	5.00%	29298.84	-63.2	223.6
19	nr agents (rural households) weight	107262	101899	112625	29287.66	-5.00%	29279.5	5.00%	29260.72	163.2	-538.8