Table of Contents

List of Acronyms

AIC	Akaike Information Criterion
CLUE	The Conversion of Land Use and its Effects
CPRM	Serviço Geológico do Brasil (Brazilian Geological Service)
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos (Center for Weather Forecasts and Climate Studies)
FUNAI	Fundação Nacional do Índio (National Indian Foundation)
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (Brazilian Institute of Environment and Renewable Natural Resources)
IBGE	Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics)
INCRA	Instituto Nacional de Colonização e Reforma Agrária (National Institute for Colonization and Agrarian Reform)
INMET	Instituto Nacional de Meteorologia (National Institute for Meteorology)
INPE	Instituto Nacional de Pesquisas Espaciais (National Institute for Space Research)
LUCC	Land-use and Land-cover change
PVM	Potential Vegetation Model
SRTM	Shuttle Radar Topography Mission

1 Introduction

The Amazon rainforest is the largest tropical rainforest in the world. It has an extent of approximately 5.5 million km^2, of which about 60% are located in Brazil (Andersen, Granger, Reis, Weinhold, & Wunder, 2002). Due to its rich biodiversity and its potential role in global climate discussions, deforestation in the Amazon is not only of local interest, but leads to questions of global environmental and economical concern (Andersen, Granger, Reis, Weinhold, & Wunder, 2002; Malhi et al., 2008; Werth & Avissar, 2002).

In the last decades the Amazon rainforest has been under increasing human pressure. The so-called Legal Amazon[1] (Amazônia Legal), which comprises nine Brazilian states, experienced a population increase from 4 million people in 1950 (Barreto, Souza Jr, Nogueron, Anderson, & Salomão, 2006) to almost 24 million people in 2007 (IBGE, 2007). Expansion of pasture areas for cattle-ranching and increasing demand for mechanized agriculture (e.g.: soybeans) are seen as the major drivers leading to massive forest clearing especially in the so-called arch of deforestation (Becker, 2005; Kaimowitz, Mertens, Wunder, & Pacheco, 2004; Nepstad & Stickler, 2006).

Since 1988 the National Institute for Space Resarch INPE (Instituto Nacional de Pesquisas Espaciais) has been monitoring deforestation in the Brazilian Amazon and provides accurate, annual deforestation maps and rates (INPE, 2010). Figure 1-1 shows the spatial patterns of deforestation as measured by INPE in 2007. In recent years the deforestation rate reached a maximum of 27423 km^2 in 2004 and had an average value of 17133 km^2 between 1997 and 2009. Figure 1-2 shows annual deforestation rates from 1997 to 2009.

[1] The Legal Amazon (Amazônia Legal) is not an uniform biome, as it initially was defined for regional planning purposes. It mainly consists of forests, savannahs/cerrados, inundated lowlands and steppes. (Andersen, Granger, Reis, Weinhold, & Wunder, 2002)

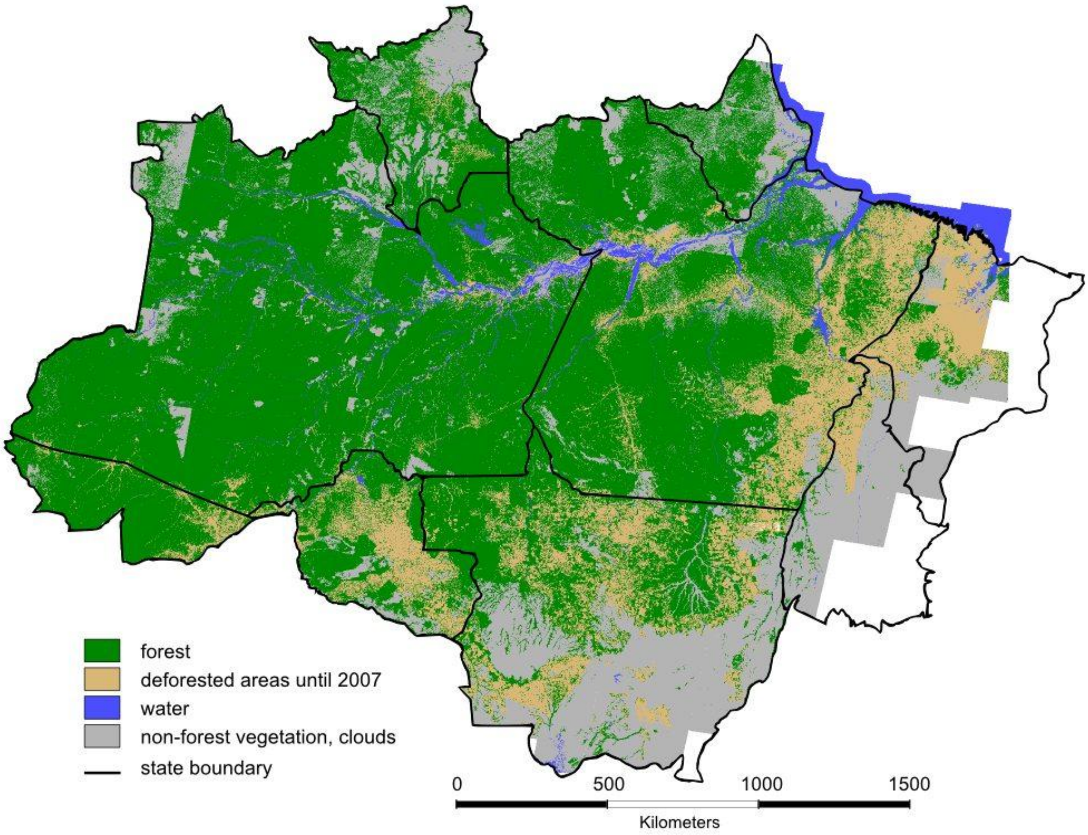

Figure 1-1: Deforestation map of the Brazilian Amazon in 2007 (INPE, 2010)

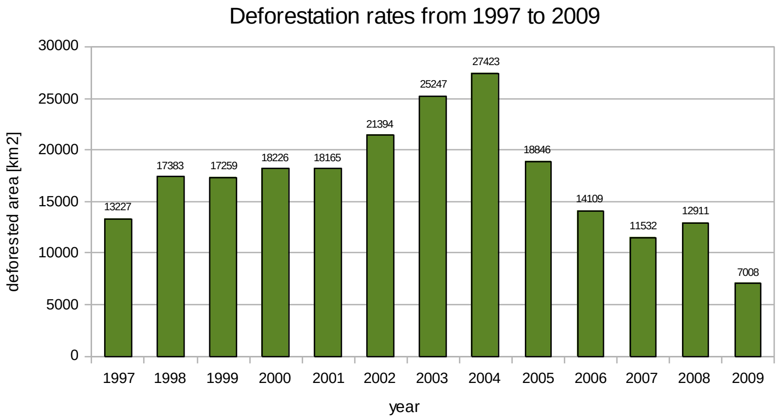

Figure 1-2: Deforestation rates from 1997 to 2009 (INPE, 2010)

Tropical deforestation is an issue not only present in South America, but can also be observed in tropics in other parts of the world, where anthropogenic impact leads to land-use and land-cover changes (LUCC). Land cover refers to the attributes of Earth's land surface and immediate subsurface (e.g.: forest, grassland etc.) and land use to the purposes for which humans exploit the land cover (e.g.: forestry, pasture etc.) (Lambin, Geist, & Rindfuss, 2006). A lot of research in the area of LUCC modeling tries to investigate and simulate the human influence on once pristine forests. (Geist & Lambin, 2001) compare various subnational LUCC case studies to analyze proximate[2] and underlying[3] causes of tropical deforestation. Their findings indicate that tropical deforestation can not be explained by a single or even a few variables, but by the interplay of several proximate and underlying factors. For Latin America they found a road-agriculture tandem as a robust, causative connection at the proximate level to be of relevant importance. In addition biophysical factors as relief or topography in combination with soil quality and water availability shape the patterns of deforestation in cases with high rates of annual deforestation (Geist & Lambin, 2001).

Various modeling approaches exist to simulate the dynamics of land-use changes. These range from empirical models, based on statistical analyses like the CLUE model (De Koning, Veldkamp, Kok, & Bergsma, 1998; Kok, Farrow, Veldkamp, & Verburg, 2001; Veldkamp & Fresco, 1996; Verburg, De Koning, Kok, Veldkamp, & Bouma, 1999) to stochastic cellular automata models like DINAMICA (Soares-Filho, 2002) or agent-based models like LUCITA (Deadman, Robinson, Moran, & Brondizio, 2004). Understanding of past deforestation processes is essential for projecting and exploring of future scenarios, to provide decision makers with reliable tools and fundamental up-to-date information.

Besides the substantial loss of biodiversity due to forest decline (Barreto, Souza Jr, Nogueron, Anderson, & Salomão, 2006; Fearnside, 2005) the interplay between climate and land-use changes is an important environmental issue. Thus several publications discuss these bi-directional interactions between climate and land-use dynamics to assess the vulnerability of the Amazon to global climate change on one side and the contribution of land-use changes to the climate on the other side (Aragão et al., 2008; Foley, Costa, Delire, Ramankutty, & Snyder, 2003; Gash, 1996; Malhi et al., 2008; Nobre, Sellers, & Shukla, 1991). No unified agreement on how the Amazonian climate might change due to deforestation has been reached so far, but most studies indicate that surface temperature has the tendency to rise, while precipitation might decrease in some parts of the Amazon, leading to significant drying in some areas during the 21st century (Malhi et al., 2008; Voldoire & Royer, 2004). Even tipping the biome-climate system towards a new drier

[2] Proximate causes of deforestation are human activities (land uses) that directly affect the environment and thus constitute proximate sources of change. They operate on the local scale and can be structured in three groups: agricultural expansion, wood extraction and infrastructure extension. (lambin, geist, 2001)
[3] Underlying causes of deforestation or driving forces may directly act at the local level or indirectly at national and global level. They are a complex of social, political, economic, technological and cultural variables, which are seen as fundamental forces which underpin proximate causes of deforestation (lambin, geist, 2001).

stable equilibrium state by land-use changes seems possible for tropical South America (Oyama & Nobre, 2003).

This work uses a Potential Vegetation Model (PVM) based on climatic variables, developed at INPE (Oyama & Nobre, 2004) to derive additional input parameters for simulation of land-use change processes in the Brazilian Amazon. This PVM incorporates a water balance model from which climate dependent variables, such as soil wetness and seasonality index, are derived. The Potential Vegetation Model and the hydrological model will be implemented in the same modeling framework as a LUCC model for the Amazon, also previously developed at INPE (Aguiar, 2006; Moreira, 2009). Thus it will be possible to examine the explanatory power of additional environmental factors in a land-use and land-cover change model, especially to discriminate agriculture from pasture land-use patterns. The combined use of such models allows for comprehensible data integration through the same database and can further be seen as a prerequisite for dynamically coupling climate-LUCC models in the future. Hence this thesis can also be seen as a step towards future research topics regarding the construction of integrated environmental models.

In this context, the scientific question of this thesis is thus to understand how such new environmental variables, derived from the water balance model, in conjunction with other environmental variables, such as slope and altimetry, can help to improve land-use change projections in the Amazon, facilitating the construction of coupled climate-LUCC integrated models in the future.

1.1 Hypothesis

The inclusion of hydrological, slope and altimetry variables improves the ability to discriminate pasture and agriculture patterns in the Brazilian Amazon.

1.2 Objectives

To address this hypothesis the following objectives were defined for this thesis:

→ Progress towards future coupling of models by implementing a potential vegetation model and its corresponding water balance model in the same modeling framework as a LUCC model

→ Verification of the adequacy of hydrological variables derived from a water balance model to simulate and discriminate agriculture and pasture patterns in the Brazilian Amazon, using statistical analysis and spatially-explicit LUCC models.

1.3 Structure of the thesis

The structure of the thesis is the following. Chapter 1 starts with an introduction to the topic, the hypothesis and the objectives of the work. In Chapter 2 a literature review will be summarized to give an overview of the state of research. Chapter 3 describes the study area and the methods that will be used to validate the proposed hypothesis. Chapter 4 shows the results of the dynamic modeling approach. In chapter 5 a summary is given and conclusions are drawn.

2 Literature Review

This chapter consists of a review of LUCC models and the coupling of climatic models to LUCC models.

2.1 LUCC Models

2.1.1 Overview of LUCC modeling approaches

(Lambin, 2004; Lambin, Rounsevell, & Geist, 2000) distinguish four categories of land-use change models: empirical-statistical, stochastic, optimization and dynamic (process-based) models. LUCC models in a fifth category combine various model types and are named integrated modeling approaches. (Briassoulis, 2000) distinguishes between statistical and econometric, spatial interaction, optimization and integrated models and a class incorporating model types which do not fall into one of these classes, while (Heistermann, Müller, & Ronneberger, 2006) classify LUCC models into geographic (empirical-statistical or rule-based/process-based), economic and integrated models. No matter which model type is used, modeling of land-use change tries to address at least one of the following questions (Lambin, 2004):

- Which socio-economic and biophysical variables contribute most to an explanation of land-use changes and why?
- Which locations are affected by land-use changes – where?
- At what rate do land-use and land-cover change progress – when?

Another differentiation between land-use change models is defined by the attribute of being spatially-explicit. (Goodchild, 2002) defines four simple tests to investigate if a LUCC model is spatially-explicit. Corresponding to one of these tests, the outcome test, the author states that the most important reason for LUCC modeling to be spatially-explicit may relate to the model outcomes, as spatial patterns resulting from the processes of

LUCC are of significant interest to policy makers. Hence, if a LUCC model is assessed through the spatial patterns it produces, it is defined as being spatially-explicit.

An overview of economic models of deforestation can be found in (Kaimowitz & Angelsen, 1998), where 150 different models are reviewed. (Barbier & Burgess, 2001) provide a survey of economic studies on tropical deforestation and land-use at the cross-country level. (Geist & Lambin, 2001) compare various subnational LUCC case studies to analyze proximate and underlying causes of tropical deforestation.

2.1.2 LUCC models in the Brazilian Amazon

Numerous LUCC studies investigate land-use changes caused by deforestation in tropical South America, ranging from local studies to regional models covering the whole Amazon. The scientific areas of the authors of these studies differ (e.g.: GIScience, economics, computer science, remote sensing etc.) and so vary the types of models and the applied modeling approaches. In the following some modeling approaches of tropical deforestation in the Brazilian Amazon are reviewed.

(Andersen & Reis, 1997) develop an econometric model of deforestation and economic development in the Brazilian Amazon to evaluate the effects of different policy instruments. The authors use a panel data set covering 316 municipalities from 1970 to 1985 in five year steps. This data set comprises economic, ecological and demographic variables. The two-sector model consists of a rural and an urban sector and six equations. Past characteristics of a region and its neighbors are used by the main equation to predict the demand for newly cleared area, while the remaining equations assess the interaction between rural and urban populations, rural and urban output and land prices. The results of this LUCC study indicate a positive trade-off between economic growth and deforestation for subsidized credit for two main reasons. The authors conclude that subsidized credit promotes higher land prices which imply more efficient land-use and that farmers are stimulated to invest in more profitable and sustainable perennial crops. On the other side they state, based on the results of their model, that road building into pristine areas is harmful, but good in already cleared areas where it improves infrastructure.

(Andersen, Granger, Reis, Weinhold, & Wunder, 2002) introduce an econometric model with an updated methodology and data of the model published in (Andersen & Reis, 1997). This LUCC model simulates land clearing and economic development considering the growth rates of clearing and the growth rates of rural GDP[1]. The model is evaluated for two different time periods. The first period is from 1980 to 1985 and the second from 1985 to 1995. Six endogenous variables are used: land-clearing, rural and urban GDP growth, rural and urban population growth and cattle herd growth. In addition to these variables models of paved and unpaved roads are included. A general-to-simple approach is used to eliminate factors out of the 74 initial potential explanatory variables. The model shows that herd growth and new land clearing are mainly affected by natural frontier spatial processes of maturation and urban demand centers. The results further indicate that building of paved roads in forested areas leads to more clearing than building of unpaved roads. In the model unpaved roads are associated mainly with land-extensive activities while paved roads correspond to more land-intensive economic activities. The authors expected to find a deforestation reducing effect in regions with high rainfall, but according to the model rainfall did not affect the growth rates of clearing and the growth of rural GDP. Simulating the Avança Brasil[2] road construction plan the LUCC study found economic gains, but no overall increase in cleared area, which the authors explain by a possible underestimation of the impact of paved roads in relatively undisturbed areas. Nevertheless, the authors recommend reducing ecological costs by paving roads only in well-established areas. They further investigate the proposal to modify the "forest law"[3] from 80% to 50% and state that the economic costs of sustaining the 80% threshold outreach the value of forest services. Hence the authors propose to change the law to 50% and to improve infrastructure in settled areas, instead of building roads through undisturbed areas.

(Laurance et al., 2002) describe an empirical-statistical land-use change model for the Brazilian Amazon. The study area is subdivided into regular grids at two spatial scales, 50x50 km^2 and 20x20 km^2. Satellite imagery from 1999 is used to estimate the proportion of forest cover, deforested area and natural water bodies in each grid cell. A set of variables comprising human-demographic factors, factors that affect physical accessibility to forests and factors that may affect land-use suitability for human occupation and agriculture is used in the statistical analysis which is carried out on a random set of 120 cells, out of the 1927 cells at scale 50x50 km^2. A robust ordination method results in the development of two major axes of variation, which lead to the conclusion that highways

[1] GDP … gross domestic product
[2] Avança Brasil is a plan by the Brazilian government which involves a lot of investment in the Amazon region including infrastructure projects, social development projects, environmental projects and information collection (Andersen, Granger, Reis, Weinhold, & Wunder, 2002)
[3] The "forest law" states that 80% of forest inside private properties must be preserved.

(paved roads), human population density and dry-season severity are the main factors leading to local deforestation, while rainfall and unpaved roads have a smaller influence. Thus the authors state that these three factors largely determine deforestation in the Brazilian Amazon.

(Soares-Filho et al., 2006) focuses on modeling conservation in the Amazon basin by using an empirically based, policy-sensitive model of deforestation. Eight different scenarios for the time from 2001 to 2050 are utilized to project the influence of conservation approaches on the future development of the Amazonian rainforest. The used method comprehends two models. The first model divides the Amazon basin into 47 socioeconomic subregions and calculates deforestation rates for these regions based on historical trends, a planned road paving schedule and existing and proposed protected areas. It then passes these rates to a second model. This spatially-explicit model uses DINAMICA (Soares-Filho, 2002) a cellular automata model on cells with 1 km^2 resolution to allocate the demand for land clearing and thus simulate the spatial patterns of deforestation using static (e.g.: distance to major rivers) and dynamic (e.g.: distance to deforested land) variables. The scenarios simulate future forest loss mainly in the eastern Amazon and along the BR-364 from Rondônia to Acre and thereby reducing the area of closed-canopy forest from 5.3 million km^2 in 2003 to 3.2 million km^2 in the "business-as-usual" scenario and to 4.5 million km^2 under a governance scenario by 2050. Areas in the northwestern part of Amazonas state and areas outside Brazil may remain largely untouched due to their remoteness. According to the authors, protected areas in the Amazon are suitable to maintain mammalian diversity to a large degree, but not to secure critical watersheds and ecoregions from impoverishment, hence they recommend improved conservation strategies outside of protected areas.

(Aguiar, 2006) introduces an empirical-statistical, spatially-explicit model in the Brazilian Amazon. Spatial lag regression models and multiple linear regression models are used at two spatial scales to find statistical relationships between different land-use types and potential land-use determining factors. These factors come from a list of variables categorized in the following groups: accessibility to markets, economical attractiveness, public policies, agrarian structure, demographical, technological, and environmental factors. The applied modeling approach allows simulating different levels of law enforcement, road paving and creation of protected areas. Various combinations of factors and model parameters are used in five exploration scenarios from 1997 to 2020. Significant variations in relative importance of land-use determining factors are found between different regions. A result of these differences is that the impact of local policies also varies across space. The author thus advises to account for this intra-regional heterogeneity in land-use change models of the Amazon. A further conclusion is that connectivity measures to national markets are amongst the most important factors to capture deforestation dynamics in the new Amazonian frontiers, but they can only explain land-use patterns in combination with other factors.

(Moreira, Costa, Aguiar, Câmara, & Carneiro, 2009) develop an approach to build a multi-scale land-use change model including top-down and bottom-up interactions and test this dynamic coupling effort with a macro model of the Brazilian Amazon and a local model of Iriri/Terra do Meio in Pará state. At the macro scale (25x25 km^2 cells) the empirical-statistical modeling approach and data as described in (Aguiar, 2006) are used. At the local scale (1x1 km^2 cells) an agent-based model is implemented. Two sets of agents are defined, each with a set of actions and decision rules. The multi-scale model contains top-down actions, e.g. to send demand for land-use types to the local scale, and bottom-up feedbacks, e.g. to notify the global scale that the demand could not be allocated due to local policy restrictions. Four combinations of scenarios at both levels are tested from 1997 to 2025. The results indicate that local conditions do not determine the pressure for land-use change alone, it is also regulated by processes acting at higher hierarchical levels. To account for local and regional land-use change processes under varying biophysical and socioeconomic conditions is thought to be one of the strengths of multi-scale models. The authors thus conclude that models using top-down and bottom-up interactions can detect processes, which might are missed considering single scale models.

2.1.3 CLUE modelling framework and its adaptation to the Amazon

The CLUE (Conversion of Land-Use and its Effects) modeling framework is a dynamic, multi-scale land-use and land-cover change model (De Koning, Veldkamp, Kok, & Bergsma, 1998; Kok, Farrow, Veldkamp, & Verburg, 2001; Veldkamp & Fresco, 1996; Verburg, De Koning, Kok, Veldkamp, & Bouma, 1999). Currently three different versions exist: CLUE for regional to global scale analysis, CLUE-CR as the first implementation of the CLUE model applied to Costa Rica and CLUE-S for regional scale analysis.

The CLUE model has the objective to provide a spatially-explicit, multi-scale, quantitative description of land-use changes. It explores possible changes in the near future under different development scenarios. The model consists of a demand and an allocation module. In the non-spatial demand module scenarios of the quantity of change define how much change takes place in every time step. The spatial allocation module calculates where the changes are likely to happen. Connections between potential explanatory variables, such as socio-economic or environmental variables and various land-use types are assessed by multiple regression analysis. The CLUE model has been applied to various regions to study a large variety of land-use change issues, e.g. agricultural intensification, urbanization or deforestation. (Verburg & Overmars, 2007)

CLUE-CR was the first dynamic multi-scale land-use/cover change model based on the CLUE framework. It was applied to Costa Rica at local, regional and national scales (Veldkamp & Fresco, 1996b). Written in PASCAL, the model incorporates five different land-use/cover classes in percent of total grid cell cover and uses a set of scale-dependent land-use/cover linear regressions as input. Altitude, temperature, relief, soil drainage, rural and urban population and other data is utilized in the nested scale analysis. Analyzing two scenarios the authors deduced that the CLUE-CR model is able to simulate effects of several driving forces on land-use/cover change in Costa Rica (Veldkamp & Fresco, 1996b).

CLUE-S (the Conversion of Land-Use and its Effects at Small regional extent) has been developed for regional scale analysis (Veldkamp et al., 2002). The major change to CLUE and CLUE-CR is the different data representation. The land-use/cover is no longer represented as a fraction of total grid cell cover, each grid cell contains only the dominant land-use/cover type. Thus CLUE-S has primarily been developed for resolutions from some meters up to 1000 meters for areas where high-resolution data is available. Two modules are incorporated in the modeling procedure. The non-spatial analysis uses driving factors of change to calculate the demand for the different land-use/cover types, while in the second module the spatial analysis and land-use/cover allocation takes place. The allocation in the CLUE-S model can be determined by four different methods or a combination of them: empirical analysis, decision rules, neighborhood functions and conversion elasticity. The CLUE-S model is a tool for analysis of land-use processes and can be used to study different mechanisms of land allocation (Verburg & Overmars, 2007).

A schematic representation of the spatial allocation module of the CLUE model is shown in Figure 2-1. The CLUE model allocates the area of each land-use type as defined by the demand module. This procedure is sequentially realized first for the coarse scale and then for the fine scale. Connections between potential explanatory variables, such as bio-physical or socio-economic variables and the land-use types are assessed by multiple regression analysis. With the help of this set of multiple regressions for all land-use types for both scales, the suitability for each cell for a certain land-use type can be calculated. This suitability ("regression" cover) is compared to the actual cover percentage. Based on this difference the value for the land-use type is changed in an iterative procedure. Competition between land-use types is incorporated in this process if the total cover percentage of all land-use types in a grid cell exceeds the total cell area. In this case the changes in each land-use type are modified corresponding to the competitive strength of each land-use type, which is based on the change in demand and the difference between present cover and "regression" cover. A detailed description can be found in (Verburg, De Koning, Kok, Veldkamp, & Bouma, 1999).

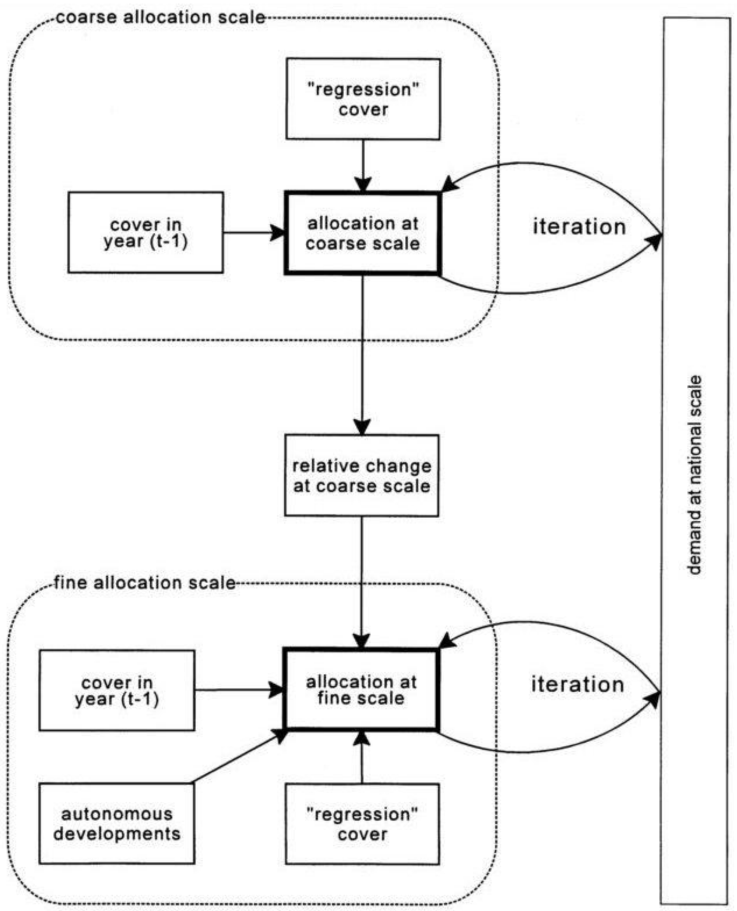

Figure 2-1: Schematic representation of the allocation at two scales (Verburg, De Koning, Kok, Veldkamp, & Bouma, 1999)

The CLUE model has been adapted by (Aguiar, 2006) to be applicable to the Brazilian Amazon. For a better distinction, this model, which has been developed at INPE, is called AmazonClueINPE. Several aspects had to be accounted for in the implementation of the model and are described in the following paragraphs.

Initial modeling decisions involved the definition of spatial and temporal scales and the choice of land-use classes. Forest and five main agricultural land-uses, namely pasture, temporary agriculture, permanent agriculture, planted forest and non-used agricultural areas, serve as dependent variables at a 25x25km^2 and a 100x100km^2 scale. Potential explanatory factors comprise accessibility to markets, economical attractiveness, demographical, technological, agrarian structure, public policies and environmental factors. In addition to regarding the whole Brazilian Amazon, the study area is also subdivided into three macro regions at the fine scale, which allows considering diverse characteristics in different regions. These regions are the Densely Populated Arch, the Central Amazon and the Oriental Amazon. The temporal settings of the model show a time span from 1997 to 2020 with a resolution of one year.

The statistical analysis led to the definition of several alternative models for each region. Log-transformation of the land-use classes and certain potential explanatory variables was used to account for non-linear relationships and thus improved the regression results. Due to correlation some variables could not be used in the same statistical model.

The AmazonClueINPE model uses a modified allocation procedure to account for specific requirements of the study area and to allow analyzing different law enforcement scenarios. When deforestation reaches a certain threshold (*lim_forest*) in a cell, a different allocation algorithm is used. This method allows simulating if the Federal Law is observed or not. The Federal Law states that 80% of forest inside private properties must be preserved. As this law is currently largely disregarded, this threshold can simulate possible law enforcement actions. A second parameter allows controlling the maximum change in a cell in one period of time. This parameter was introduced because the AmazonClueINPE model initially concentrated changes only in a few cells with high suitability for change. With the *max_change* parameter an upper limit for the possible change in each cell in a given period of time is established.

In addition to *lim_forest* and *max_change*, some other adjustable parameters exist. The scale factor (*scale_fact*) gives the possibility to increase the importance of one scale in respect to the other and thus favor one of the two scales. A value of 1 treats both scales likewise. The *max_iter* value defines the maximum number of iterations in the allocation process. The *max_demand_diff* parameter indicates the maximum allowed difference between demand and allocated change. It is defined in terms of the demand.

Table 2-1: Parameters for the AmazonClueINPE allocation module

parameter	description
lim_forest	forest threshold to preserve 20% of cell area from deforestation
max_change	upper limit for change in one period of time
scale_fact	to favor one scale in respect of the other
max_iter	maximum number of iterations
max_demand_diff	maximum allowed difference between demand and allocated change

Various combinations of demand and allocation scenarios are defined to explore the influence of potential land-use determining factors on land-use changes under certain policy conditions and market constraints in the Brazilian Amazon.

The exploration of different statistical models for the three macro-regions led to an important conclusion. The model results indicate that using the statistical model of the Densely Populated Arch (arch25) in all spatial regions produces more realistic spatial patterns of the deforestation process than using regression models from other macro-regions or the whole Amazon. (Aguiar, 2006) states that this model (arch25), due to the inclusion of a distance to roads and a connection to markets measure delivers better results of the AmazonClueINPE model. The author points out that applying the arch25 model also to the other macro-regions should not lead to the assumption that the process in the Arch is likely to happen in other regions in the same way, but that it "captures the current and possible axes of development". Important other variables in this model include a protected areas variable, distance to timber production areas and percentage of fertile soils.

The first version of the AmazonClueINPE model was written in C++ and tested in the before mentioned study (Aguiar, 2006). (Moreira, 2009) implemented the AmazonClueINPE model in the TerraME modeling language, which is part of the TerraME modeling framework and described in section 2.3. The modular implementation allows top-down and bottom-up interactions between multiple scales through the integration of spatial, temporal and analytical couplers. A general description of the two studies was given in section 2.1.2.

2.2 Climatic Models coupling to LUCC models

2.2.1 Overview

Climate affects vegetation, but vegetation also has the potential to affect climate (Cox et al., 2004; Foley, Costa, Delire, Ramankutty, & Snyder, 2003; Oyama & Nobre, 2004). Vegetation models have been developed to investigate these complex bidirectional interactions. According to (Cook & Vizy, 2008) two kinds of vegetation models are currently in use. The first type is the potential vegetation model, which determines vegetation in equilibrium with a given climate. Due to other vegetation type determining factors apart from climate (i.e. topography, soil type, etc.) there is a discrepancy between the spatial distribution of potential and natural vegetation. Nevertheless reasonable agreement between the global distribution of potential and natural biomes at large spatial scales can be reached (Oyama & Nobre, 2004). The second type of vegetation models is the dynamic vegetation model, which is fully interactive and simulates the impact of vegetation on the exchange of moisture, heat and momentum between the atmosphere and the land surface (Cook & Vizy, 2008). Researchers can draw important conclusions about vegetation and climate interactions by coupling both kinds of vegetation models to atmospheric general circulation models (AGCM).

The biophysical environment is continuously altered by human influence, which is in general not represented in dynamic global vegetation models (GLP, 2005). The Lund-Potsdam-Jena managed Land model (Bondeau et al., 2007) builds an exception, as it integrates dynamic land-use at a global scale.

Numerous LUCC models, as discussed in section 2.1, can be used to study the human involvement in land-use change processes. Consideration of anthropogenic impact is a prerequisite for the construction of integrated land system models as conceptualized in (GLP, 2005) or (Schaldach & Priess, 2008). These land system models consist of human and environment sub-systems, which influence each other through land-use and environmental change. Constructing such an integrated land system model, which incorporates biophysical characteristics and biogeochemical cycles, as well as a land-use model as a representation of human decision-making is an ambitious challenge. Due to the complexity of the involved interactions between human and environment subsystems, land system models are still rare in literature (Schaldach & Priess, 2008). Further collaborative work between various research areas is inevitable to successfully couple LUCC models to climate or vegetation models, which would be an important step to further explore the Earth System.

2.2.2 CPTEC-PVM

CPTEC-PVM is a potential vegetation model, developed at INPE and introduced in (Oyama & Nobre, 2004). It comprehends a water balance model to derive water-related quantities from meteorological input data to distinguish between potential vegetation types. In its second generation CPTEC-PVM2 (Lapola, Oyama, & Nobre, 2009) considers CO_2-plant interactions through plant physiological processes and their interactions with the water cycle. The water balance model of CPTEC-PVM2 slightly differs from the first version as it calculates canopy resistance, which is used to estimate evapotranspiration, in terms of net primary productivity and atmospheric CO_2 (Lapola, Oyama, & Nobre, 2009).

Oyama and Nobre coupled the CPTEC-PVM to an AGCM to look for climate-vegetation equilibrium states for Tropical South America under present-day climatic conditions (Oyama & Nobre, 2003). Two equilibria were identified. The first equilibrium state is the current biome distribution. In the second equilibrium state savanna replaces forests in eastern Amazonia and a semi-desert area appears in the driest portion of Northeastern Brazil. The authors point out that tropical Brazil lies under increasing land-use pressure and that deforestation and other land cover and land-use changes could weaken the hydrological cycle in Amazonia and Northeastern Brazil and could tip by itself the climate-vegetation system towards this new drier equilibrium state with savannization and desertification (Oyama & Nobre, 2003). On the contrary (Malhi et al., 2008) point out that resilience of Amazonian forest ecosystems to climatic drying is currently underestimated in vegetation-climate models. According to (Oliveira, Dawson, Burgess, & Nepstad, 2005) drought stress is partly being avoided through hydrological redistribution in tropical forests, which corresponds to the water transfer by roots to drier regions of the soil profile. (Nepstad, Stickler, Soares-Filho, & Merry, 2008) mention that some coupled vegetation-climate models show savannization in parts of the Amazon, but that the majority of these modelling approaches does not support this theory, while noting that effects from factors like fire activity or land-use are not included in these models.

(Salazar, Nobre, & Oyama, 2007) used CPTEC-PVM to asynchronously couple it to fifteen Coupled Ocean-Atmosphere General Circulation Models to reveal the effect of projected climate change on vegetation based on two emission scenarios. The authors compared the projected distribution of biomes to the potential vegetation forced by present-day climate. They mention that though various climate change studies show that climate points towards a warmer future for South America, there is yet uncertainty how rainfall, evapotranspiration and amount of soil water will change due to the changing climate especially in Amazonia and Northeastern Brazil. The authors findings indicate that vegetation in tropical South America will mainly change through the conversion of tropical forest into savanna, due to an increase in dry season length and/or decrease of annual soil moisture, mainly concentrated in southeastern Amazonia. (Salazar, Nobre, & Oyama, 2007).

(Cook & Vizy, 2008) used CPTEC-PVM to asynchronously couple it to a regional atmospheric model to study the effects of twenty-first-century climate change on the tropical and subtropical climate and vegetation of Southern America. The results of the coupled region model simulations indicate a 70% loss of the Amazon rain forest by the end of the twenty-first century with much of the forest being replaced by savanna vegetation and a southward and westward expansion of caatinga vegetation, which refers to the semiarid vegetation of mixed shrubland and grassland that primarily exists in the drought-prone northeastern region, into present day savanna regions.

(Lapola, Oyama, & Nobre, 2009) used CPTEC-PVM2 driven by meteorological input data from fourteen coupled ocean-atmosphere global climate models under two different greenhouse gas emission scenarios to investigate the role of the CO_2 fertilization effect on future biome distribution in South America. Results show that there must be substantial biome shift in the Amazon, including substitution of forest by savanna if the CO_2 fertilization effect does not play a role in tropical ecosystems or if the dry season length exceeds four months. Otherwise, the CO_2 fertilization effect could prevent major biome changes in the Amazon.

In this thesis a water balance model will be used to derive variables that represent wetter or drier climatic conditions, which can serve as input for the LUCC model. The water balance model is a submodel of CPTEC-PVM as described in (Oyama & Nobre, 2004). Although, according to (Lapola, Oyama, & Nobre, 2009), CPTEC-PVM is limited for future climate-vegetation simulations and its successor CPTEC-PVM2 was already available, in this work the water balance model of CPTEC-PVM is used because it outputs two water-related factors which can be better compared to already existing environmental variables of the LUCC model and thus give comprehensible results.

2.2.2.1 Water Balance Model

The water balance model of CPTEC-PVM is based on the one from (Willmott, Rowe, & Mintz, 1985). It produces a consistent global distribution of soil moisture (Oyama & Nobre, 2004) by estimating the water balance over a homogeneous soil layer covered by short grass. Different soil and vegetation types are not taken into consideration. The difference to (Willmott, Rowe, & Mintz, 1985) is that actual evapotranspiration is calculated using the Penman-Monteith equation, instead of Thornthwaite's equation and that the possibility of soil freezing has been included. The water balance model is used to calculate two moisture variables which later on serve as input in the PVM classification algorithm. The wetnessindex (H) is used to distinguish between wet and dry climates and the seasonalityindex (D) to represent the soil moisture seasonality. The variables are defined according to

$$H = \frac{\sum_{i=1}^{12} g_i E_i}{\sum_{i=1}^{12} g_i E_{\max,i}} \qquad D = 1 - \frac{\sum_{i=1}^{12} F(0.5 - w_i)}{6}$$

Formulas 2-1 and 2-2

$$g = \begin{cases} 1, & unfrozen \\ 0, & frozen \end{cases} \qquad F(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Formulas 2-3 and 2-4

where E is the actual and $E_{\max}$ the maximum evapotranspiration, w the soil water degree of saturation (ratio between soil water storage and soil water availability) and i corresponds to the i^{th} month. The detailed formulation of the water balance model can be found in (Oyama & Nobre, 2004).

2.2.2.2 CPTEC-PVM Classification

The two moisture variables serve as input for the algorithm to define the potential biome distribution based on the classification from (Dorman & Sellers, 1989). The three other variables in the PVM classification process are related to temperature. They are the mean temperature of the coldest month (T_C) and the number of growing degree days using a 0°C (G_0) and a 5°C (G_5) threshold. The five input variables for the PVM are calculated for each cell or grid point after every run of the water balance model. The algorithm to obtain the potential biome is shown in Figure 2-2.

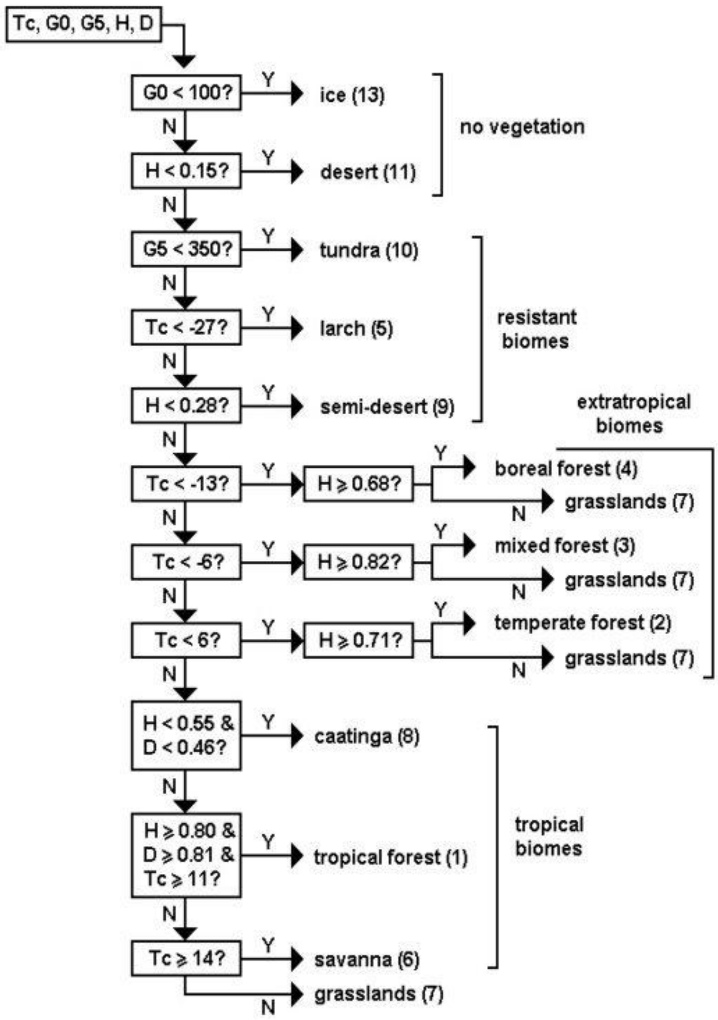

Figure 2-2: Algorithm to obtain the potential biome from environmental variables (Oyama & Nobre, 2004) [4]

Following the potential biome classification algorithm the CPTEC-PVM outputs the current potential vegetation as shown in Figure 2-3.

[4] T_c: temperature of the coldest month, G0: growing degree days with 0°C threshold, G5: growing degree days with 5°C threshold, H: wetnessindex, D: seasonalityindex

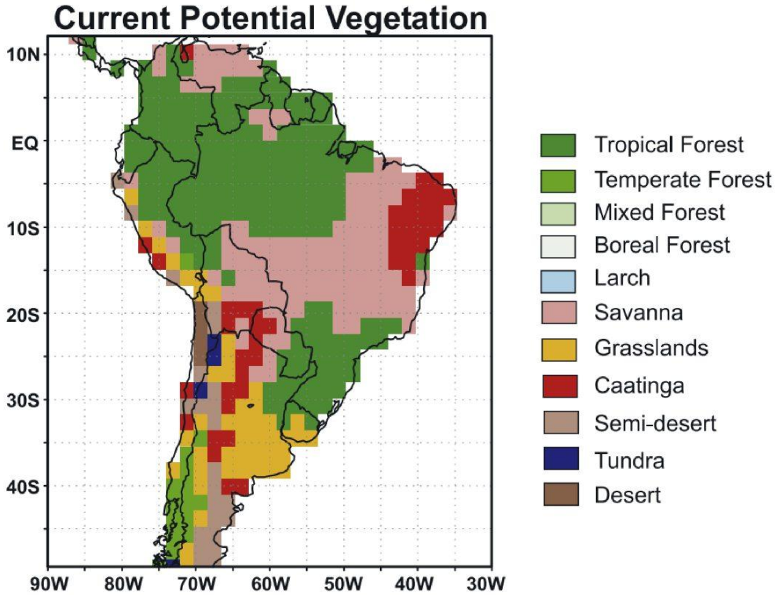

Figure 2-3: Current potential biomes for South America (Salazar, Nobre, & Oyama, 2007)

In this work the biome classification of the CPTEC-PVM is not used. The water balance model is used to derive the two variables wetness- and seasonalityindex (Formulas 2-1 and 2-2) which are added to the other potential determining factors (Table 3-1) which are used in the analysis of land-use and land-cover changes in the Legal Amazon (Chapter 4).

2.3 The TerraME modeling environment

TerraME[5] is a programming environment for spatial dynamical modeling in various application areas (Carneiro, 2006). It is based on TerraLib[6], an open source GIS classes and functions library for large-scale environmental and socio-economic applications (Câmara et al., 2008). TerraME provides a nested cellular automata model and services for spatiotemporal data analysis and management, model development, simulation and assessment. It supports cellular automata, agent-based models and network models (TerraME Website, 2010). Land-use change models and hydrological models belong to the typical applications of TerraME. The modules and services provided by TerraME are visualized in Figure 2-4. It shows a typical TerraME program sequence. A TerraME model can be written in any text editor. The model source code is syntactically checked and executed by the TerraME interpreter, which retrieves the required data from a TerraLib database and stores it afterwards. TerraView[7] can be used to visualize and analyze the data.

[5] www.terrame.org
[6] www.terralib.org
[7] www.dpi.inpe.br/terraview

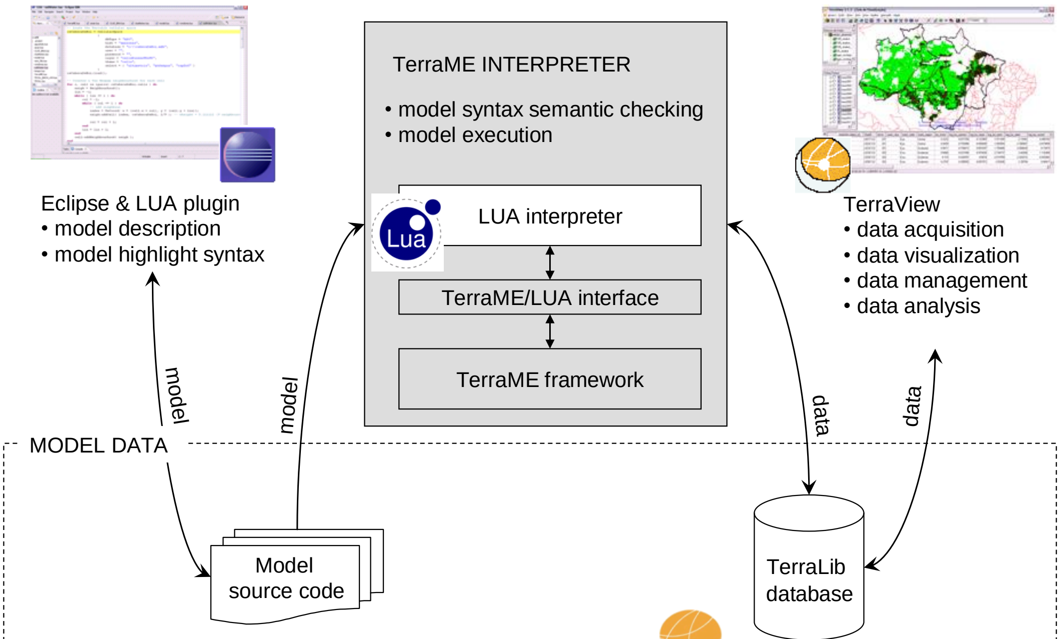

Figure 2-4: TerraME modules and services (Carneiro, 2006)

The TerraME modeling language is an extension of the Lua scripting language (Ierusalimschy, de Figueiredo, & Celes, 1996) and has been designed to allow the development of models in a comprehensible way, also for non-professional programmers.

The current version of TerraME (RC4 for TerraLib 3.2) works under Windows (XP and Vista) and supports Access and MySQL databases. TerraME has been developed as a joint effort among TerraLab (Laboratory for Modelling and Simulation of Land Systems), at Federal University of Ouro Preto, with Image Processing Division (DPI) and Earth System Science Center (CCST), at INPE (TerraME Website, 2010).

TerraView is an open source GIS application based on the TerraLib GIS library for visualization and analysis of geographical data. The software supports various raster and vector data formats. The data is stored in relational or geo-relational databases as ACCESS, PostgreSQL, MySQL or Oracle (TerraView Website, 2010). Additional functionality can be reached by adding TerraView plugins, which are constantly developed by the TerraView community.

aRT[8] (Andrade, Junior, & Fook, 2005) is a R[9] package which provides the integration between the statistical software R and the GIS library TerraLib. Thus it allows accessing and analyzing of geospatial data from a TerraLib database in R.

[8] www.leg.ufpr.br/doku.php/software:art

TerraME and TerraLib are freely available under the GNU Lesser General Public License, TerraView, R and aRT under the GNU General Public License.

[9] www.r-project.org

3 Methods

In this chapter the study area is introduced and the methods are described, followed by sections about the CPTEC-PVM implementation and Land-use change modeling, including database construction, statistical analysis and the dynamic land-use model AmazonClueINPE.

3.1 Study area

The study area is the Brazilian Amazon (Legal Amazon, Amazônia Legal) which consists of the states Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima, Tocantins and a part of Maranhão. The Legal Amazon comprises an area of approximately 5 million km^2 (58% of Brazilian territory) and was initially introduced for regional planning purposes (Andersen, Granger, Reis, Weinhold, & Wunder, 2002). It can be divided into three macro regions as defined by (Becker, 2005). The first macro region is the Densely Populated Arch, where most of the high population density areas, roads and centers of economy are. The second is the Central Amazon, which is the most vulnerable area in the Amazon, because of major roads crossing its interior. The third macro region is called Occidental Amazon, which is the most preserved area, because it is not cut by main roads and people settled to a large extent only in the area around Manaus, while the rest of the region remained mostly abandoned.

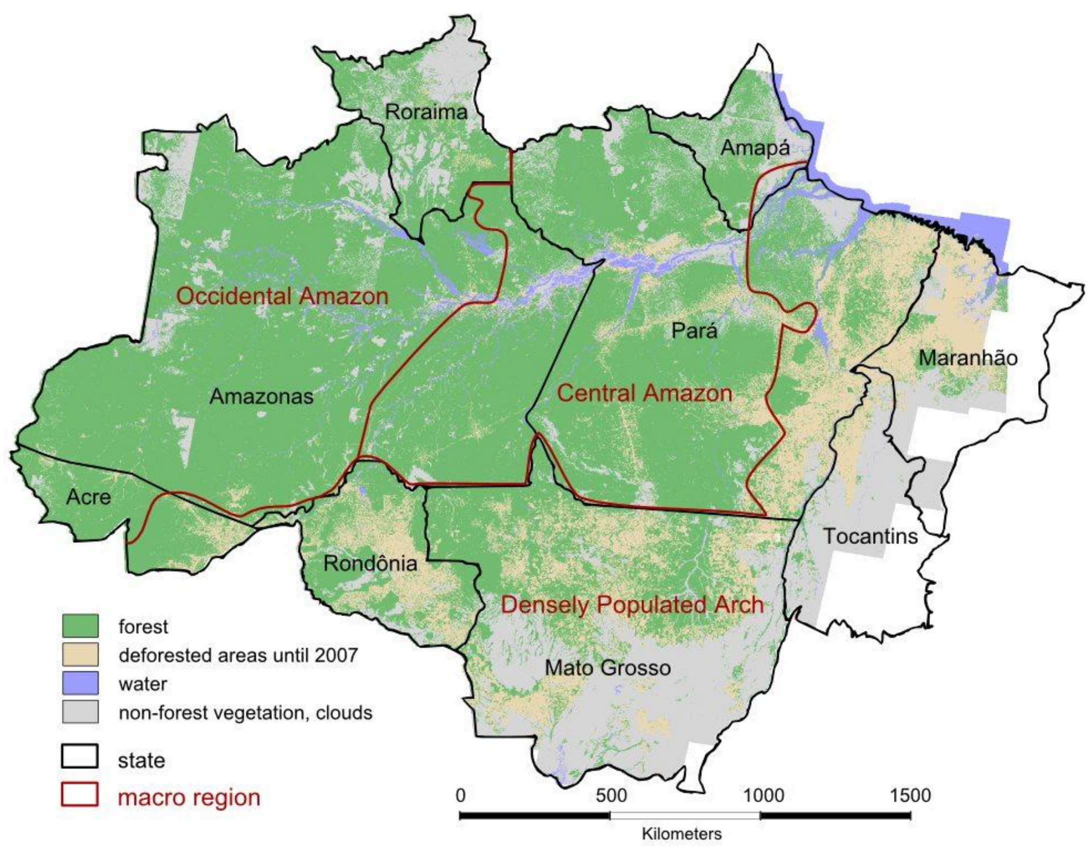

Figure 3-1: The Brazilian Amazon and its three macro regions (Becker, 2005; INPE, 2010)

The study area is subdivided into regular cells at two spatial resolutions. The cells have an extent of $100x100km^2$ at the coarse scale and $25x25km^2$ at the fine scale. Based on a deforestation map from 1997 derived by INPE through the PRODES project (INPE, 2010) cells with a large amount of non-forest vegetation or mainly covered by clouds are excluded from further analysis (Aguiar, 2006). This results in the generation of 5228 cells at the fine scale and 363 cells at the coarse scale. The amount of deforestation in each cell is taken from the deforestation map (INPE, 2010), while the proportion of the different land-use types in each cell is computed from IBGE Agricultural Census 1996 (IBGE, 1996). The detailed process can be found in (Aguiar, 2006).

3.2 CPTEC-PVM implementation

The CPTEC-PVM and the corresponding water balance model were implemented in the TerraME modeling language based on the original version (Oyama & Nobre, 2004) written in FORTRAN 77. The general description can be found in section 2.2.2. The full mathematical description is available in (Oyama & Nobre, 2004)[1]. Although only the

[1] (Oyama, 2005) changed some threshold values in the biome classification algorithm after the publication of this paper, which lead to the generation of CPTEC-PVM v2. The implementation in TerraME and the

results of the water balance model are used in this thesis, the PVM was fully implemented to allow for future coupling to other models written in the TerraME modeling language.

The input data for the potential vegetation model, namely the mean values per month for precipitation, surface temperature and surface pressure as well as a land-sea mask are obtained from the climate data archive "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999) (V 1.02)" (Willmott & Matsuura, 2001). A preliminary routine stores it in the database to have simple access to the data. The result of this step is that the data can easily be accessed and processed within the TerraME environment and be visualized with TerraView.

The source code is split up into four modules to provide a comprehensible structure.

main.lua sets the directories, loads the three modules and starts the simulation by executing the TerraME environment *env*.

Code 3-1: main.lua

```
-- set directories
----------------------------------------
      DIR = "projects\\pvm\\source\\"
      DATABASEDIR = "database\\"

-- load files
----------------------------------------
      dofile (DIR.."func.lua")
      dofile (DIR.."wbm.lua")
      dofile (DIR.."pvm.lua")

-- run simulation
----------------------------------------
      defBiome:build();
      env:add(cs);
      env:add(defBiome);
      env:execute(1);
```

wbm.lua incorporates the water balance model. From monthly meteorological data of precipitation, surface temperature and surface pressure it calculates the two water-related environmental variables wetness- and seasonalityindex, which are used as potential land-use determining factors in the LUCC analysis. In addition to these two variables the three other variables needed for the biome classification algorithm, namely the temperature of the coldest month, the growing degree days with 0°C threshold and the growing degree day with 5°C threshold are as well calculated in this module. The module consists of functions

screenshot in this chapter correspond to this version (v2), which results in minor differences between the figure from the original paper (v1, Figure 4-2) and the screenshot from TerraView (v2, Figure 4-1).

for the water balance model, the surface water budget for a month, evapotranspiration and runoff calculations.

pvm.lua comprehends the definition of the cellular space (Code 3-2) and the automaton to calculate the potential biome number of each cell on basis of the environmental variables delivered by the water balance model module.

Code 3-2: pvm.lua: defining the cellular space

```
-- define cellular space
      cs = CellularSpace{
            dbType = "ADO",
            host = "localhost",
            database = DATABASEDIR.."amazonia.mdb",
            user = "",
            password = "",
            layer = "cells25",
            theme = "cells25",
      }
      cs:load();
```

The biome classification algorithm is implemented as an Automaton, which consists of a state for each biome. These states are composed of Jump and Flow conditions. If the Jump condition applies to the cell it is send to another State. If it does not apply, the Flow condition is executed, which assigns the corresponding value to the potential biome number variable (*bpot*).

Code 3-3: pvm.lua: defining the automaton (extract)

```
-- define automaton
      defBiome = Automaton {
            it = Trajectory {
                  cs,
                  -- only cells that contain data
            function (cell) return true;
                  end,
                  },
...
            State {
                  id = "tropical forest",
                  Jump {
                        function (event, automaton, cell)
                              return cell["weti"] < 0.84 or
                                    cell["seai"] < 0.86 or
                                    cell["tmin"] < 11;
                        end,
                        target = "savanna"
                  },
                  Flow {
                        function(event, automaton, cell)
                              cell.bpot = 1;
                        end
                  }
            },
            State {
```

```
                   id = "savanna",
                   Jump {
                         function (event, automaton, cell)
                               return cell["tmin"] < 14;
                         end,
                         target = "grasslands"
                   },
                   Flow {
                         function(event, automaton, cell)
                               cell.bpot = 6;
                         end
                   }
             }
...
}
```

The environment *env* is defined in the PVM module. It defines a timer which includes only one event. This event first calls the water balance model module for each cell to calculate the environmental variables and then executes the *defBiome* automaton to calculate the potential biome numbers. In the last step the five variables and the potential biome number are stored into the new table *env_var* in the database.

Code 3-4: pvm.lua: defining the environment

```
-- define environment
      env = Environment{
      id = "env",
      -- define timer
      time = Timer{
            Pair{
                  Event{ time = 1, period = 1, priority = 0},
                  Message {
                        function (event)
                              -- run water balance model
                              for i, cell in pairs(cs.cells) do
                                    wbm(cell);
                              end;
                              cs:synchronize();

                              -- run potential vegetation model
                              defBiome:setTrajectoryStatus(true);
                              defBiome:execute(event);
                              cs:synchronize();
                              cs:save(event:getTime(), "env_var",
{"tmin",
                              "gdd0", "gdd5", "weti", "seai", "bpot"})
                        end
                  }
            },
      }
      }
```

The functions module (func.lua) includes basic auxiliary routines.

The potential biome (bpot) and the five variables needed in the biome classification algorithm, namely wetnessindex (weti), seasonalityindex (seai), the mean temperature of the coldest month (tmin) and the number of growing degree days using a 0°C (gdd0) and a 5°C (gdd5) threshold are stored in the database. In this thesis only the wetnessindex and seasonalityindex will be used for further analysis. The results of the CPTEC-PVM implementation are presented in section 4.1.

3.3 Land-use change modeling

3.3.1 Database Construction

3.3.1.1 Land-use classes

Basically two land-use types are used in this work: forest and deforested areas. For further analysis the class deforested areas is divided into five different subclasses. These classes are pasture, temporary agriculture, permanent agriculture, non-used agricultural land and planted forest. Every cell contains the proportion of the area covered by each land-use class inside the cell by total cell area, thus the values of the six land-use classes in each cell sum up to a value of 1.

Forest

The land-use class forest consists of all areas that are classified as primary forest by the PRODES project.

Deforested Areas

This class comprehends all deforested areas detected by INPE until 1997. The PRODES project (INPE, 2010) detects clear-cut areas greater than 6.25 ha. A short description of the land-use types based on definitions from the Census of Agriculture 1996 (IBGE, 1996) follows.

Pasture

The land-use class pasture includes all areas defined as planted pasture. These areas are especially cultivated for cattle ranching. In 1996 approximately 70% of deforested areas fell into this class, hence it was the major land-use type after forest.

Temporary agriculture

The class temporary agriculture includes areas used for planting or being prepared for planting short-term crops, which require new seeding after each harvest (rice, manioc, maize, soybeans, sugarcane etc.). Areas which have previously been used for planting short-term crops, but have not been utilized for no longer than four years, fall also into this class.

Permanent agriculture

The class permanent agriculture includes areas used for planting or being prepared for planting long-term crops, which keep producing during successive years without the necessity of new seeding (cacao, coffee, cotton etc.). Also nurseries of permanent-crop seedlings fall into this class.

Non-used agricultural land

The class non-used agricultural land summarizes areas that are abandoned or fallow, as they have not been used for a period of more than four years, although being suitable for crops, pasture or woods.

Planted forest

This land-use class summarizes areas which are cultivated or being prepared for planting trees like black acacia, eucalyptus, pine etc.. Also seedling nurseries of forest essences fall into this category.

3.3.1.2 Spatial land-use patterns

Deforested Areas

According to measurements by INPE (INPE, 2010) around 180.000 km^2 of forest have been cut down between 1997 and 2006. The major part of deforestation took place in the Densely Populated Arch in the states of Mato Grosso, Pará, Rondônia and Maranhão. Figure 3-2 shows the deforestation map of 1997 and Figure 3-3 the relative deforestation map from 1997 to 2006.

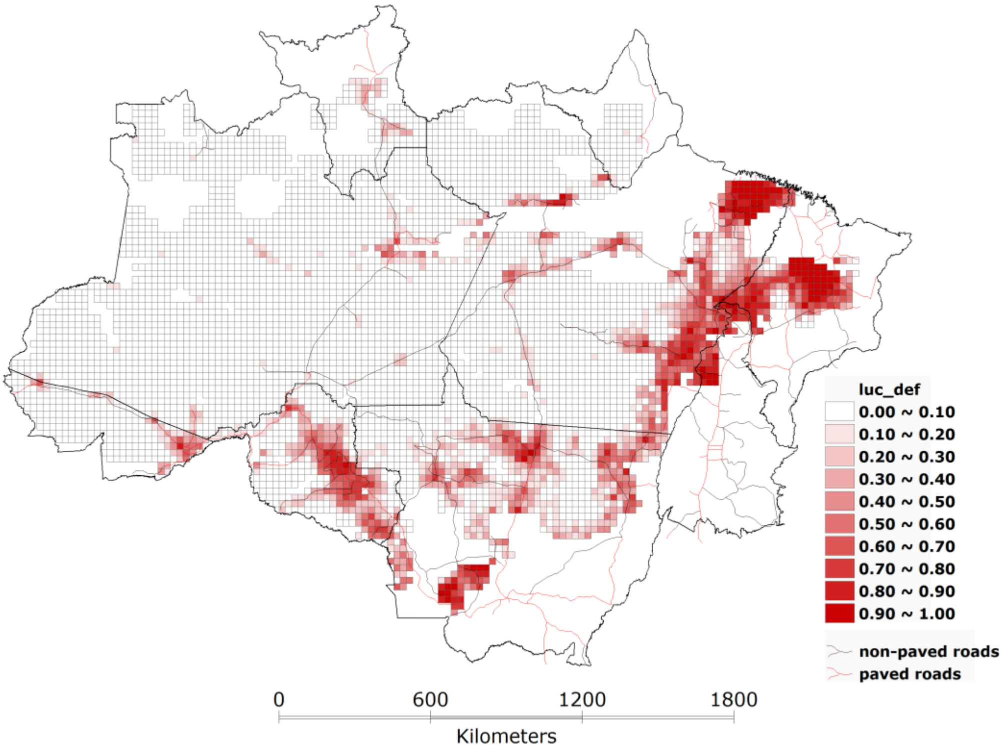

Figure 3-2: Deforestation in 1997 (INPE, 2010)

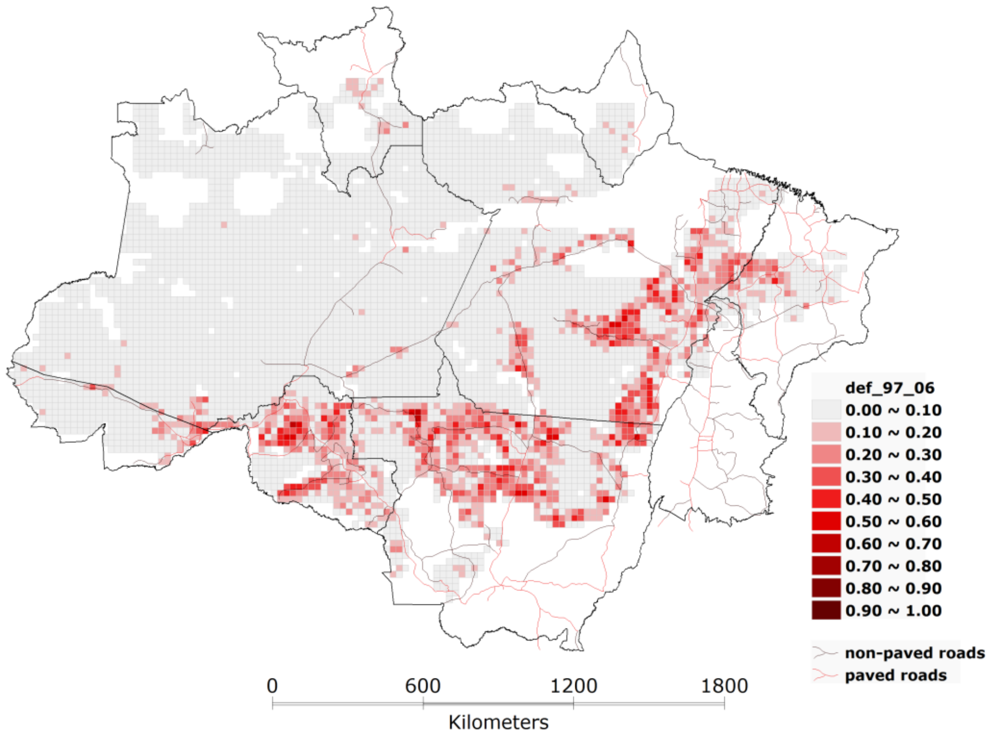

Figure 3-3: Deforestation from 1997 to 2006 (INPE, 2010)

Pasture and temporary agriculture

Pasture is the major non-forest land-use type in the Amazon and accounted for almost 70% of deforested areas in 1997. Hence the pattern in 1997 (Figure 3-4) shows high similarity to the deforestation map. Figure 3-5 shows the relative change of pasture between 1997 and 2006 on basis of the Agricultural Census data from IBGE (IBGE, 1996; IBGE, 2009). A general increase of pasture can be seen in the Densely Populated Arch in the states of Pará, Mato Grosso, Rondonia and Maranhão, since usage as pasture is the most important land-use type for deforested areas. There is also a clear pasture increase along the Transamazônica (BR-230) and along the BR-163 in Pará. There are only a few cells in Mato Grosso and Pará with a decrease in pasture.

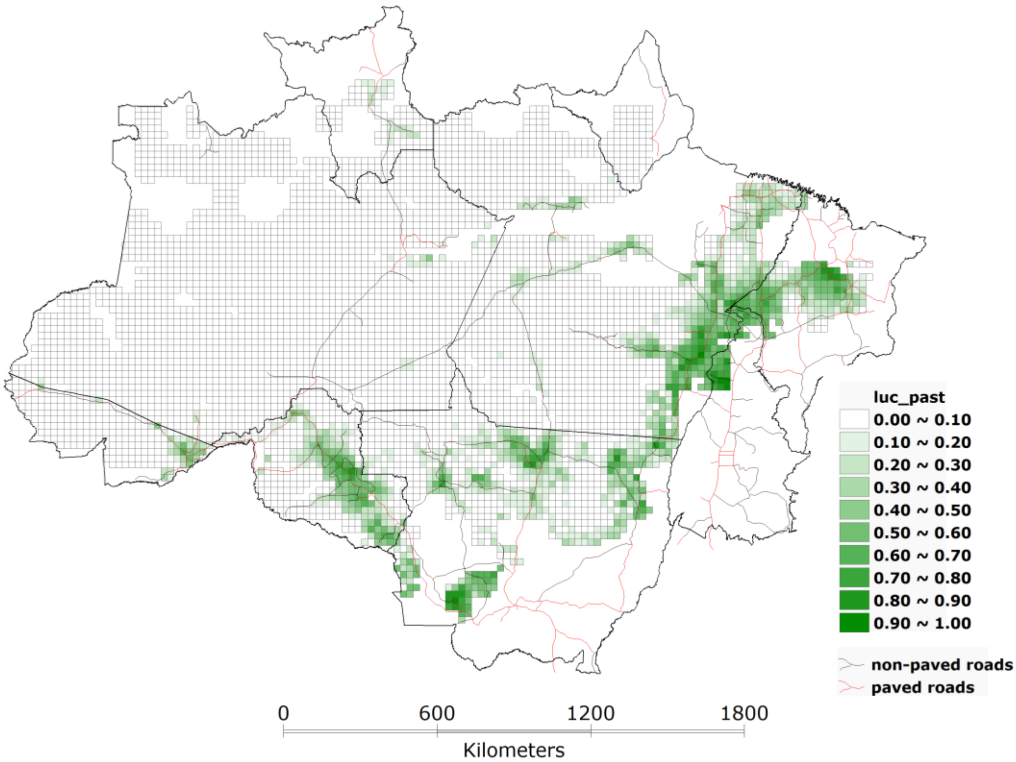

Figure 3-4: Pasture in 1997 (IBGE, 1996; INPE, 2010)

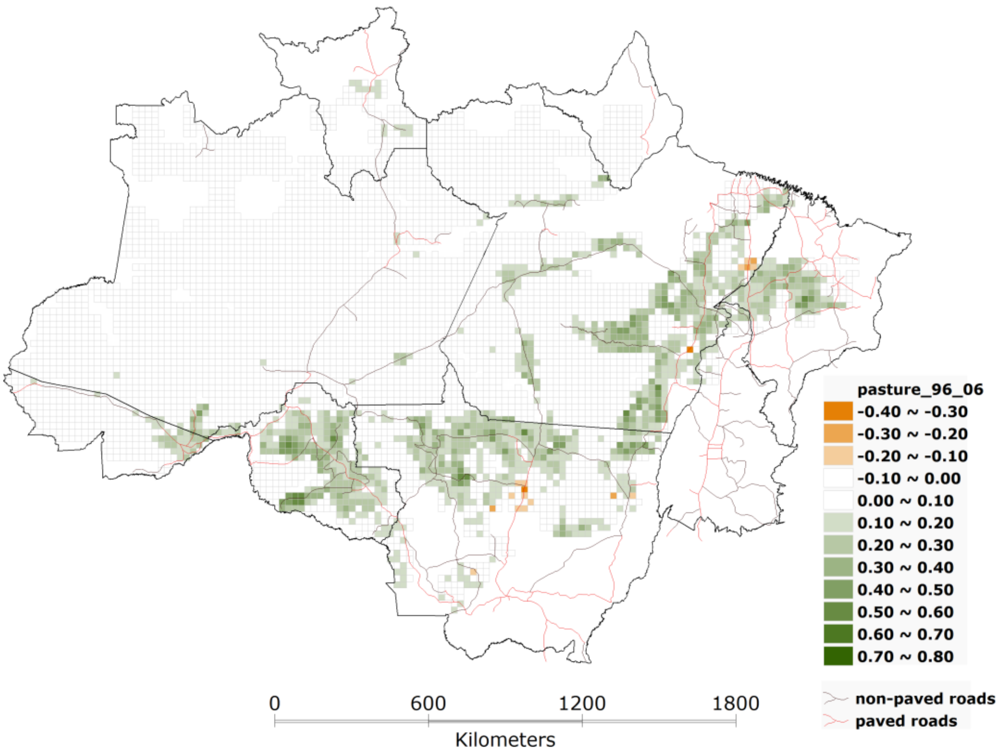

Figure 3-5: Pasture from 1997 to 2006 (IBGE, 1996; IBGE, 2009; INPE, 2010)

Temporary agriculture patterns add up to 14% of deforested area in 1997 (Figure 3-6). Consideration of this land-use class is important to detect land-use change processes related to the cultivation of temporary crops, e.g. soybeans. The analysis of this land-use type in the period of 1997 to 2006 (Figure 3-7) shows two distinct processes, a pattern of decreasing temporary agriculture in Maranhão and an increasing pattern in Mato Grosso. This increasing pattern in the south of the Brazilian Amazon reflects the expansion of mechanized agriculture in central Mato Grosso, which might also explain some of the decreasing pasture values in this area.

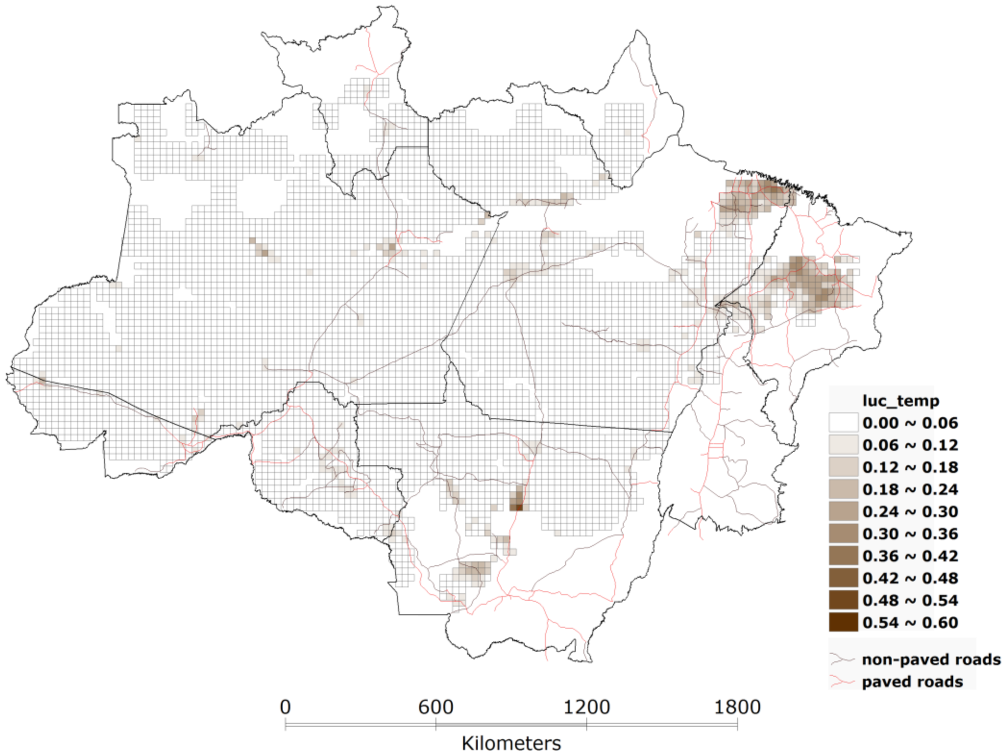

Figure 3-6: Temporary agriculture in 1997 (IBGE, 1996; INPE, 2010)

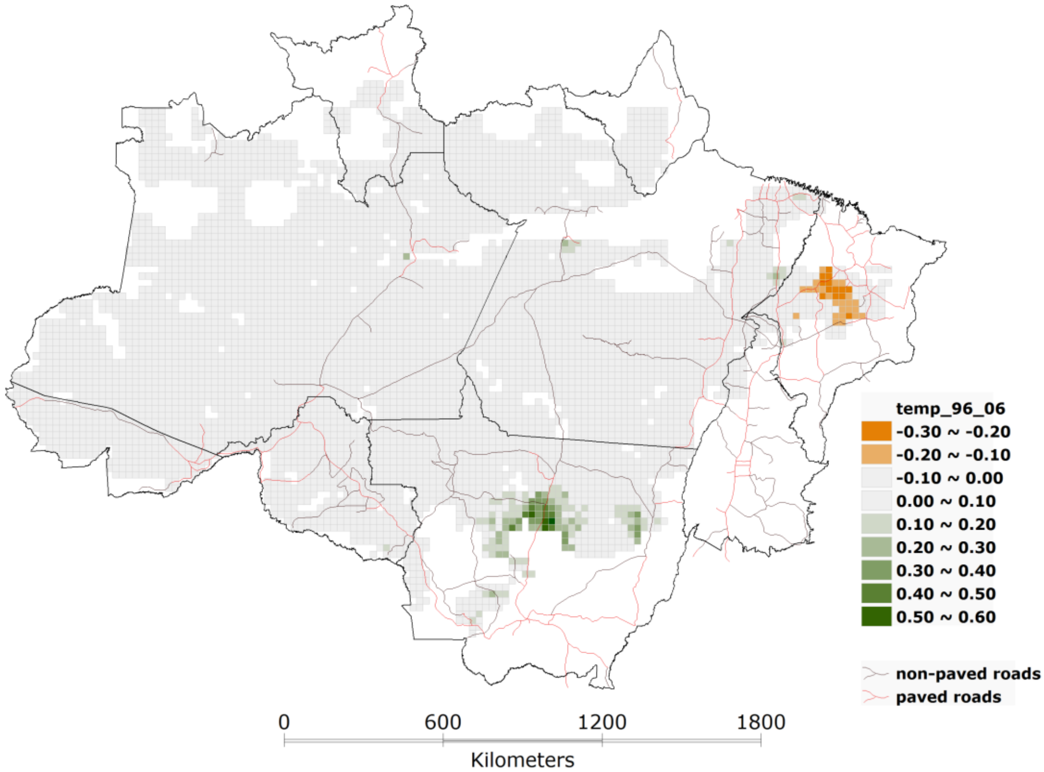

Figure 3-7: Temporary agriculture from 1997 to 2006 (IBGE, 1996; IBGE, 2009; INPE, 2010)

3.3.1.3 Potential land-use determining factors

The potential land-use determining factors used in this work come from a pool of variables defined in (Aguiar, 2006) and can be grouped into different classes. The six classes are Accessibility, Economic Attractiveness, Public Policies, Demographics, Agrarian Structure and Environment. The variables derived from the Water Balance Model of the CPTEC-PVM – wetnessindex and seasonalityindex – and the slope and altimetry variables are added to the Environment group. Some explanatory variables – as well as the land-use variables – are logarithmic transformed to account for non-linear relationships between them. The variables are listed in Table 3-1. The socioeconomic factors are visualized in Figure 3-8 to Figure 3-20 and the environmental factors in Figure 3-21 to Figure 3-29 for the scale 25x25 km^2.

Table 3-1: potential land-use determining factors (adapted from (Aguiar, 2006))

Category	Variablename	Description	Unit	Source
Accessibility	log_dist_urban_areas	euclidean distance to urban centers (log)	km (log)	IBGE
	log_dist_roads	euclidean distance to roads (log)	km (log)	IBGE
	log_dist_paved_roads	euclidean distance to paved roads (log)	km (log)	IBGE
	log_dist_non_paved_roads	euclidean distance to non paved roads	km (log)	IBGE
	log_dist_large_rivers	euclidean distance to large rivers	km (log)	IBGE
	conn_markets	indicator of strength of connection to national markets through roads network	-	IBGE
	conn_sp	indicator of strength of connection to São Paulo through roads network	-	IBGE
	conn_ports	indicator of strength of connection to ports through roads network	-	IBGE
Economic Attractiveness	log_dist_wood_extr_poles	Euclidean distance to wood extraction poles (log)	km (log)	IBAMA
	log_dist_min_deposits	euclidean distance to mineral deposits (log)	km (log)	CPRM
Public policies	prot_all1	percentage of protected areas	% of cell area	IBAMA FUNAI
Demographics	log_pop_dens_96	population density in 1996	people/km^2	IBGE
	log_setl_nfamilies_70_99	number of settled families from 1970 to 1999 (log)	number of families (log)	INCRA
Agrarian Structure	agr_area_small	percentage of area of small properties	% of cell area	IBGE
Environment	soils_fert_B1	percentage of high and medium to high fertility soils	% of cell area	IBGE
	soils_fert_B3	percentage of wetland soils	% of cell area	IBGE
	clima_humi_min_3_ave	average humidity in the three drier months of the year	%	INMET

	weti	wetnessindex	-	CPTEC-PVM
	seai	seasonalityindex	-	CPTEC-PVM
	altitude_avg	average elevation	m	SRTM
	slope_flat	percentage of flat areas (0°-5°)	% of cell area	SRTM
	slope_mod	percentage of moderately sloped areas (5°-15°)	% of cell area	SRTM
	slope_steep	percentage of steeply sloped areas (>15°)	% of cell area	SRTM

Socioeconomic factors

Accessibility

This category comprises factors describing the accessibility of a given cell. Variables describe the Euclidean distance to the closest road, urban center or large river and indicate the strength of connection to national markets or ports through the roads network.

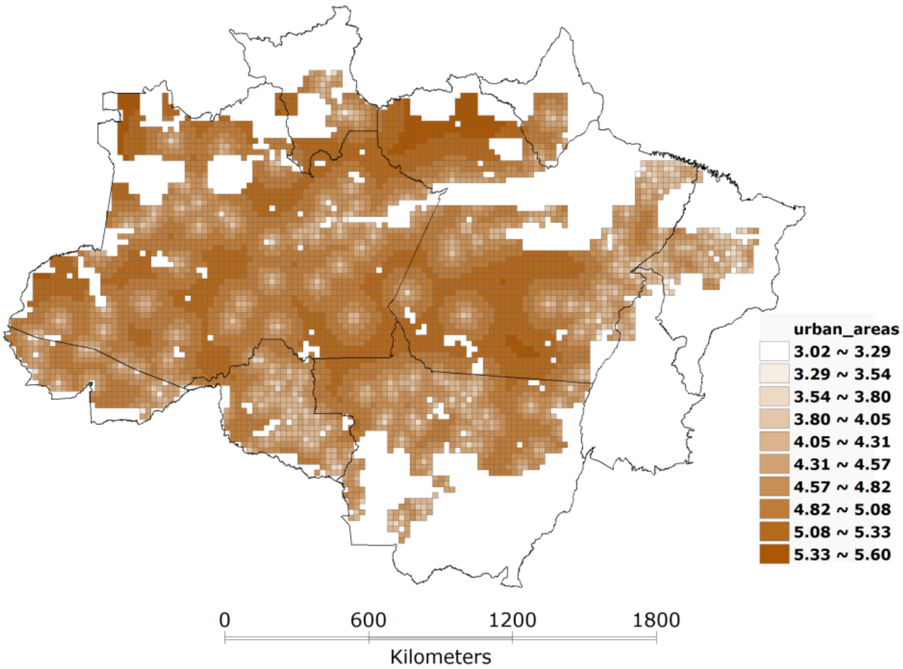

Figure 3-8: distance to urban areas (log_dist_urban_areas)

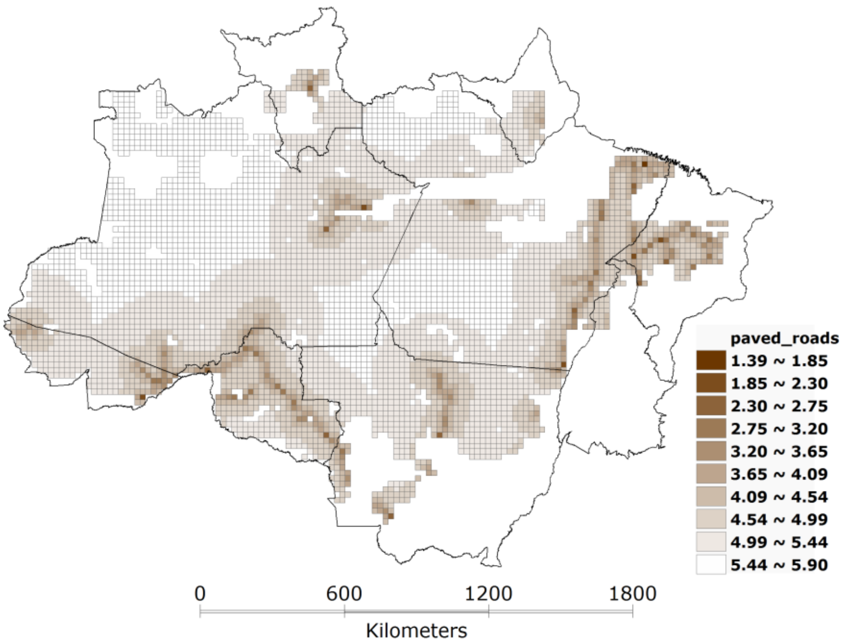

Figure 3-9: distance to paved roads (log_dist_paved_roads)

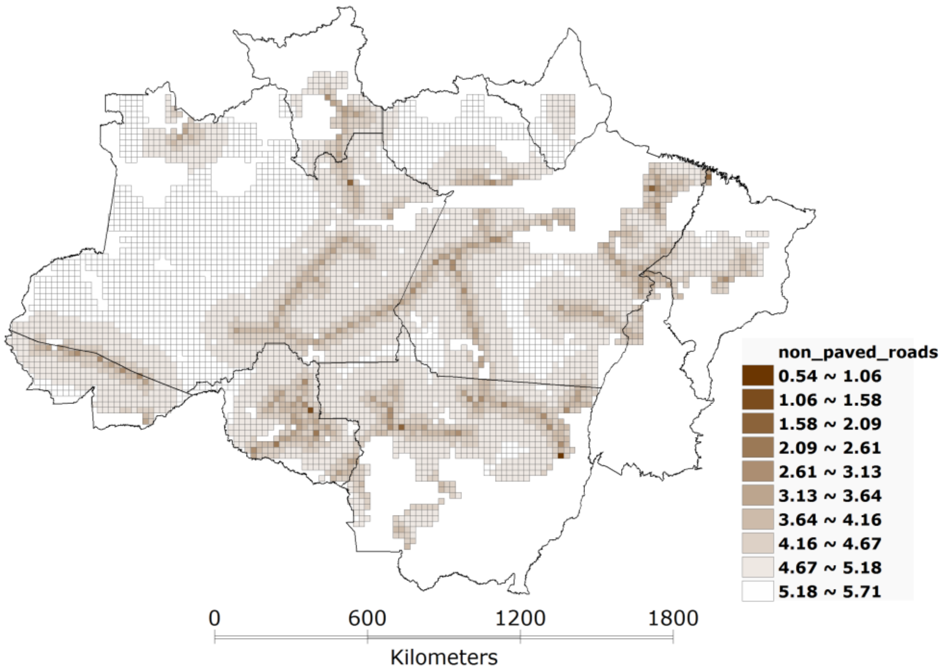

Figure 3-10: distance to non-paved roads (log_dist_non_paved_roads)

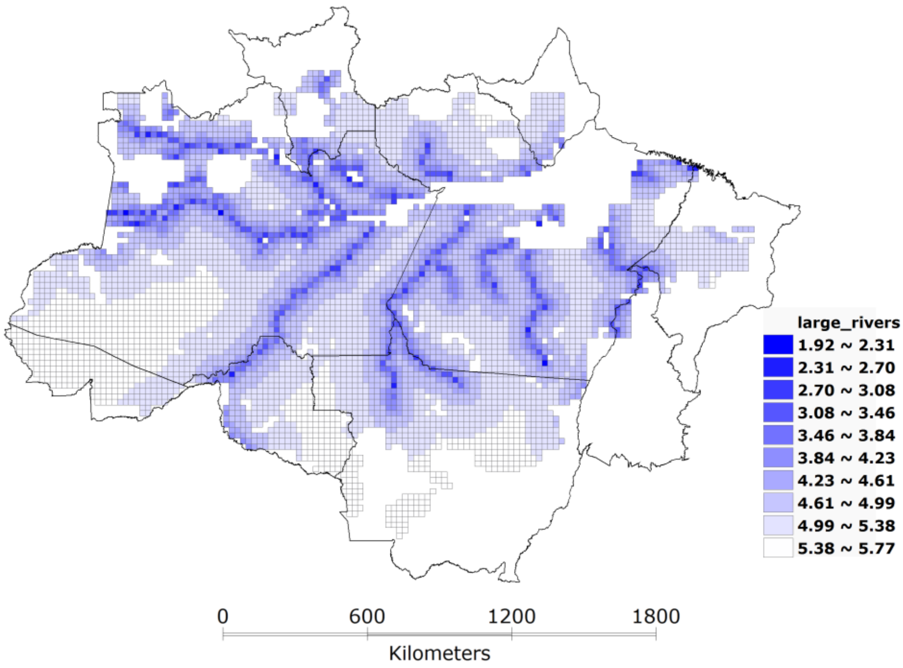

Figure 3-11: distance to large rivers (log_dist_large_rivers)

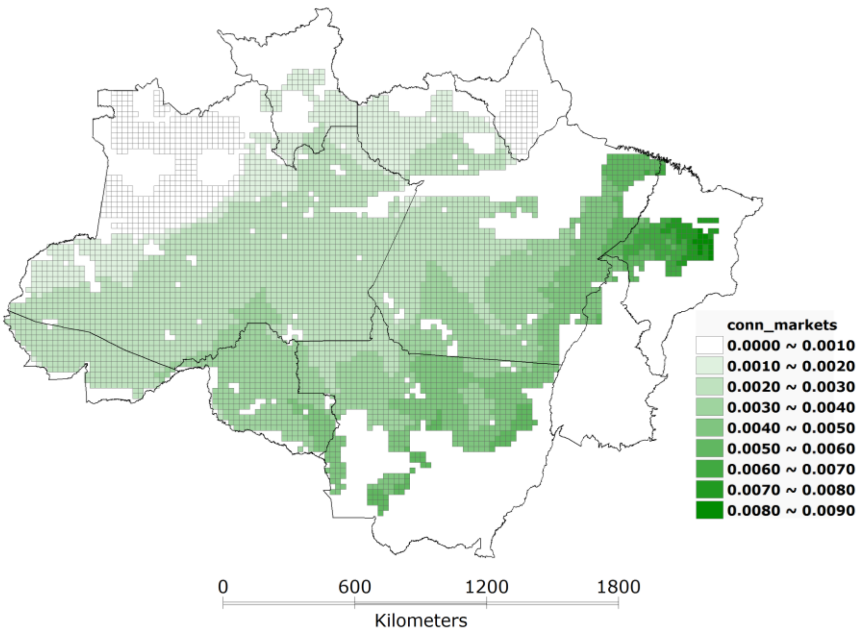

Figure 3-12: connection to markets (conn_markets)

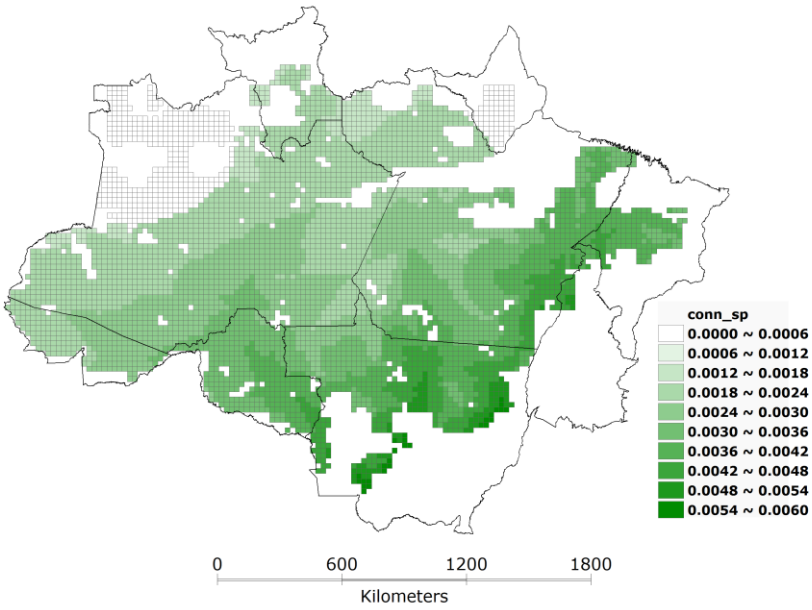

Figure 3-13: connection to São Paulo (conn_sp)

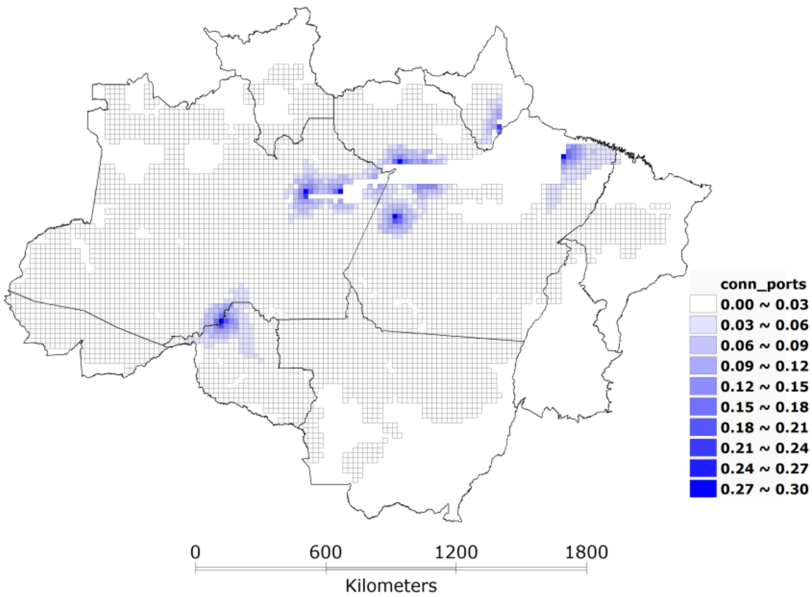

Figure 3-14: connection to ports (conn_ports)

Economic Attractiveness

These factors determine economic attractiveness through the distance to wood extraction poles and mineral deposits.

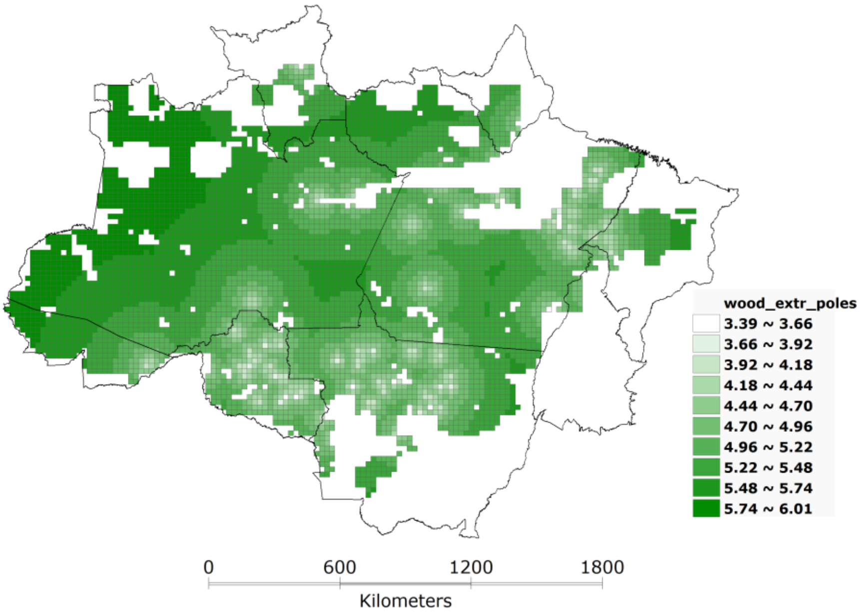

Figure 3-15: distance to wood extraction poles (log_dist_wood_extr_poles)

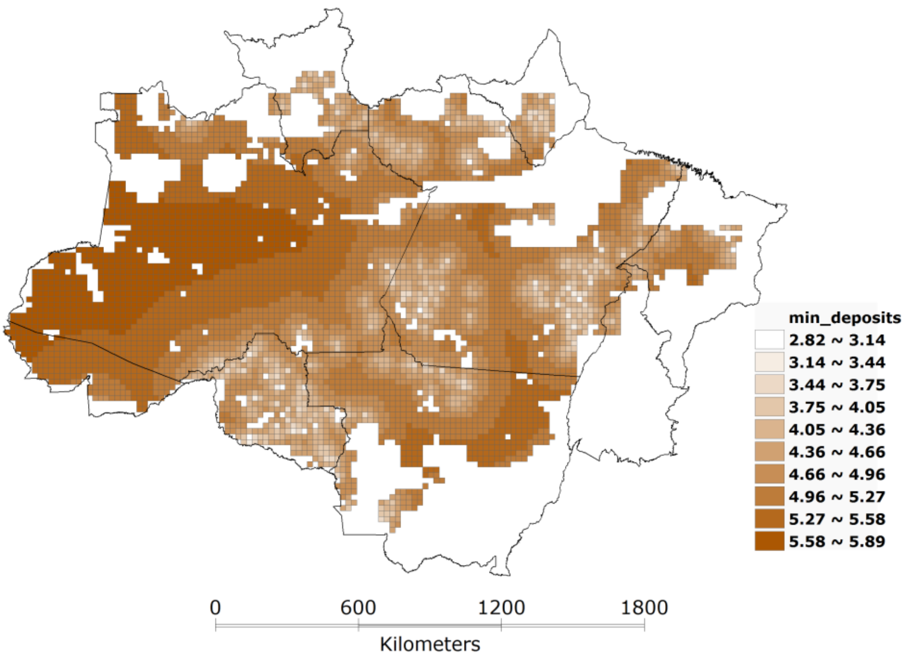

Figure 3-16: distance to mineral deposits (log_dist_min_deposits)

Public Policies

Various areas in the Brazilian Amazon are declared indigenous land, nature reserve or in some other way protected, which is represented by a protected area variable.

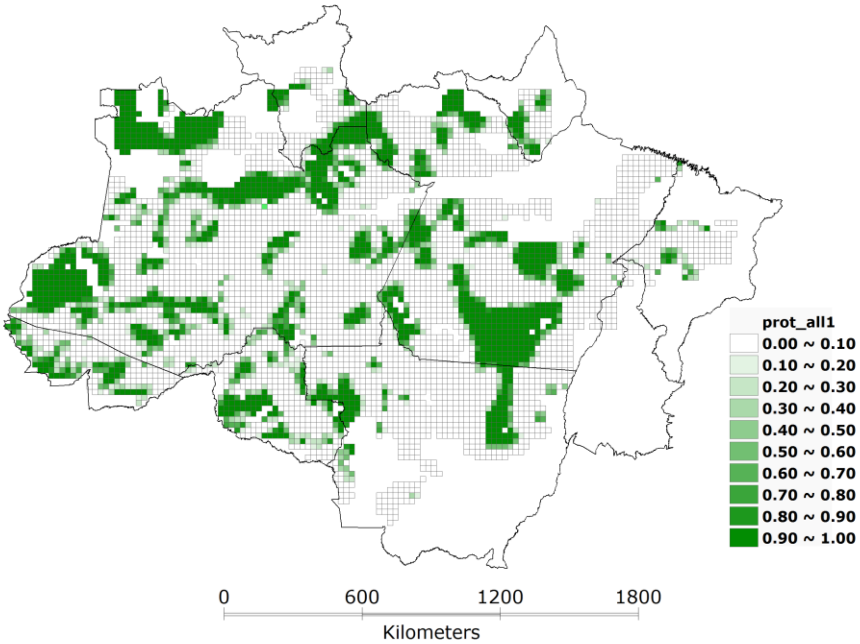

Figure 3-17: protected areas (prot_all1)

Demographics

These variables describe the demographic structure in a cell, from population density to the number of settled families.

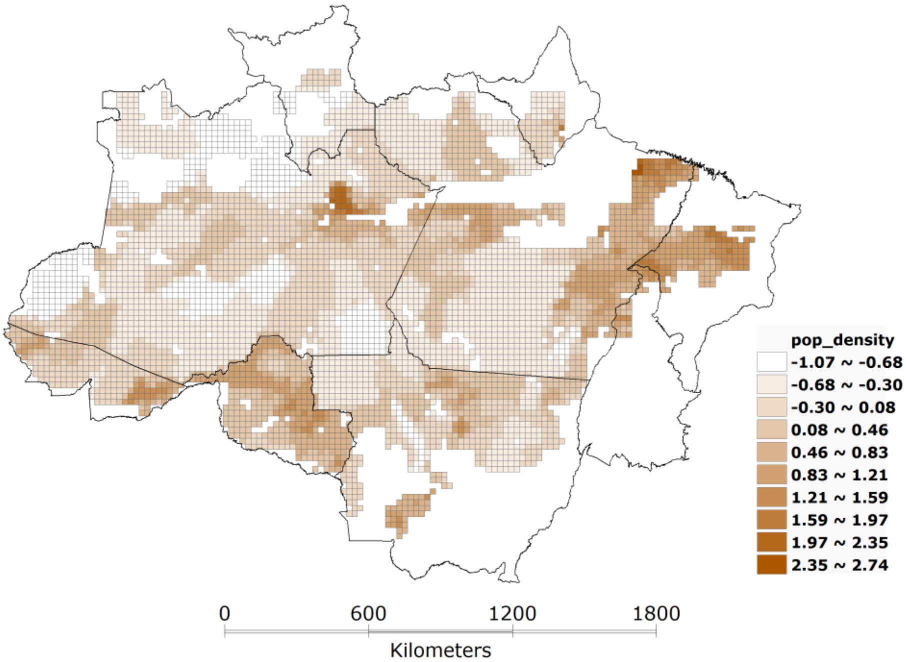

Figure 3-18: population density (log_pop_dens_96)

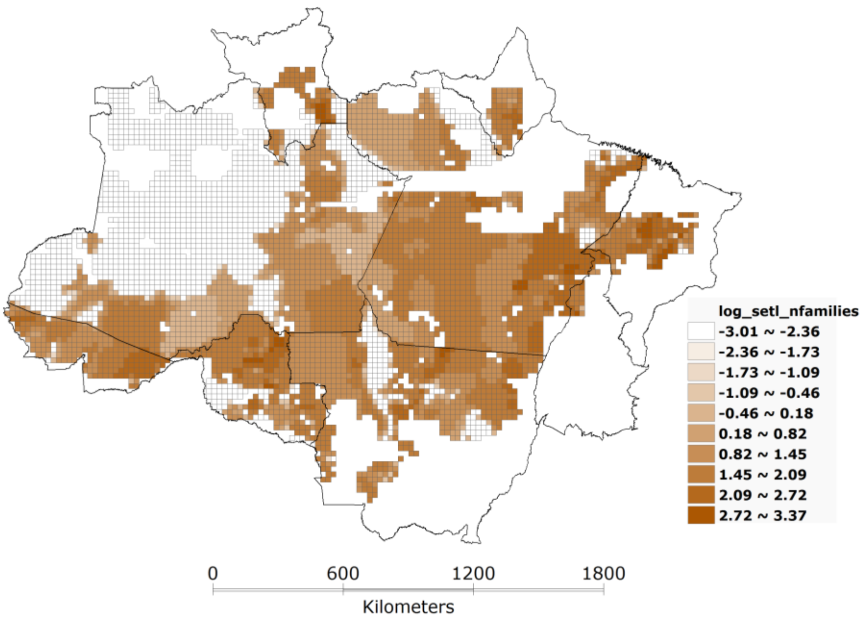

Figure 3-19: number of settled families (log_setl_nfamilies_70_99)

Agrarian Structure

The agrarian structure is represented by a variable summarizing the percentage of small properties in a cell.

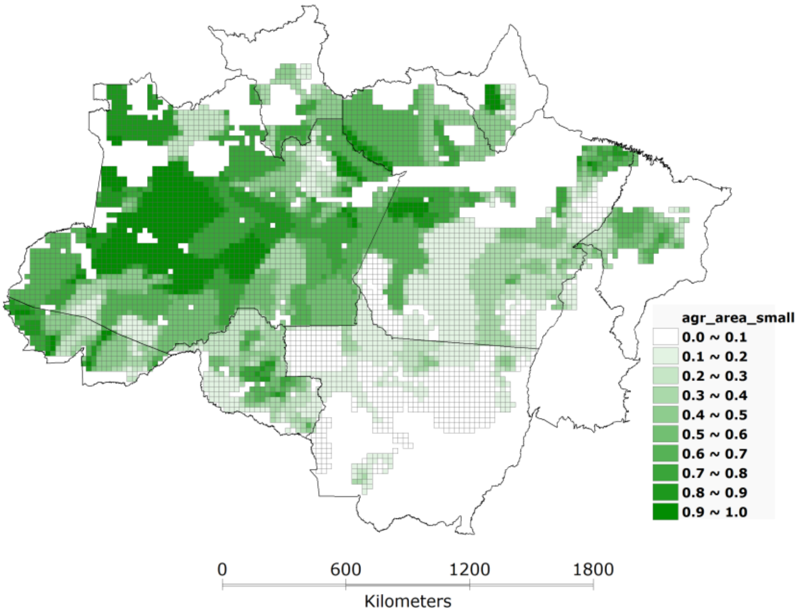

Figure 3-20: areas of small properties (agr_area_small)

Environmental factors

The environmental factors are composed of variables describing bio-physical characteristics (soil fertility, moisture), climatological conditions (humidity, wet-dry climate) and topographic properties (altitude, slope).

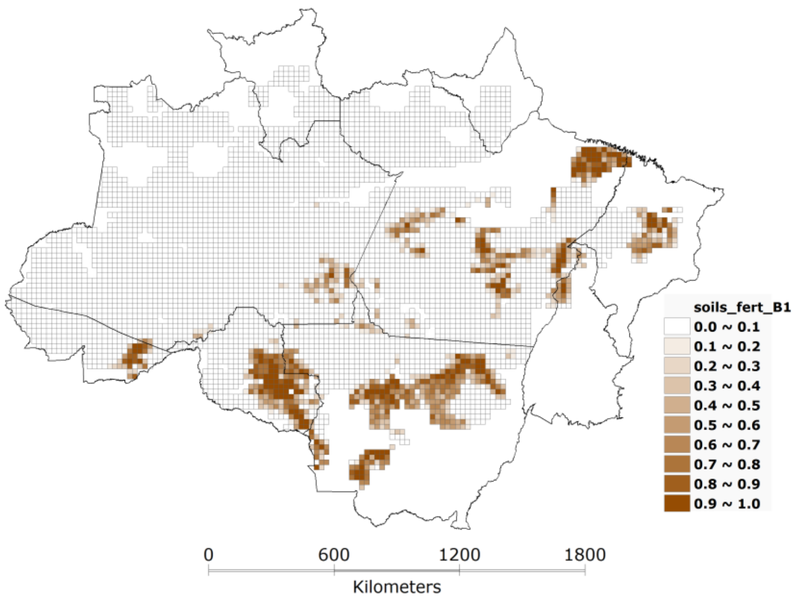

Figure 3-21: fertile soils (soils_fert_B1, %)

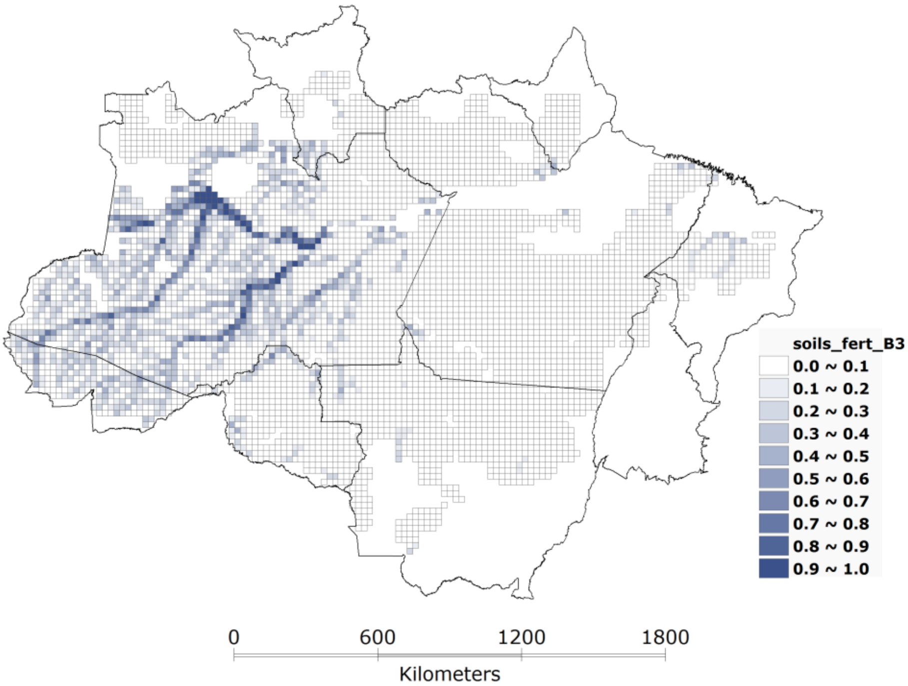

Figure 3-22: wet soils (soils_fert_B3, %)

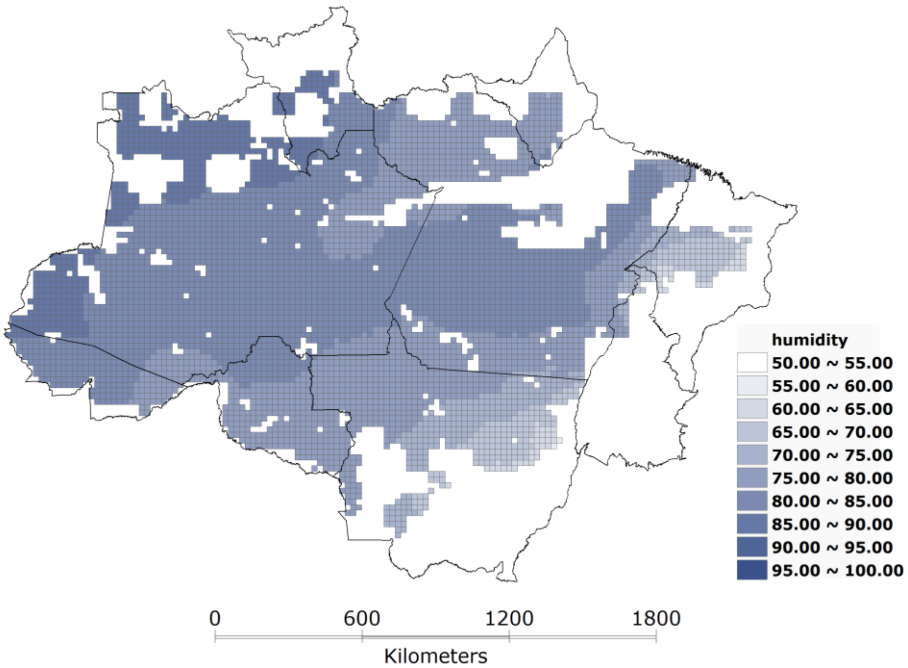

Figure 3-23: humidity (clima_humi_min_3_ave, %)

Environmental factors from CPTEC-PVM

These two environmental variables are derived by the water balance model of the CPTEC-PVM. The wetnessindex relates to soil wetness to distinguish between wet and dry climate conditions, the seasonalityindex represents soil moisture seasonality.

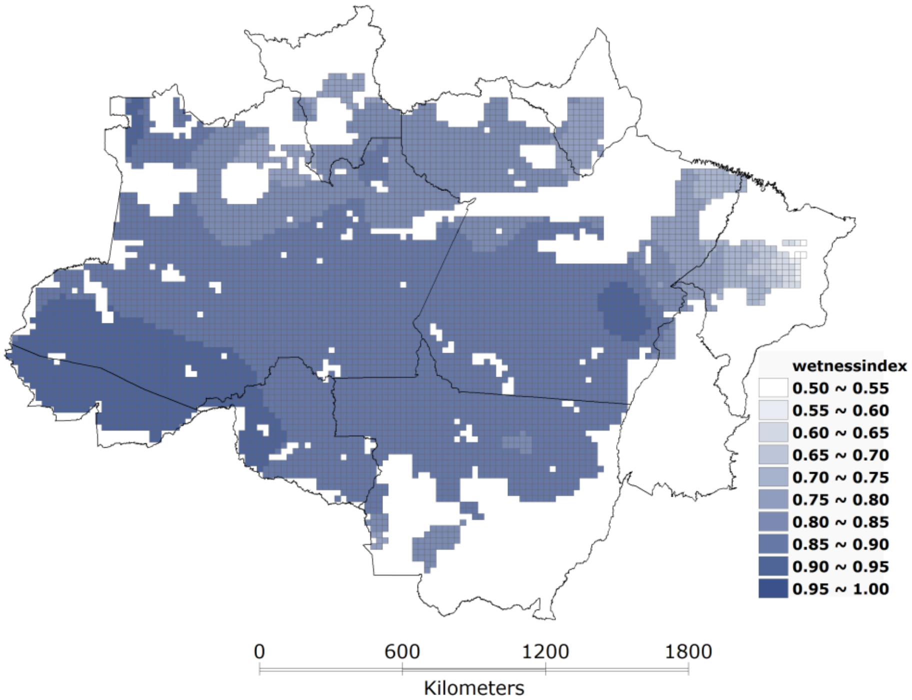

Figure 3-24: wetnessindex (weti)

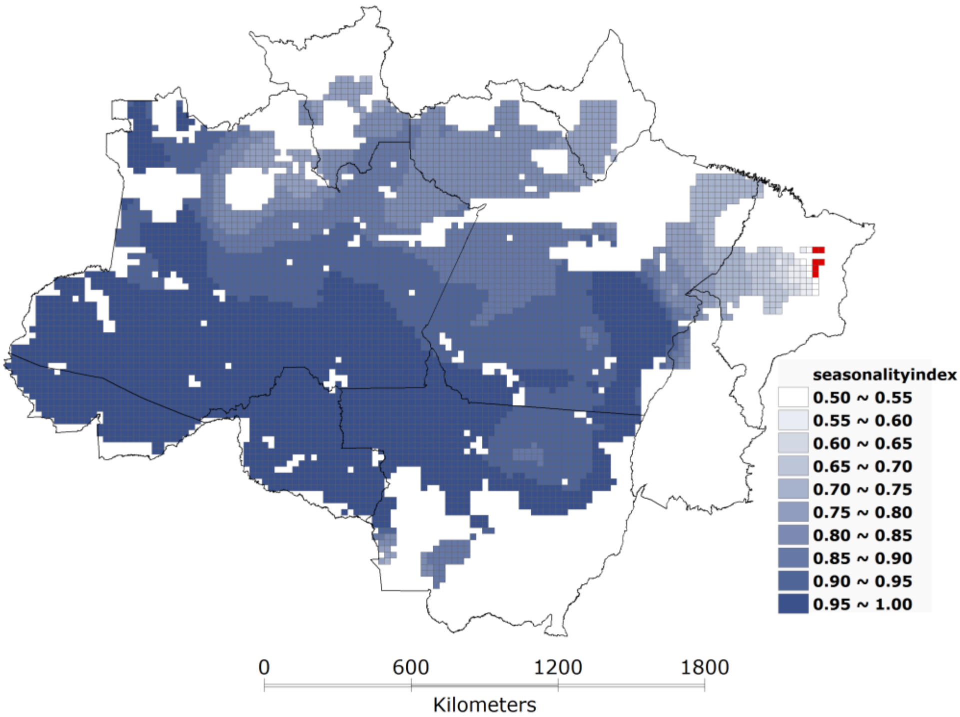

Figure 3-25: seasonalityindex (seai)

Integrating altimetry and slope data

For a possible improvement in discriminating pasture and agriculture patterns altimetry and slope data are included as additional environmental factors at scale 25x25km^2. The variables are derived from data of the Shuttle Radar Topography Mission (SRTM), which had the objective to generate the most complete high-resolution digital topographic database of the Earth by obtaining elevation data (SRTM Website, 2010). The data for the Brazilian Amazon is freely available with a spatial resolution of 3 arc-seconds (approximately 90 meters).

Four variables are introduced, one regarding altitude and three regarding slope. The altitude variable contains the average elevation in meters in each cell. For the generation of the slope variables three classes are defined: flat (0°-5°), moderately sloped (5°-15°) and steeply sloped (>15°). The three variables: slope_flat, slope_mod, slope_steep represent the fraction of each of the corresponding classes in each cell in terms of total cell area.

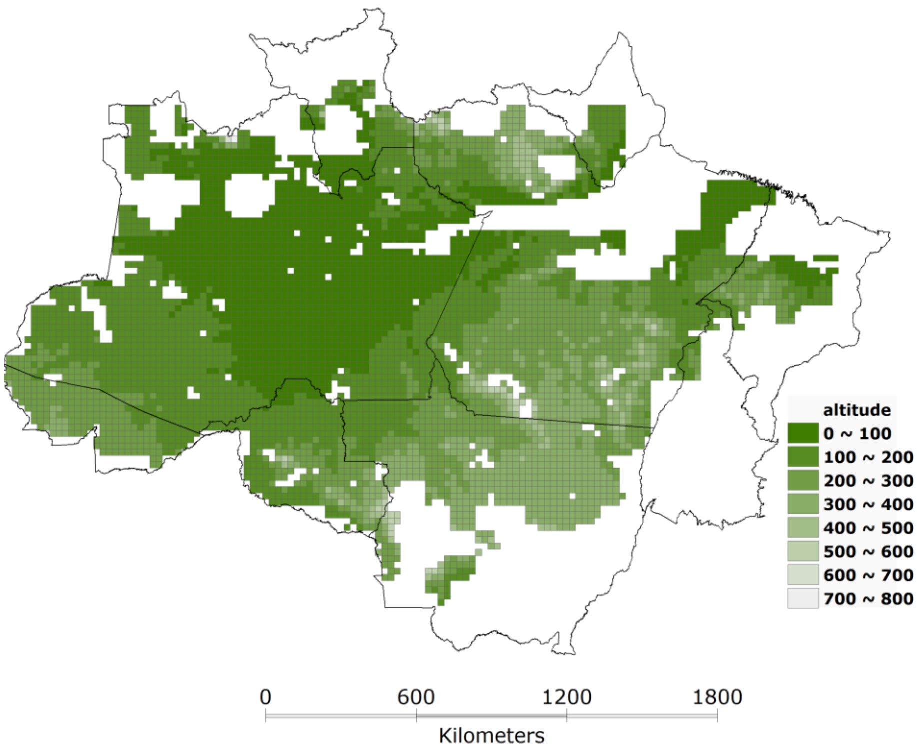

Figure 3-26: altitude (altitude_avg)

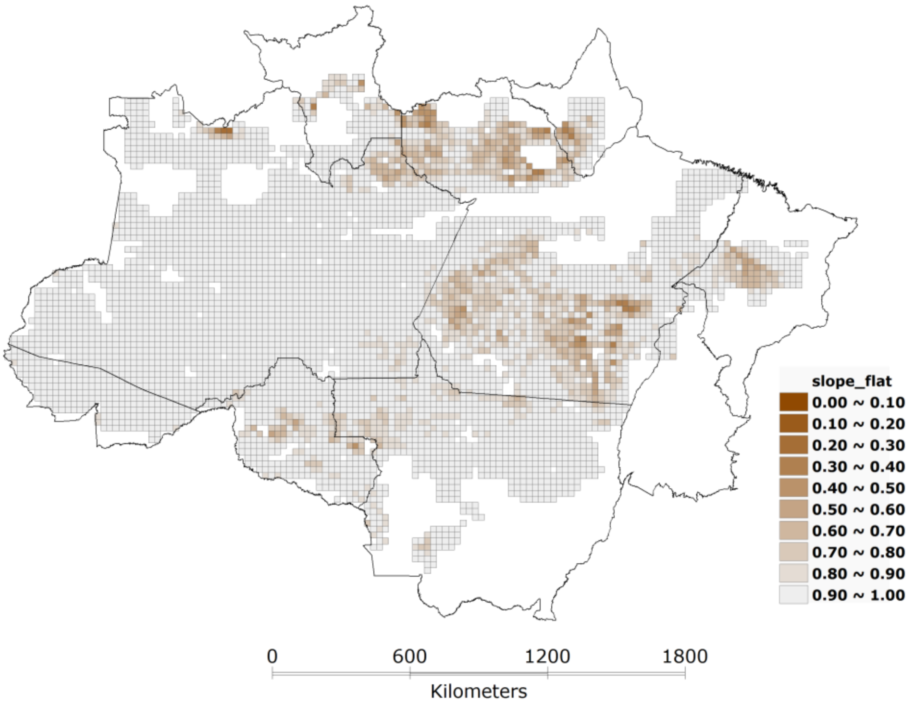

Figure 3-27: flat areas (slope_flat)

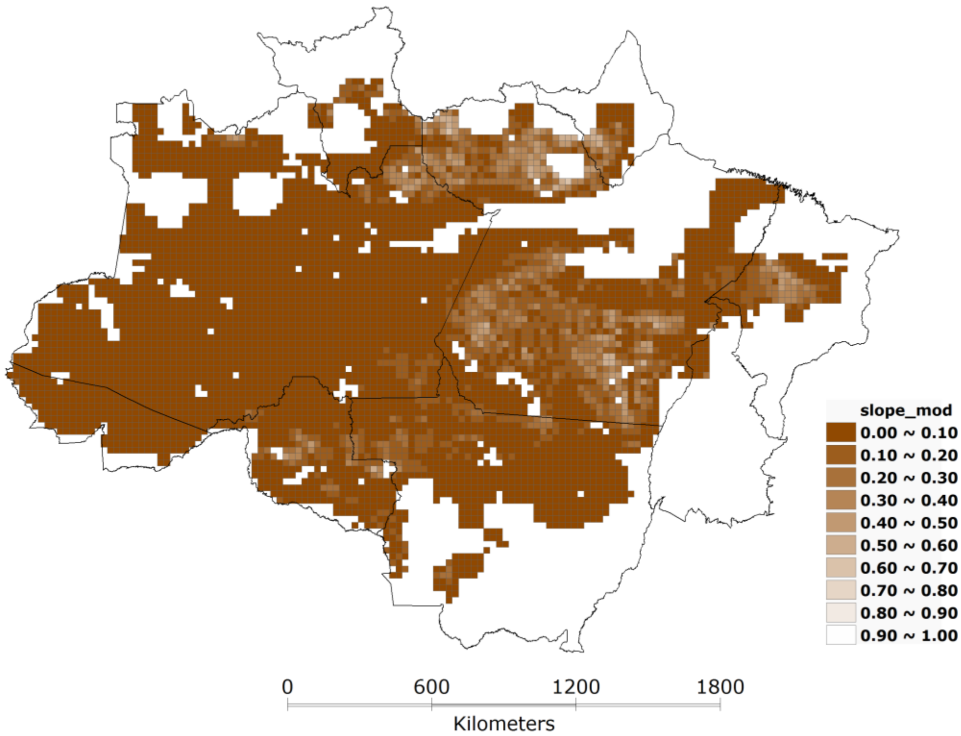

Figure 3-28: moderately sloped areas (slope_mod)

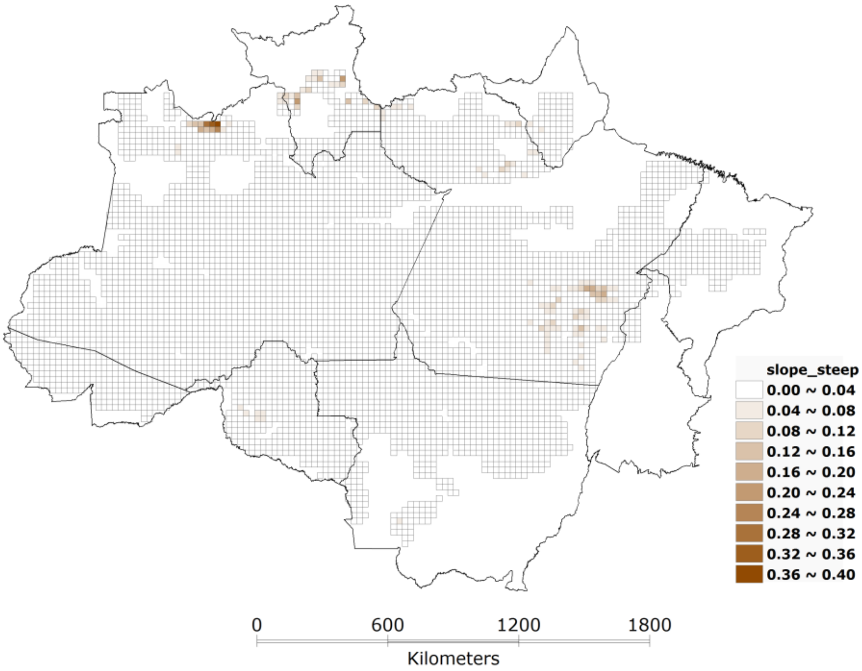

Figure 3-29: steeply sloped areas (slope_steep)

3.3.2 Statistical Analysis

Followed by a preliminary survey of the variables from the list of factors defined in (Aguiar, 2006), the potential land-use determining factors listed in Table 3-1 are selected for further analysis. The objective of the statistical analysis is to find a statistical model that is capable of explaining the deforestation and land-use patterns in the Brazilian Amazon with the help of explanatory variables. Correlation analysis and linear regression analysis are used to detect relationships between the dependent land-use variables (deforested areas, pasture, temporary and permanent agriculture, non-used agricultural land, planted forest and forest) and the independent explanatory variables. Several regression models are defined and compared with special emphasis on the newly integrated environmental variables. Thereafter the most suitable models are selected for the AmazonClueINPE model to test their abilities for allocation and projection of several land-use types.

3.3.2.1 Exploratory Data Analysis

Before carrying out the regression analysis, the potential land-use determining factors and the land-use variables are visually analyzed and undergo a correlation analysis. The Pearson correlation coefficient is defined by $\frac{Cov(X,Y)}{\sqrt{Var(X)*Var(Y)}}$, where Cov(X,Y) stands for the covariance of the two random variables X and Y and Var(X) and Var(Y) for their variance. It is a measure for the linear stochastic relation of X and Y, but it does not describe explicit causal relations (Navratil & Staudinger, 2006).

3.3.2.2 Regression Analysis

Based on the exploratory data analysis, multiple linear regression analysis is used to investigate the importance of a number of independent factors on the dependent land-use variables. The mathematical model for each land-use variable is defined by $Y = X\beta + \varepsilon$, where Y is a (n x 1) vector of values for the dependent land-use variable in n cells, X is a (n x k) matrix of a column of ones and (k-1) columns of explanatory variables in n cells, β is an (k x 1) vector of regression coefficients (including intercept) and ε an (n x 1) disturbance vector. The regression coefficients are estimated using the method of linear least squares.

To compare the regression models different criteria are used. One of them is the coefficient of determination R^2, the fraction of variance explained by the model (R Reference Index, 2009), which is defined by $R^2 = 1 - \frac{\sum R[i]^2}{\sum (y[i] - \bar{y})^2}$, where $\sum R[i]^2$ is the residual sum of squares and $\sum (y[i] - \bar{y})^2$ the total sum of squares. The adjusted R^2, which penalizes for a higher number of explanatory variables, is also used. The Akaike Information Criterion (AIC) is another measure used to rank statistical models and hence applicable for model selection. It is defined by $AIC = -2 * \log L + k * edf$, where L is the likelihood and edf the equivalent degrees of freedom. The smaller the value of the AIC, the better is the model. The Akaike Information Criterion is used in a stepwise procedure as a criterion for the selection of variables. To analyze the importance of the factors for a certain land-use the standardized regression coefficients (Beta values) are calculated. They can be interpreted as a measure of how many standard deviations of change in the dependent variable are related to a one standard deviation increase in the independent variable (Lesschen, Verburg, & Staal, 2005). Thus the importance of the different variables can be compared in terms of standard deviation units, disregarding their originally diverse units.

3.3.2.3 Alternative model construction

The potential explanatory variables are grouped into different models for the regression analysis. Difficulties arise when trying to distinguish the effects of explanatory variables due to their tendency to be highly correlated, which is fairly common in land-use analysis (Lesschen, Verburg, & Staal, 2005). Thus the limitation that the correlation coefficients between variables in the same model are not allowed to have an absolute value higher than 0.5, is considered. Hence probable important factors, for example distance to urban areas, distance to roads, connection to markets or the wetnessindex can not all be used together in the same model. This leads to the definition of several models with as little correlation between the factors but as much explanatory power as possible.

One of the objectives of this thesis is to further improve the statistical models defined in (Aguiar, 2006) for the Densely Populated Arch macro region, to allow better discrimination of pasture and temporary agriculture patterns. This is attempted by including additional environmental variables from the water balance model of the CPTEC-PVM and altimetry and slope data. The decision to build the statistical models at the fine scale for the Densely Populated Arch and not for the whole Brazilian Amazon is based on results in (Aguiar, 2006) where using a regression model of the Densely Populated Arch in all spatial regions led to more realistic spatial patterns in the dynamical modeling results. (compare section 2.1.3)

The models are shown in Table 3-2 and Table 3-3 and are named after the probable most influential variables. The label *environment* includes the wetnessindex (*weti*), the altitude (*altitude_avg*) and the slope (*slope_flat*) variable.

Table 3-2: Models at scale 100x100km^2

100.A[2]	100.B[2]	100.C	100.D
urban + humidity	roads + humidity	urban + wetness	roads + wetness
log_dist_urban_areas	log_dist_roads	log_dist_urban_areas	log_dist_roads
agr_area_small	agr_area_small	agr_area_small	agr_area_small
conn_ports_inv_p	conn_ports_inv_p	conn_ports_inv_p	conn_ports_inv_p
soils_fert_B1	soils_fert_B1	soils_fert_B1	soils_fert_B1
soils_fert_B3	soils_fert_B3	soils_fert_B3	soils_fert_B3
prot_all1	prot_all1	prot_all1	prot_all1
log_dist_large_rivers	log_dist_large_rivers	log_dist_large_rivers	log_dist_large_rivers
clima_humi_min_3_ave	clima_humi_min_3_ave	weti	weti

[2] adapted from (Aguiar, 2006)

Table 3-3: Models at scale 25x25km^2

25.A[2]	25.B	25.C	25.D
roads + connection to markets	roads + wetness + connection to São Paulo	roads + environment + connection to São Paulo	roads + environment + humidity
log_dist_non_paved_road	log_dist_non_paved_road	log_dist_non_paved_road	log_dist_non_paved_road
log_dist_paved_roads	log_dist_paved_roads	log_dist_paved_roads	log_dist_paved_roads
agr_area_small	agr_area_small	agr_area_small	prot_all1
log_setl_nfamilies_70_99	log_setl_nfamilies_70_99	log_setl_nfamilies_70_99	soils_fert_B1
prot_all1	prot_all1	prot_all1	soils_fert_B3
log_dist_large_rivers	log_dist_large_rivers	log_dist_large_rivers	weti
log_dist_min_deposits	log_dist_min_deposits	log_dist_min_deposits	conn_sp_inv_p
conn_ports_inv_p	conn_ports_inv_p	conn_ports_inv_p	altitude_avg
soils_fert_B1	soils_fert_B1	soils_fert_B1	slope_flat
soils_fert_B3	soils_fert_B3	soils_fert_B3	clima_humi_min_3_ave
conn_markets_inv_p	weti	weti	
log_dist_wood_extr_poles	log_dist_wood_extr_poles	log_dist_wood_extr_poles	
	conn_sp_inv_p	conn_sp_inv_p	
		altitude_avg	
		slope_flat	

3.3.3 Dynamical modeling: AmazonClueINPE

The LUCC model used in this work is based on the CLUE (Conversion of Land-Use and its Effects) model (De Koning, Veldkamp, Kok, & Bergsma, 1998; Kok, Farrow, Veldkamp, & Verburg, 2001; Veldkamp & Fresco, 1996; Verburg, De Koning, Kok, Veldkamp, & Bouma, 1999) and was adapted by (Aguiar, 2006) to be applicable for the Brazilian Amazon. It has the objective to provide a spatially-explicit, multi-scale, quantitative description of land-use changes. AmazonClueINPE is available in C++ and TerraME. In this thesis the TerraME version introduced in (Moreira, 2009) has been used. A review of the CLUE modeling framework and its adaptation to the Amazon is given in section 2.1.3.

The AmazonClueINPE model consists of a demand and an allocation module. In the non-spatial demand module scenarios of the quantity of change define how much change takes place every year in each land-use type. These demand values are calculated by multiplying the yearly deforestation rate, published by INPE (INPE, 2010), with a ratio for each land-use type (Table 3-4).

Table 3-4: area of land-use type per deforestation area in 1997

Pasture	Temporary agriculture	Permanent agriculture	Non-used agricultural areas	Planted forest
0.68	0.14	0.03	0.14	0.01

These ratios – area of land-use type per deforestation area – are calculated on basis of the Census of Agriculture in 1996 (IBGE, 1996) and assumed to not change in the time of study. Though values of the Census of Agriculture 2006 (IBGE, 2009) could have been used to interpolate the ratios for the intermediate years, it would have lead to problems, since the land-use type non-used agricultural land, which accounted for approximately 14% of total deforested area in 1996, was not included in the agricultural census in 2006.

The following parameters are used in the allocation module (Table 3-5). The model starts in 1997 and runs in one year time-steps until 2006. The deforestation threshold (lim_forest) that tries to slow down deforestation after a certain limit is reached in each cell is set to 0.2, which means that if less than 20% of the cell area is forest, a different allocation algorithm, which limits further deforestation is used. The maximum change value (max_change) per year is set to 0.5, which limits the change of a land-use type to a maximum of 50% of the cell area per time-step. The minimum elasticity (min_elasticity) is 0.1 for all land-use types and has an influence on the magnitude of change in the cells. The scale factor can be used to provide one of the two scales with higher importance. It is set to 1, hence no scale is favored. The number of iterations per time step is set to 2000. The maximum allowed difference (max_demand_diff) between demand and allocated areas is 0.01 times the demand.

Table 3-5: Parameters for the AmazonClueINPE allocation module

parameter	description	value
lim_forest	forest threshold to preserve 20% of cell area from deforestation	0.2
max_change	upper limit for change in one period of time	0.5
scale_fact	to favor one scale in respect of the other	1
max_iter	maximum number of iterations	2000
max_demand_diff	maximum allowed difference between demand and allocated change	0.01

Some of the potential land-use determining variables are updated during the AmazonClueINPE model runs. These are the connection to markets and ports variables (in 2000), as well as the distance to roads (in 2000) and the protected areas variables (in 2005).

3.3.3.1 Exploration of alternative regression models

An AmazonClueINPE model test plan is developed to test the different models of the regression analysis in a dynamical modeling approach. The attempt is to use as much complementary information at both scales as possible to test the ability of regression model combinations to simulate deforestation patterns and to discriminate different land-use types. Due to dissimilarities between the statistical models different spatial patterns are reproduced for each test. The objective of defining a test plan is to analyze the strengths and constraints of the different regression models and their combinations to find models that are capable of reproducing the actual deforestation and land-use patterns in Legal Amazon.

The hypothesis of this thesis states that the inclusion of hydrological, slope and altimetry variables improves the ability to discriminate pasture and agriculture patterns in the Brazilian Amazon. The following test plan (Table 3-6) is defined to explore this assumption.

- Analysis A uses regression models without newly integrated variables.
- Analysis B uses regression models which include the wetnessindex from CPTEC-PVM.
- Analysis C uses regression models which include the wetnessindex from CPTEC-PVM as well as slope and altimetry variables.

Table 3-6: AmazoniaClueINPE model test plan

	Test #	regressionmodel 100	regressionmodel 25
Analysis A	**A1**	100.A urban + humid	25.A roads + connection to markets
	A2	100.B roads + humid	25.A roads + connection to markets
Analysis B	**B1**	100.C urban + wetness	25.B roads + wetness + connection to São Paulo
	B2	100.D roads + wetness	25.B roads + wetness + connection to São Paulo
Analysis C	**C1**	100.C urban + wetness	25.C roads + environment + connection to São Paulo
	C2	100.C urban + wetness	25.D roads + environment + humidity

3.3.3.2 Evaluation of dynamic modeling results

The results of the AmazonClueINPE model runs are presented for the Legal Amazon at scale 25x25km^2. The maps for deforested areas, pasture and temporary agriculture are compared with special emphasis on some of the current hotspots of change (section 3.3.1.2). Due to the absence of the land-use type non-used agricultural land in the agricultural census data of 2006 and the fact that this class remains in the

AmazonClueINPE model, quantitative analysis is difficult to accomplish. Hence the evaluation of the AmazonClueINPE model results is mainly carried out visually and not on a cell to cell basis as in different quantitative map comparison methods. Nevertheless, ratios of change are compared per state where appropriate. Thus the analysis focuses more on the formation of new spatial patterns than on comparing specific cell values. Besides looking at the spatial patterns in general, it will be analyzed if the selected variables are capable of forcing the model to allocate the land-use changes in reasonable areas, i.e. areas which currently are or recently have been under human pressure.

3.4 Software

The statistical analysis is carried out in R[3], a free software environment for statistical computing and graphics. The AmazonClueINPE model is run with TerraME[4] RC4 (for TerraLib[5] 3.2), the DBMS used is Microsoft Access. The results of the AmazonClueINPE model are analyzed and visualized in TerraView[6] 3.2.0.

[3] www.r-project.org
[4] www.terrame.org
[5] www.terralib.org
[6] www.dpi.inpe.br/terraview

4 Results

In this chapter the results of the CPTEC-PVM implementation and the LUCC modeling are presented.

4.1 CPTEC-PVM

The CPTEC-PVM and its water balance model are implemented in the TerraME modeling environment as described in section 3.2. As a result the environmental variables from the water balance model are stored in the database and can be used by the AmazonClueINPE model. In addition to that the potential biome distribution is stored and visualized using TerraView 3.2.0 in Figure 4-1 on a global scale and in Figure 4-3 for the Brazilian Amazon. The spatial patterns of the wetnessindex and the seasonalityindex are shown in section 3.3.1.3. Figure 4-2 shows the potential vegetation map of the original version (Oyama & Nobre, 2004).

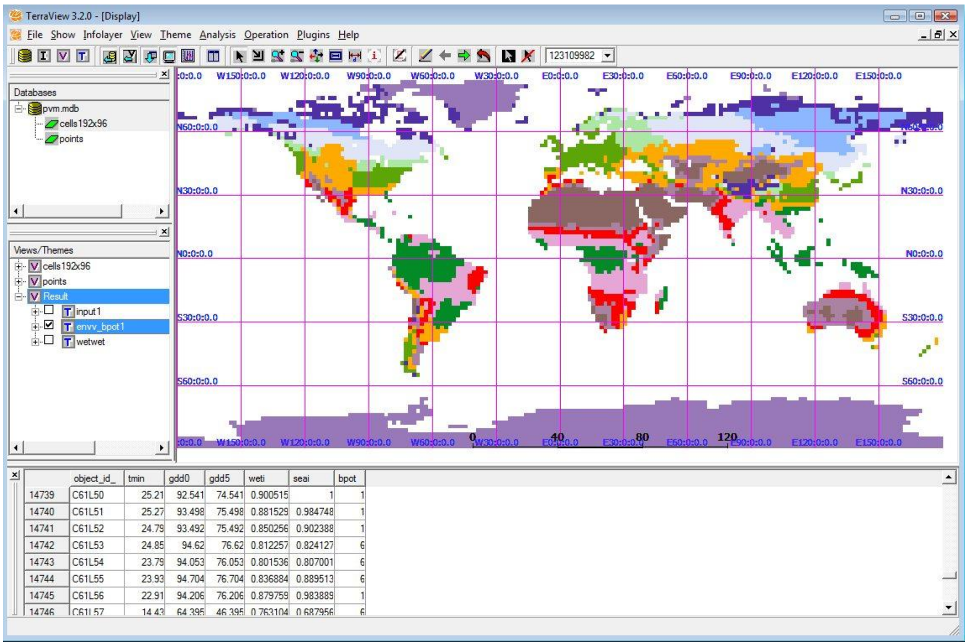

Figure 4-1: TerraView screenshot visualizing the result of the PVM implementation in TerraME

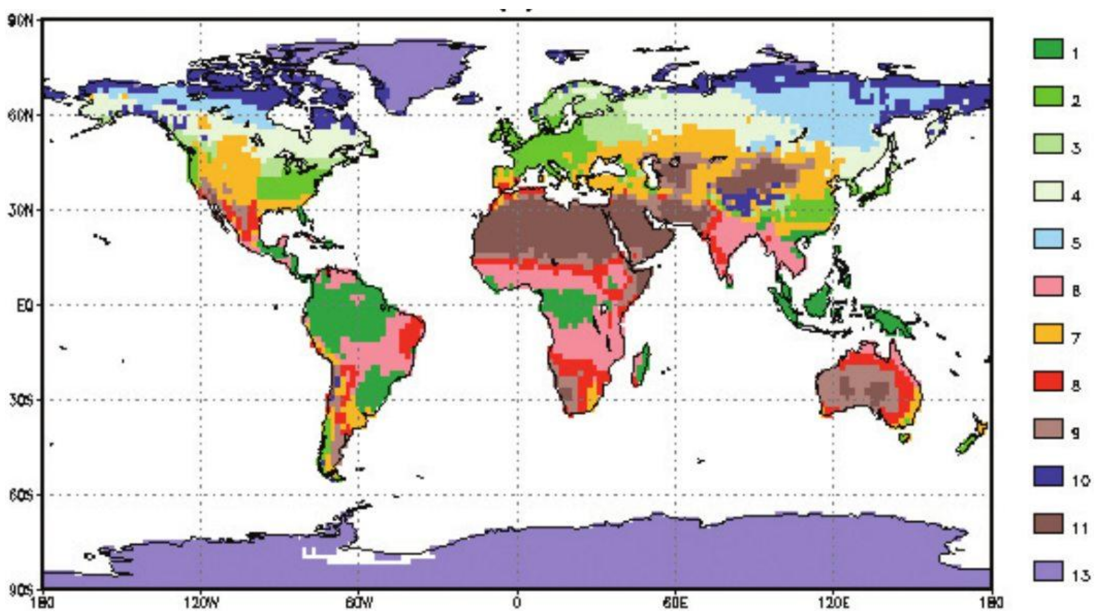

Figure 4-2: Potential vegetation (Oyama & Nobre, 2004) [1]

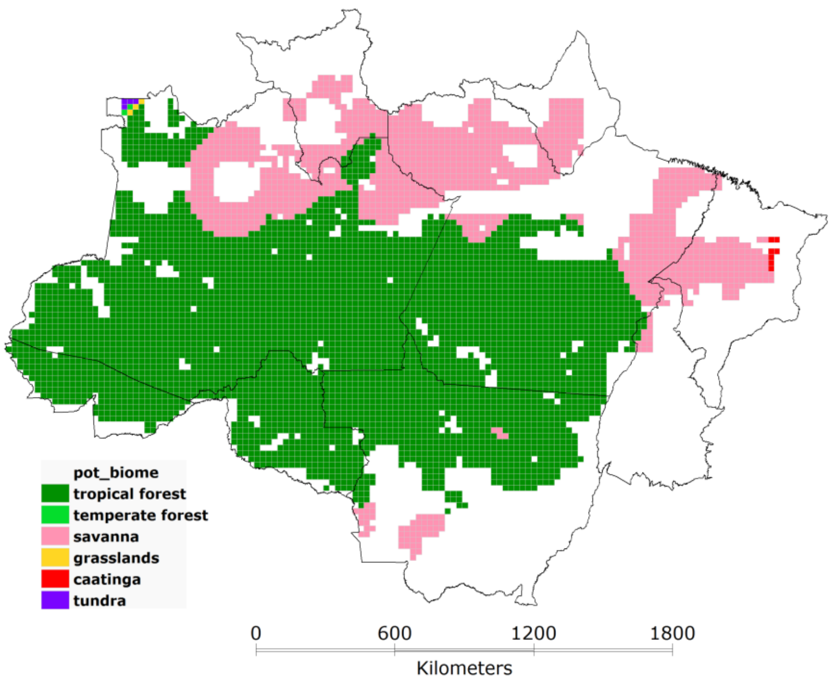

Figure 4-3: Potential vegetation in Legal Amazon

Only the two environmental variables from the water balance model of CPTEC-PVM are used for analysis in this thesis. Hence the potential biome classification has no effect on the land-use change modeling in the following section.

[1] (Oyama, 2005) changed some threshold values in the biome classification algorithm after the publication of (Oyama & Nobre, 2004), which lead to the generation of CPTEC-PVM v2. The implementation in TerraME and the screenshot in this section correspond to this version (v2), which results in minor differences between the figure from the original paper (v1, Figure 4-2) and the screenshot from TerraView (v2, Figure 4-1).

4.2 Land-use change modeling

The hypothesis, that the newly integrated environmental variables can improve the discrimination of pasture and temporary agriculture patterns, is under investigation in this chapter. The following sections show the results of the statistical analysis and a comparison of the land-use change model outcomes.

4.2.1 Statistical Analysis results

The statistical analysis is carried out for two scales, 100x100km^2 and 25x25km^2, as described in section 3.3.2. The descriptions of the used potential land-use determining factors can be found in Table 3-1. The statistical models are listed in Table 3-2 and Table 3-3.

4.2.1.1 Exploratory Data Analysis

The potential land-use determining factors are tested using the Pearson Correlation Coefficient (compare section 3.3.2.1) to find stochastic relations with special emphasis on the new environmental variables.

The two variables that were derived from the water balance model of the CPTEC-PVM wetnessindex and seasonalityindex are almost perfectly positive correlated (0.99), which is strengthened by looking at their spatial patterns (Figure 3-24 and Figure 3-25). Due to their similarity only one of the two indices is used for further analysis. The decision for the wetnessindex is based on a comparison of the correlation coefficients of the two environmental variables with the different land-use types. The correlation coefficients between the wetnessindex and the land-use variables show in general higher absolute values than the corresponding coefficients for the seasonalityindex (Table 4-1). Thus the assumption is made that the wetnessindex is likely to have more explanatory power and therefore used in the successive analysis.

Table 4-1: Correlation coefficients for the wetnessindex and the seasonalityindex with land-use variables in the Legal Amazon 25

	forest	deforestation(log)	pasture(log)	temp agric(log)	perm agric(log)	non-used land(log)	planted forest(log)
wetnessindex	0.38	-0.21	-0.21	-0.30	-0.16	-0.34	-0.16
seasonalityindex	0.31	-0.16	-0.15	-0.25	-0.13	-0.30	-0.15

In the Densely Populated Arch at scale 25x25km^2 the only variable, apart from the seasonalityindex, that has an absolute correlation coefficient higher than 0.5 with the wetnessindex is the connection to markets variable (-0.70, Table 4-2).

Table 4-2: Highest correlation coefficients for the wetnessindex in the Arch 25

	seai	conn_mkt	log_pop_dens_96	altitude_avg	clima_humi_min_3_ave	log_dist_urban_areas
weti	0.99	-0.70	-0.48	0.40	0.38	0.37

Hence the wetnessindex can be combined with any other variable in the regression models. The connection to São Paulo variable is not correlated to the wetnessindex (-0.04) and can therefore be used as a connection measure to the southeastern market. The correlation coefficient with the former climate variable clima_humi_3_ave is the forth largest value (0.38) after the population density in 1996 (-0.48) and the average altitude (0.40). As Figure 3-24 and Figure 3-23 illustrate, the wetnessindex and the humidity variable (average humidity in the three driest consecutive months) do not show the same spatial patterns. To test whether the wetnessindex is the more adequate environmental variable in terms of explanation of deforestation and land-use patterns or delivers additional information will be under investigation. It could lead to an improvement of the statistical models and furthermore to better simulations of deforestation and land-use processes.

The correlation coefficients for the soils variables (fertile and wet soils) with the wetnessindex are small (Table 4-3). There is almost no correlation with the wet soils variable (0.05) and only a small negative correlation with the fertile soils variable (-0.14).

Table 4-3: Correlation between the wetnessindex and other environmental variables in the Arch 25

	clima_humi_min_3_ave	fertile soils	wet soils
weti	0.38	-0.14	0.05

Considering the slope and altitude variables, the correlation coefficients (Table 4-4) indicate that the average altitude is negatively correlated with the fraction of flat slopes (-0.23) and positively correlated with the fraction of moderately (0.23) and steeply sloped areas (0.20). The slope variables show comprehensible correlations, e.g. slope_flat is almost perfectly negative correlated with slope_mod (-0.99) and highly negative correlated with slope_steep (-0.58). Hence only the slope_flat and the altitude_avg variable are used in the regression models.

Table 4-4: Correlation coefficients for the environmental variables in the Arch 25

	altitude_avg	slope_flat	slope_mod	slope_steep
altitude_avg	1	-0.23	0.23	0.20
slope_flat	-0.23	1	-0.99	-0.58
slope_mod	0.23	-0.99	1	0.50

slope_steep	0.20	-0.58	0.50	1

The potential important variables in the regression models do not show high correlation coefficients with the new environmental variables (Table 4-5). Only the connection to São Paulo variable almost reaches the 0.5 limit for excluding variables due to their correlation.

Table 4-5: Correlation coefficients for the environmental variables with important variables in the Arch 25

	log_dist_paved_roads	log_dist_non_paved_roads	prot_all1	conn_sp_inv_p
altitude_avg	0.25	-0.05	0.12	0.45
slope_flat	0.03	-0.06	-0.09	0.04

Correlation coefficients of the environmental variables with the land-use types (Table 4-6) reveal a strong negative correlation between the wetnessindex and the land-use temporary agriculture (-0.51) and weaker correlations to deforestation (-0.34) and pasture (-0.27). The altitude variable shows a negative correlation to temporary agriculture (-0.18) and almost no correlation to deforestation (-0.03) and pasture (0.03). For the slope variable the correlation coefficients show only small values and thus no clear indication for a connection to the land-use types in the Densely Populated Arch.

Table 4-6: Correlation coefficients for the environmental variables with land-use types in the Arch 25

	forest	deforestation (log)	pasture (log)	temp agric (log)	perm agric (log)	non-used land (log)	planted forest (log)
weti	0.49	-0.34	-0.27	-0.51	-0.19	-0.17	-0.52
altitude_avg	0.21	-0.03	0.03	-0.18	-0.29	-0.16	-0.21
slope_flat	0.02	0.00	-0.02	-0.01	0.06	-0.02	0.06
clima_humi_min_3	0.23	-0.26	-0.27	-0.28	0.13	0.05	-0.24
soils_fert_B1	-0.35	0.31	0.28	0.33	0.34	0.18	0.33
soils_fert_B3	0.08	-0.08	-0.10	-0.04	-0.05	0.00	-0.04

In general potential important variables like distance to urban areas, distance to roads, connection to markets and population density are to a high degree correlated (Table 4-7).

Table 4-7: Correlations for the Arch 25

	log_dist_urban_areas	log_dist_paved_road	log_dist_n_pav_road	conn_mkt	log_pop_dens_96	weti
log_dist_urban_areas	1	0.51	0.44	-0.46	-0.56	0.37
log_dist_paved_road	0.51	1	0.12	-0.44	-0.63	0.32
log_dist_n_paved_road	0.44	0.12	1	-0.27	-0.29	0.09
conn_mkt	-0.46	-0.44	-0.27	1	0.48	-0.70
log_pop_dens_96	-0.56	-0.63	-0.29	0.48	1	-0.48
weti	0.37	0.32	0.09	-0.70	-0.48	1

4.2.1.2 Regression Analysis

The results of the multiple linear regression analysis are presented for the Densely Populated Arch at scale 25x25km^2 for deforestation and the land-use types pasture and temporary agriculture with special emphasis on the environmental variables. Variables that are not significant at the 5% level, i.e. their p-values are greater than 0.05, are eliminated.

Twelve variables are included in the model 25.A – roads + connection to markets. The two soil variables (soil_fert_B1 and soil_fert_B3) are the only environmental variables in this model. The humidity variable (clima_humi_min_3_ave) can not be used because of its correlation with the connection to markets variable (-0.59). The 25.B – roads + wetness + connection to São Paulo model uses the wetnessindex from the CPTEC-PVM and replaces the connection to markets with the connection to São Paulo variable. The 25.C – roads + environment + connection to São Paulo model further adds the altimetry and slope factors. The 25.D – roads + environment + humidity model uses all environmental variables, including the humidity variable, but excludes some factors describing economic attractiveness and accessibility.

Deforestation

Considering the adjusted coefficient of determination (adj. R^2) and the Akaike Information Criterion (AIC) the performance of the regression models for deforestation does not vary significantly (Table 4-8). The adjusted R^2 ranges from 0.64 to 0.66. The most important factors are the accessibility variables like connection to markets or connection to São Paulo, which show a positive relation to deforestation, and distance to roads with a negative relation to deforestation. The highest absolute values in terms of standardized regression coefficients in all four models are represented by the protected areas (prot_all1) variable, which shows negative beta values between -0.36 and -0.38 and hence strong indications for a deforestation avoiding factor. The most important environmental variable in the models is the wetnessindex with a maximum value of -0.19 in models 25.B – roads + wetness + connection to São Paulo and 25.C – roads + environment + connection to São Paulo. The soils variables do not exceed beta values of 0.10 and also the slope variable shows small statistical relation to deforestation. Some variables are not significant and therefore excluded from the regression models like connection to ports, distance to large rivers and average altitude.

Table 4-8: regression models of deforestation for the Arch 25

25.A - roads + connection to markets			25.B - roads + wetness + connection to São Paulo		
	beta	p-value		beta	p-value
log_dist_non_paved_road	-0.21	0.00	log_dist_non_paved_road	-0.18	0.00
log_dist_paved_roads	-0.21	0.00	log_dist_paved_roads	-0.20	0.00
agr_area_small	-0.08	0.00	log_setl_nfamilies_70_99	0.03	0.02

log_setl_nfamilies_70_99	0.05	0.00	prot_all1	-0.36	0.00
prot_all1	-0.37	0.00	log_dist_min_deposits	-0.06	0.00
log_dist_min_deposits	-0.05	0.00	soils_fert_B1	0.09	0.00
soils_fert_B1	0.10	0.00	soils_fert_B3	0.07	0.00
soils_fert_B3	0.04	0.01	weti	-0.19	0.00
conn_markets_inv_p	0.34	0.00	log_dist_wood_extr_poles	-0.10	0.00
log_dist_wood_extr_poles	-0.11	0.00	conn_sp_inv_p	0.33	0.00
	adj. R^2	AIC		adj. R^2	AIC
	0.64	-1786.79		0.65	-1879.95
25.C - roads + environment + connection to São Paulo			25.D - roads + environment + humidity		
	beta	p-value		beta	p-value
log_dist_non_paved_road	-0.19	0.00	log_dist_non_paved_road	-0.22	0.00
log_dist_paved_roads	-0.20	0.00	log_dist_paved_roads	-0.24	0.00
prot_all1	-0.37	0.00	prot_all1	-0.38	0.00
log_dist_min_deposits	-0.04	0.00	weti	-0.15	0.00
soils_fert_B1	0.10	0.00	soils_fert_B1	0.10	0.00
soils_fert_B3	0.07	0.00	soils_fert_B3	0.06	0.00
weti	-0.19	0.00	conn_sp_inv_p	0.31	0.00
log_dist_wood_extr_poles	-0.10	0.00	slope_flat	-0.07	0.00
conn_sp_inv_p	0.33	0.00			
slope_flat	-0.06	0.00			
	adj. R^2	AIC		adj. R^2	AIC
	0.66	-1893.54		0.65	-1843.59

Pasture

The major land-use type in the Brazilian Amazon in 1997 was pasture with approximately 68% of all deforested areas (compare section 3.3.3). Hence the regression for this land-use type shows similar results to the regression results of deforestation. The adjusted R^2 reaches values between 0.63 and 0.66. The most important factors are again the protected areas variable and the accessibility factors. In comparison to the standardized regression coefficients for deforestation the distance to roads variables lost some explanatory power, whereas the connection to markets and the connection to São Paulo increased their importance. As a result the connection measures exceed the protected areas variable in absolute beta value in three of the models. The wetnessindex and the soils variable show also smaller absolute values for the standardized regression coefficients, while the values for the slope factor slightly increase.

Table 4-9: regression models of pasture for the Arch 25

25.A - roads + connection to markets			25.B - roads + wetness + connection to São Paulo		
	beta	p-value		beta	p-value
log_dist_non_paved_road	-0.22	0.00	log_dist_non_paved_road	-0.17	0.00
log_dist_paved_roads	-0.20	0.00	log_dist_paved_roads	-0.17	0.00

agr_area_small	-0.14	0.00	log_setl_nfamilies_70_99	0.05	0.00
log_setl_nfamilies_70_99	0.07	0.00	prot_all1	-0.35	0.00
prot_all1	-0.36	0.00	log_dist_min_deposits	-0.08	0.00
log_dist_min_deposits	-0.08	0.00	soils_fert_B1	0.06	0.00
conn_ports_inv_p	-0.04	0.02	soils_fert_B3	0.06	0.00
soils_fert_B1	0.09	0.00	weti	-0.13	0.00
soils_fert_B3	0.04	0.02	log_dist_wood_extr_poles	-0.09	0.00
conn_markets_inv_p	0.35	0.00	conn_sp_inv_p	0.41	0.00
log_dist_wood_extr_poles	-0.11	0.00			
	adj. R^2	AIC		adj. R^2	AIC
	0.63	-1837.88		0.65	-1980.98
25.C - roads + environment + connection to São Paulo			25.D - roads + environment + humidity		
	beta	p-value		beta	p-value
log_dist_non_paved_road	-0.18	0.00	log_dist_non_paved_road	-0.21	0.00
log_dist_paved_roads	-0.18	0.00	log_dist_paved_roads	-0.21	0.00
log_setl_nfamilies_70_99	0.05	0.00	prot_all1	-0.37	0.00
prot_all1	-0.35	0.00	weti	-0.07	0.00
log_dist_min_deposits	-0.07	0.00	soils_fert_B1	0.07	0.00
soils_fert_B1	0.08	0.00	soils_fert_B3	0.05	0.00
soils_fert_B3	0.07	0.00	conn_sp_inv_p	0.41	0.00
weti	-0.13	0.00	altitude_avg	-0.04	0.03
log_dist_wood_extr_poles	-0.09	0.00	slope_flat	-0.10	0.00
conn_sp_inv_p	0.40	0.00			
slope_flat	-0.08	0.00			
agr_area_small	-0.04	0.03			
	adj. R^2	AIC		adj. R^2	AIC
	0.66	-2007.27		0.65	-1946.34

Temporary agriculture

For the land-use type temporary agriculture the adjusted coefficients of determination have values around 0.70, only the 25.D – roads + environment + humidity model shows a smaller value of 0.65. The connection to markets and the connection to São Paulo variables are together with the protected area variable and the wetnessindex among the most important factors. A strong positive factor for temporary agriculture is the agrarian structure variable (agr_area_small). The paved roads variable shows significantly more impact than the non-paved roads variable. Considering the environmental variables the wetnessindex shows large negative beta values. Out of the other environmental factors the soil fertility factor (soil_fert_B1) has the most influence on the land-use type temporary agriculture.

Table 4-10: regression models of temporary agriculture for the Arch 25

25.A - roads + connection to markets			25.B - roads + wetness + connection to São Paulo		
	beta	p-value		beta	p-value

log_dist_non_paved_road	-0.08	0.00	log_dist_non_paved_road	-0.07	0.00
log_dist_paved_roads	-0.19	0.00	log_dist_paved_roads	-0.20	0.00
agr_area_small	0.18	0.00	agr_area_small	0.23	0.00
log_setl_nfamilies_70_99	0.07	0.00	log_setl_nfamilies_70_99	0.06	0.00
prot_all1	-0.27	0.00	prot_all1	-0.26	0.00
soils_fert_B1	0.11	0.00	log_dist_large_rivers	0.03	0.04
soils_fert_B3	0.04	0.01	soils_fert_B1	0.10	0.00
conn_markets_inv_p	0.46	0.00	soils_fert_B3	0.06	0.00
log_dist_wood_extr_poles	-0.13	0.00	weti	-0.32	0.00
			log_dist_wood_extr_poles	-0.11	0.00
			conn_sp_inv_p	0.32	0.00
	adj. R^2	AIC		adj. R^2	AIC
	0.69	-3184.99		0.70	-3196.60
25.C - roads + environment + connection to São Paulo			25.D - roads + environment + humidity		
	beta	p-value		beta	p-value
log_dist_non_paved_road	-0.07	0.00	log_dist_non_paved_road	-0.16	0.00
log_dist_paved_roads	-0.20	0.00	log_dist_paved_roads	-0.30	0.00
log_setl_nfamilies_70_99	0.06	0.00	prot_all1	-0.27	0.00
prot_all1	-0.26	0.00	weti	-0.26	0.00
log_dist_large_rivers	0.03	0.04	soils_fert_B1	0.14	0.00
soils_fert_B1	0.10	0.00	soils_fert_B3	0.06	0.00
soils_fert_B3	0.06	0.00	conn_sp_inv_p	0.23	0.00
weti	-0.32	0.00	altitude_avg	-0.11	0.00
log_dist_wood_extr_poles	-0.11	0.00	slope_flat	-0.08	0.00
conn_sp_inv_p	0.32	0.00	clima_humi_min_3_ave	-0.06	0.01
agr_area_small	0.23	0.00			
	adj. R^2	AIC		adj. R^2	AIC
	0.70	-3196.60		0.65	-2962.19

The regression analysis shows that the beforehand defined potential leading factors like distance to roads and connection to markets are amongst the most important variables in the various models. The variable for protected areas (prot_all1) plays a decisive role for all land-use types. The wetnessindex and the agrarian structure variable show especially for temporary agriculture high importance. An interesting aspect is that the agrarian structure variable has negative regression coefficients for deforested areas and pasture, but positive coefficients for temporary agriculture. The relatively low importance of the humidity variable in model 25.D – roads + environment + humidity might be a result of the simultaneous usage of this environmental variable and the connection to São Paulo variable, since the latter already carries a lot of information, which might be redundant, although with different sign. These two variables share a correlation coefficient of -0.51, but nevertheless it was decided to include both variables to make the environmental variables comparable without excluding a potential important variable.

The regression results indicate that the environmental variables are – apart from the wetnessindex in some models – not amongst the main land-use determining factors. Although they seem to carry only a small fraction of information at this scale, they can contribute as additional variables to help explaining land-use decisions.

4.2.2 Dynamic modeling: AmazonClueINPE

The AmazonClueINPE model is used as dynamic modeling approach to simulate land-use change processes in the Brazilian Amazon. The LUCC model and its parameters are described in section 3.3.3.

The results of the AmazonClueINPE model are presented for deforestation and the land-use types pasture and temporary agriculture for scale 25x25km^2. To allow for visual comparison with the actual maps of change between the years 1997 and 2006 (Figure 3-3, Figure 3-5 and Figure 3-7), the relative change between these years has been computed.

4.2.2.1 Analysis of deforestation patterns

Analysis A

The regression models in Analysis A do not include the newly integrated variables. The spatial outcomes of model A1 (Figure 4-4) and A2 (Figure 4-5) vary due to the different regression models at the coarse scale. The results show spatial patterns that are mostly aligned along the roads because of the importance of the distance to paved and unpaved roads variables in the regression models. The two models do not reproduce the actual intensification of forest destruction in northern Mato Grosso and eastern Pará, but have especially in the result of A1 a strong diffusive pattern close to Porto Velho (RO) and Humaitá (AM) spreading along the BR-319 road to Manaus (AM). The BR-319 is currently in bad condition and for the most part not feasible for transportation purposes (Fearnside & Graça, 2006), but according to the governmental plans to reconstruct this connection, new deforestation areas could emerge as (Barni, Fearnside, & Graça, 2009) point out. The strong deforestation patterns around São Félix do Xingu in Pará that advance into the region of Terra do Meio are in a limited amount reproduced by the models, but their simulated pattern sprawls more along the road in comparison to the diffusive pattern present at the actual deforestation map (Figure 3-3). The forest destruction along the BR-163 connecting Cuiabá (MT) and Santarem (PA) is to some extent captured.

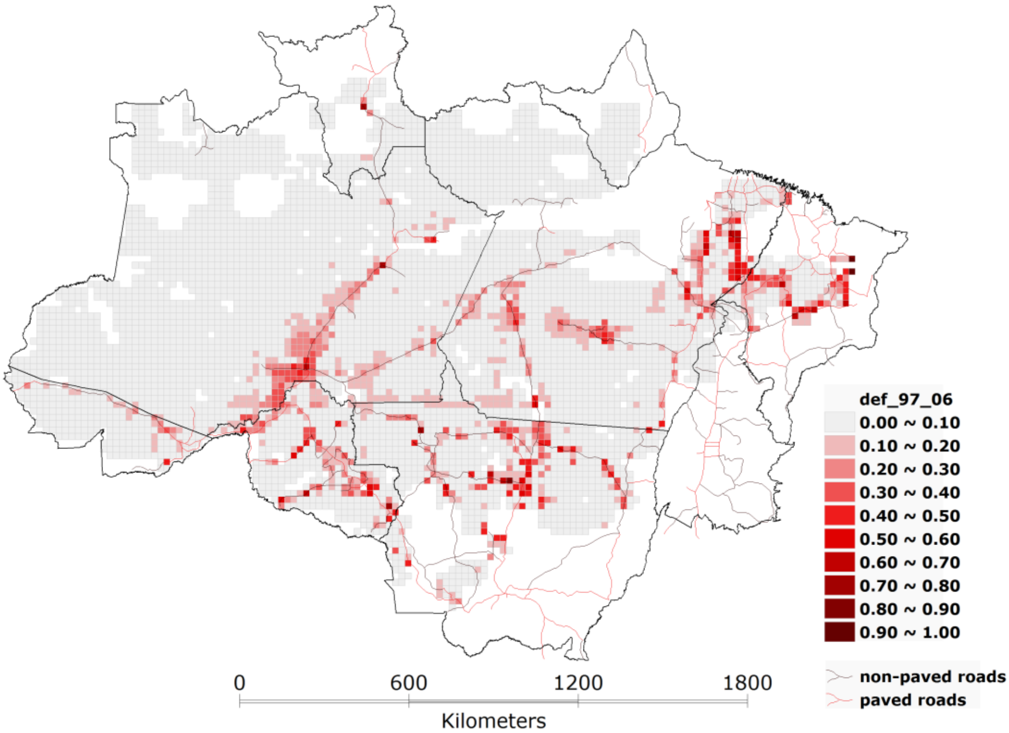

Figure 4-4: A1 – deforestation

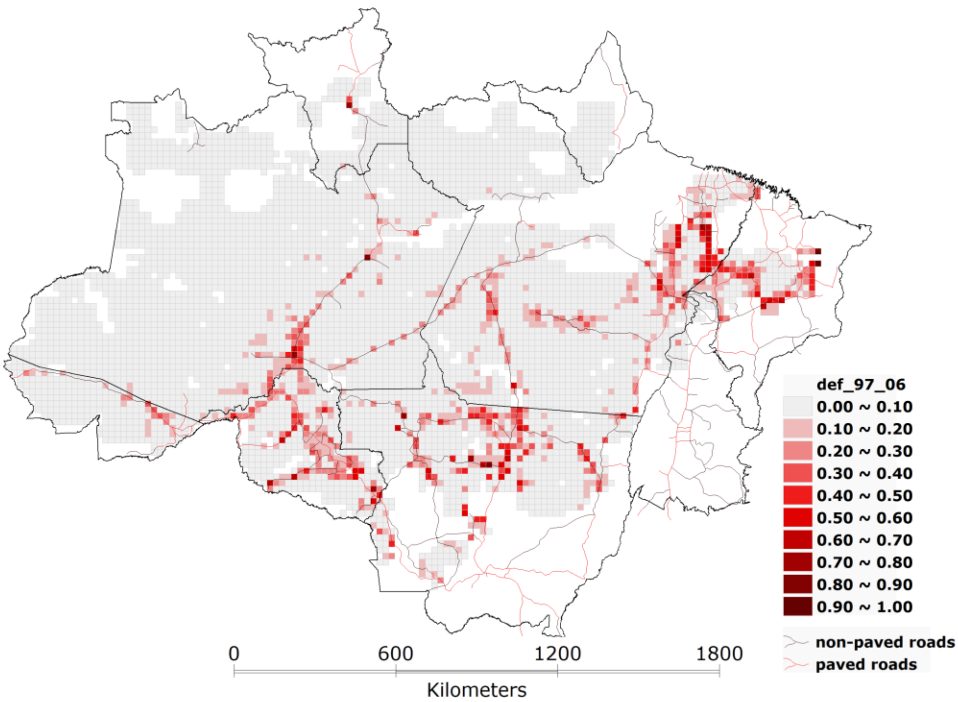

Figure 4-5: A2 – deforestation

Analysis B

The models B1 (Figure 4-6) and B2 (Figure 4-7) apply regression models including the wetnessindex. In addition to the environmental variable a connection measure to the south-eastern market is used at the fine scale. The consequence is that deforestation remains at a low level in the states of Roraima, Amapá and Acre, due to their distance to São Paulo. In states which are better connected to São Paulo, deforestation is allocated close to roads, e.g. in Mato Grosso and Rondônia. The outcomes are more reasonable results in the mentioned states in consideration of the actual reference deforestation map (Figure 3-3). But apart from these desirable changes simulated deforestation patterns also exist in the state of Amazonas around Humaitá and along the BR-319, where only small changes in forest cover took place in this period.

The humidity variable has higher absolute standardized regression coefficient values in the coarse scale models of Analysis A than the wetnessindex in Analysis B and thus more influence in the regression models. This fact is not obvious when analyzing the results of the AmazonClueINPE model. The difference between using the wetnessindex or the humidity variable in the coarse scale models is quite small and hence influences the results only to a small extent. The resulting deforestation patterns are almost the same, only the intensity slightly differs. Probably the two environmental variables are in the coarse scale regression models not important enough to produce significant differences in the results of the AmazonClueINPE model for deforestation.

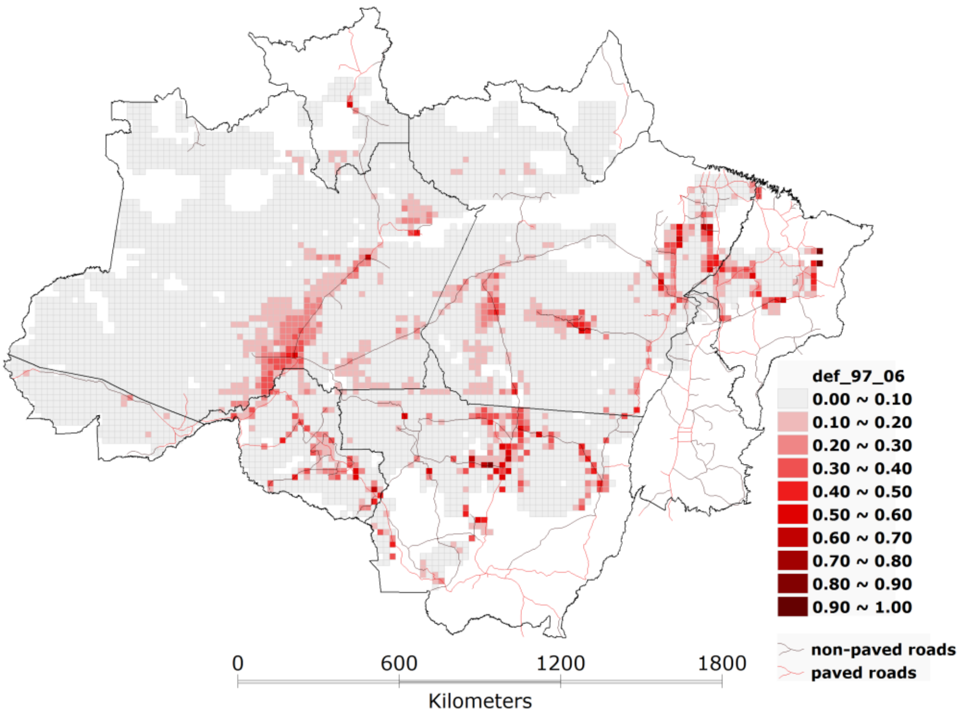

Figure 4-6: B1 – deforestation

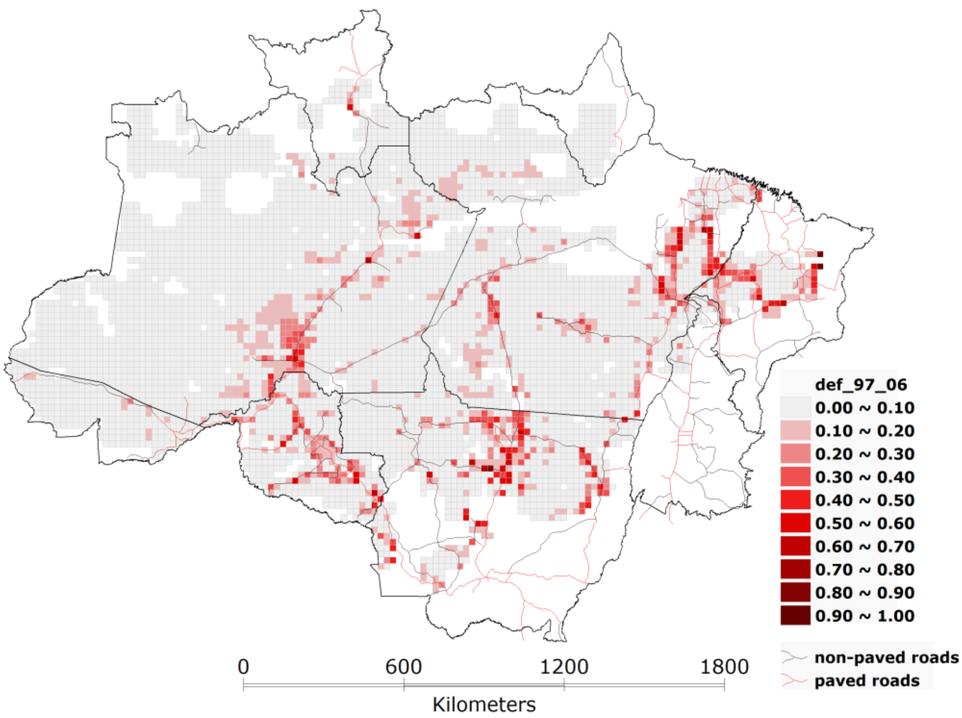

Figure 4-7: B2 – deforestation

Analysis C

In Analysis C the slope and altimetry variables are included in regression models at the fine scale of model C1 (Figure 4-8) and C2 (Figure 4-9).

Considering the analysis of the regression results (section 4.2.1) it was assumed that the effect of adding the slope and the altitude variable would not bring major changes in the results of the AmazonClueINPE model because these environmental variables were not among the main land-use determining factors. This is verified by the comparison of deforestation patterns of model B1 and C1. No new patterns emerge, only the intensity of land-use change slightly differs in some areas. Comparing model C1 and C2, it can be noticed that even more concentrated deforestation patterns around the BR-319 road in Amazonas state exist in C2. In Mato Grosso, north of Sinop and close to the BR-163, patterns in the tests are similar, but more intense in model C1. In other parts of Mato Grosso C2 allocates more deforestation, which leads to a greater deforestation area in this state. In Pará and Maranhão the spatial patterns again do not differ significantly in the results of the two model runs.

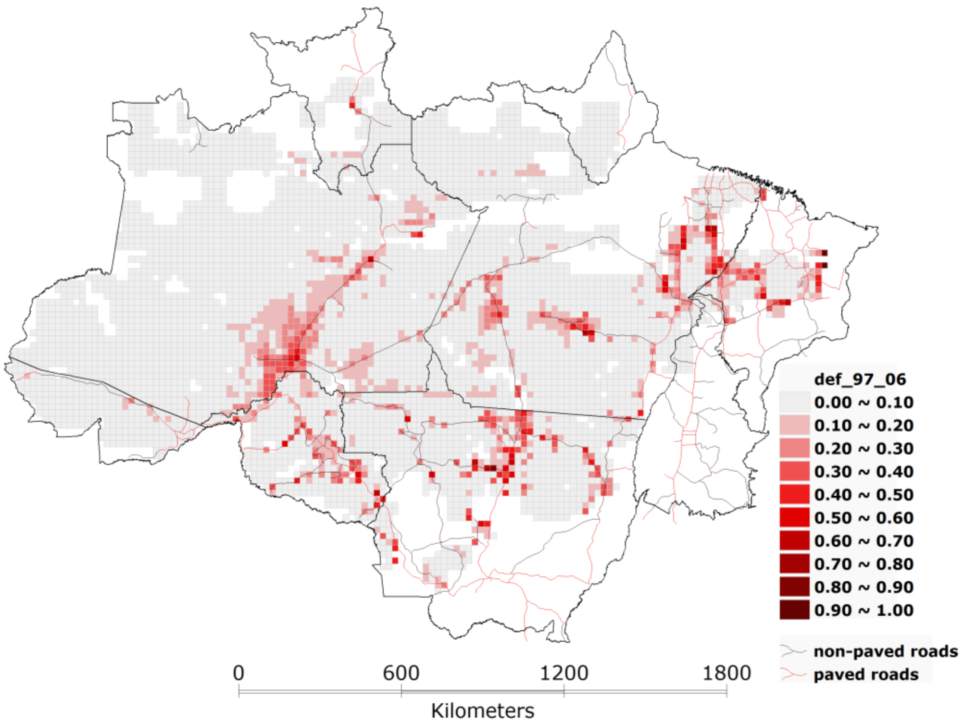

Figure 4-8: C1 – deforestation

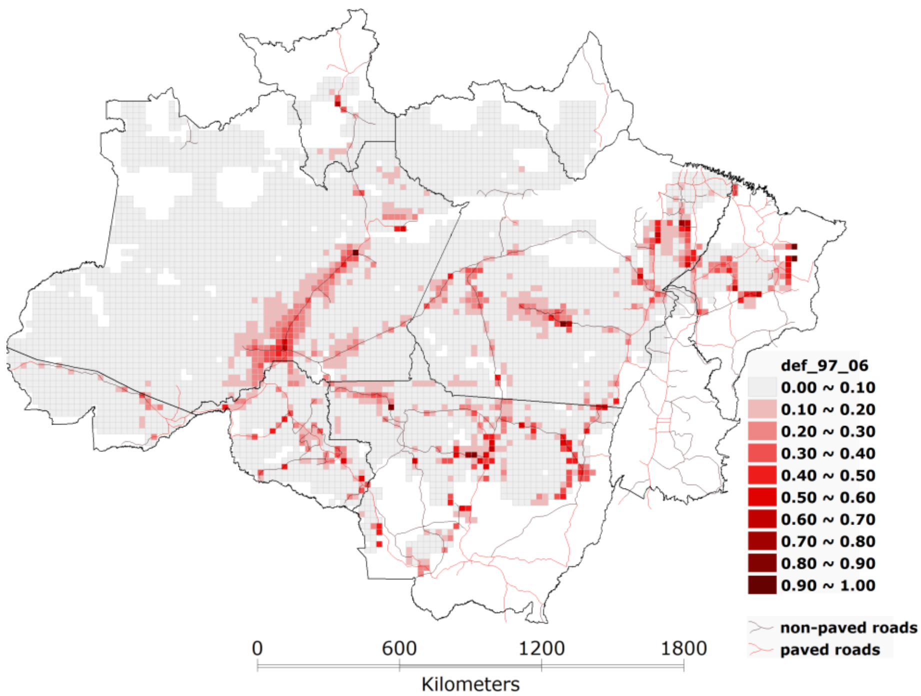

Figure 4-9: C2 – deforestation

The results of the AmazonClueINPE model for deforestation cannot fully explain the role of the newly integrated environmental factors on the allocation of specific land-use changes. Hence the resulting patterns for pasture and temporary agriculture are analyzed for some of the models separately to further determine the effect of environmental factors on the discrimination of pasture and temporary agriculture patterns.

4.2.2.2 Analysis of pasture and temporary agriculture patterns

Analysis A

The models in Analysis A serve as comparison to the models including the newly integrated variables in Analysis B and C to verify whether an improvement in explaining land-use patterns takes place.

The pasture patterns of model A1 (Figure 4-10) and A2 show only minor differences, as they use the same fine scale regression model, but different models at the coarse scale. The land-use type pasture increases in all states and does not show any decreasing patterns in this period. Compared to the pasture reference map (Figure 3-5), which is based on the agricultural census, too little change is allocated in eastern Para and northern Mato Grosso. This corresponds to the analysis of deforestation patterns for these models in the previous section.

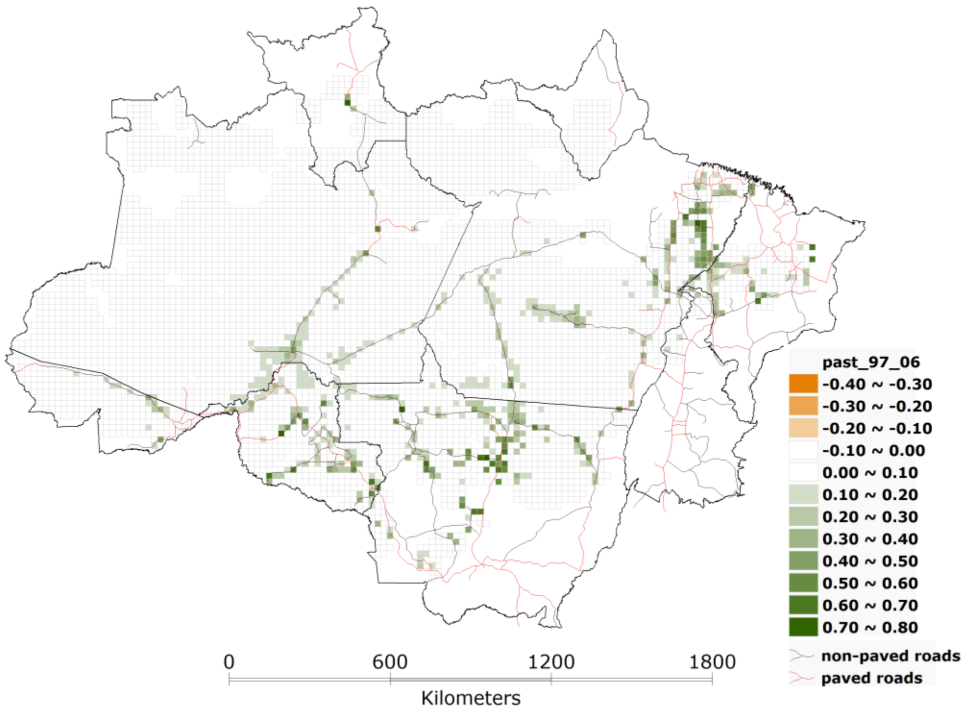

Figure 4-10: A1 – pasture

The temporary agriculture patterns of the models A1 (Figure 4-11) and A2 can not capture the increasing pattern in Mato Grosso. The change allocated in this state shows decreasing values. On the contrary temporary agriculture is favored in Rondônia and Amazonas state. Increasing patterns can also be found in north-eastern Pará and Maranhão.

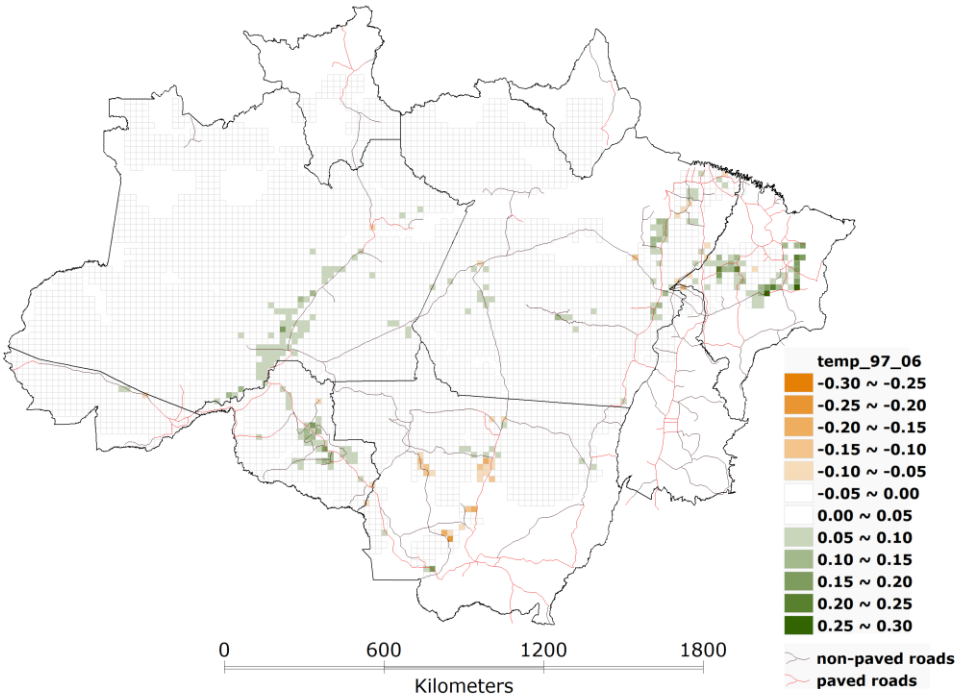

Figure 4-11: A1 – temporary agriculture

Analysis B

The wetnessindex is included in the regression models used in Analysis B. The pasture patterns do not significantly differ from the patterns in Analysis A in most states. Acre builds an exception, as almost no change is allocated in models B1 (Figure 4-12) and B2 in this state.

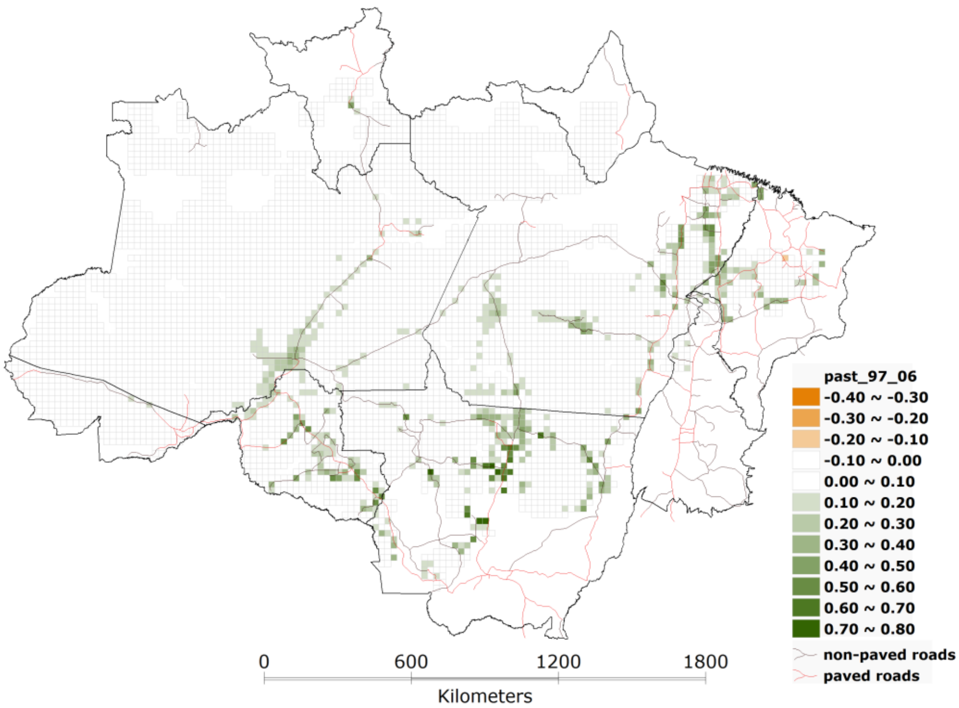

Figure 4-12: B1 – pasture

The models of temporary agriculture present similar results to the previous models. They again show the increasing spatial patterns in the southern part of the BR-319 road in Amazonas state as well as decreasing patterns in Mato Grosso.

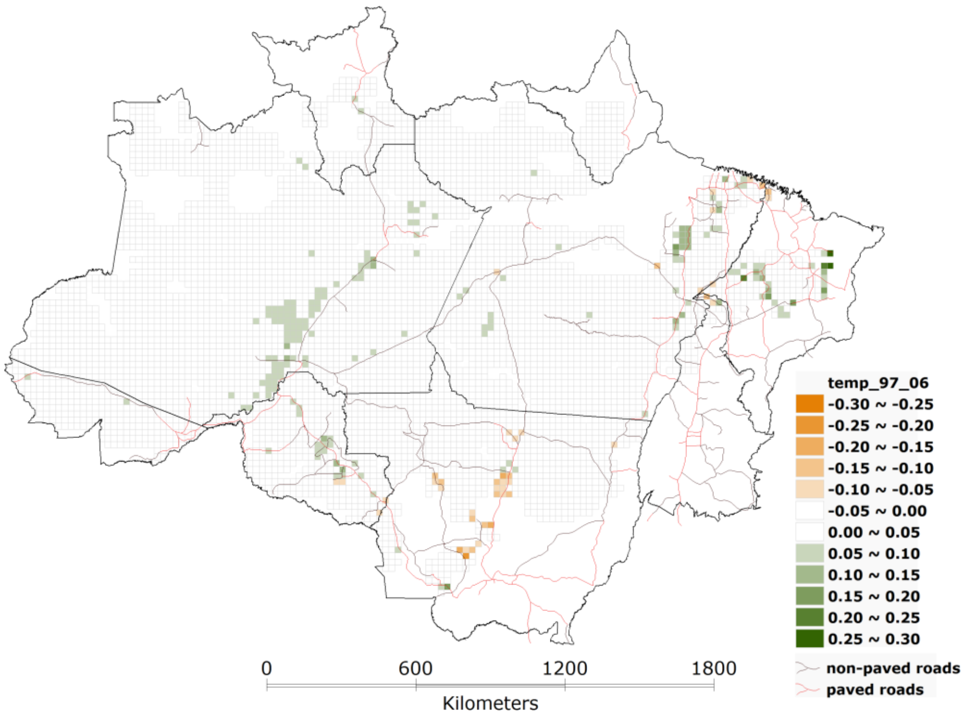

Figure 4-13: B1 – temporary agriculture

Analysis C

The increasing spatial pattern around Rio Branco in Acre in the pasture reference map (Figure 3-5) is not reproduced by the models C1 (Figure 4-14) and C2 (Figure 4-15). In Amazonas state the importance of the BR-319 road is overestimated by some variables as previously noticed in previous models that used the connection to São Paulo variable. Apparently some modification of the factors in this area needs to be incorporated to diminish this effect. In Rondônia the models show a significant pasture increase concentrated around the roads, while the pasture reference map shows a widespread pattern in almost the whole state, apart from an area in the west of the state. The positive pasture change in Mato Grosso state is to some extent captured by the models, but again it does not show the diffusive character visible in the reference map in the north of the state. In central Mato Grosso, where the census data shows a decrease in pasture, no decline is simulated in the models C1 and C2. In Pará the models show some increasing pasture patterns along the BR-163 road, but no distinctive pattern as in the census map. Also the spatial patterns concentrated along the BR-230 are, apart from some changes close to the BR-163, not captured by the models. In northeastern Pará the increasing patterns are reproduced quite well, but the decreasing pasture pattern in this area is not correctly simulated by the models. The spatial patterns in the area around São Félix do Xingu are limited to areas close to roads, a possible spreading northeastwards as shown in the reference map is possibly suppressed by the protected areas in this region and the importance of the corresponding variable in the regression models.

There is a discrepancy in the amount of land-use change between the AmazonClueINPE model results and the census reference map, as the AmazonClueINPE model uses fixed demand ratios based on the agricultural census of 1996. Thus the relative change per state in terms of total amount of land-use change in Legal Amazon is calculated and compared. Figure 4-16 confirms the overestimation of change of land-use type pasture in the time of study in the state of Amazonas, which accounts for 18% (model C1) and 25% (model C2) of the overall pasture change in the Legal Amazon, while the census data shows only an increase of 5%. In Pará and Rondonia too little change is allocated, compensating for the surplus in Amazonas. The Mato Grosso values fit quite well. The other states consist of too few cells to make an assumption.

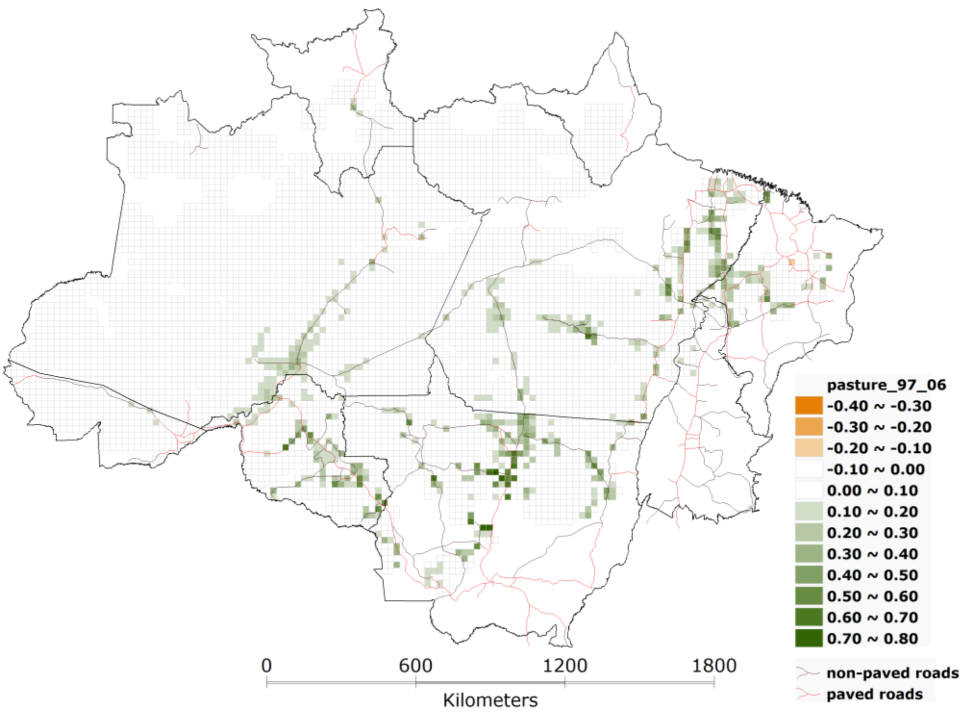

Figure 4-14: C1 - pasture

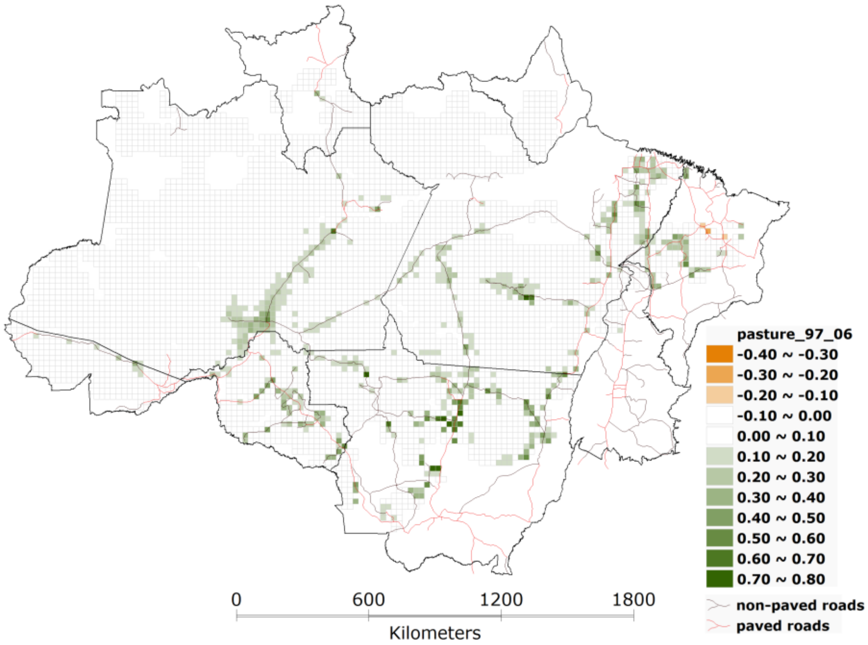

Figure 4-15: C2 - pasture

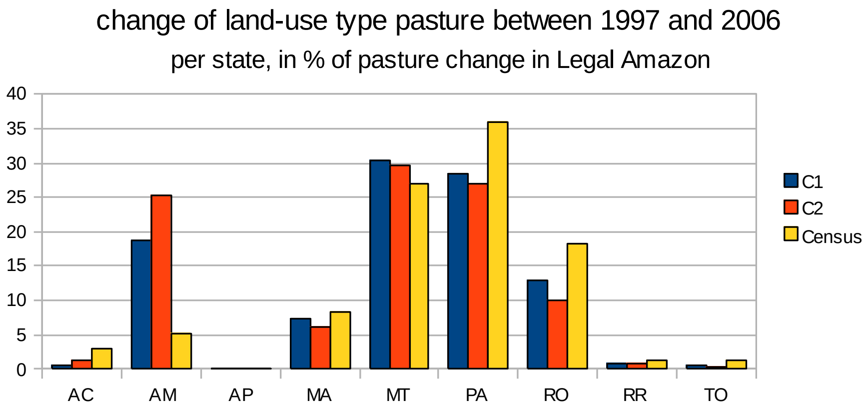

Figure 4-16: change of land-use type pasture per state between 1997 and 2006

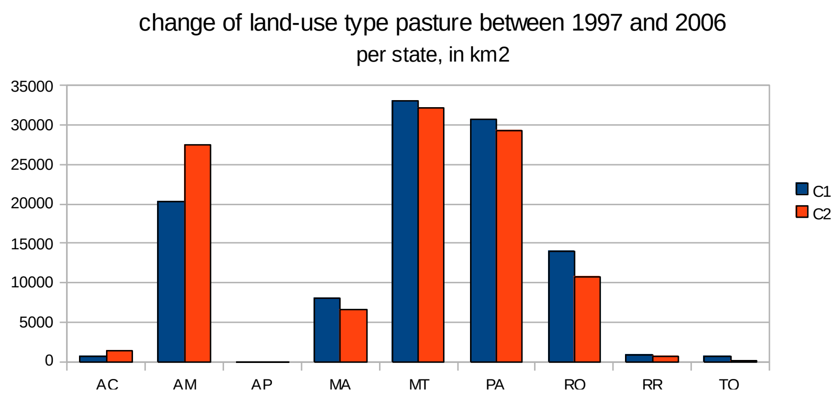

Figure 4-17: change of land-use type pasture per state between 1997 and 2006

In the census reference map for the land-use type temporary agriculture (Figure 3-7) two clear spatial patterns are visible. The first one is located in Mato Grosso and shows an increase in temporary agriculture for the period of 1997 to 2006, while the other one in

Maranhão shows a decrease in temporary agriculture. These two distinct processes are not captured by the models C1 (Figure 4-18) and C2 (Figure 4-19). Both models show a general increase around the BR-319 road in Amazonas state, in northeastern Pará and Maranhão, while some of the temporary agriculture patterns in Mato Grosso show a decrease, especially in model C1. Comparing the land-use changes per state (Figure 4-20), it can be seen that model C1 allocates more change of temporary agriculture in Amazonas than model C2. The opposite is true for pasture (Figure 4-17) where model C2 promotes better conditions than model C1. This might be due to the exclusion of the variable *agr_area_small* in the fine scale regression model of C2. This variable, an indicator for the agrarian structure, has a standardized regression coefficient of 0.23 (Table 4-10) for temporary agriculture in the fine scale regression model of C1 and relative high values in Amazonas state. The other dominant variables in the two regression models do not show significant differences in their regression coefficients.

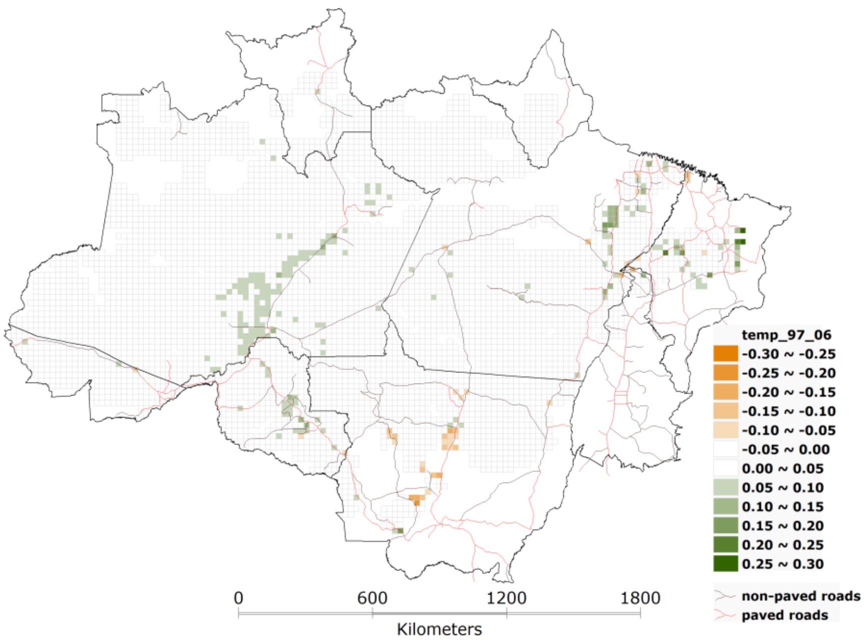

Figure 4-18: C1 - temporary agriculture

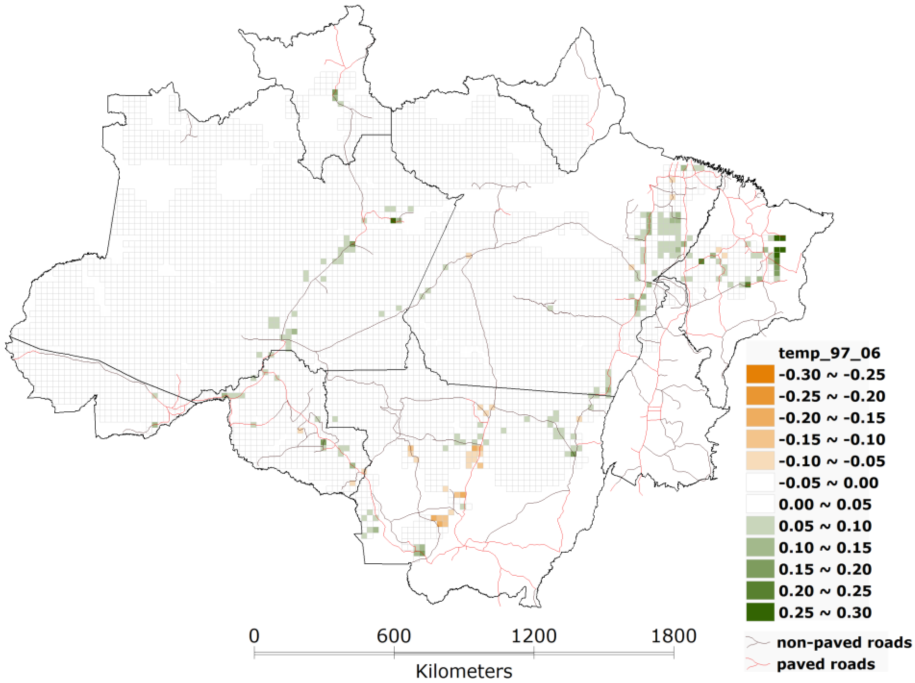

Figure 4-19: C2 - temporary agriculture

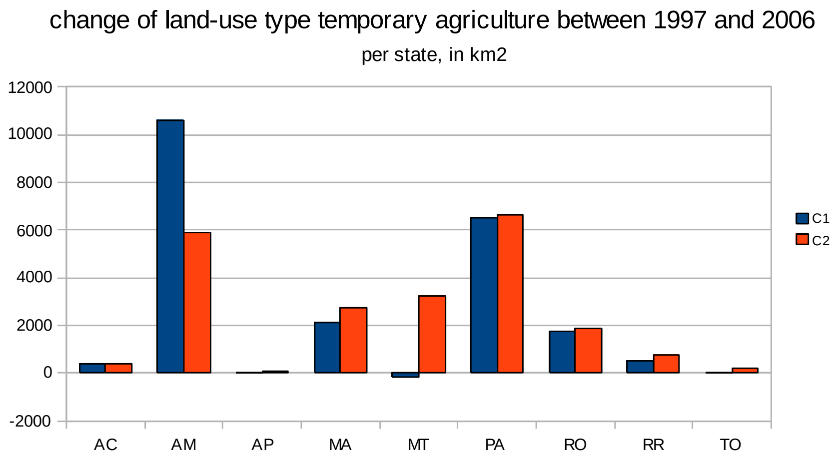

Figure 4-20: change of land-use type temporary agriculture per state between 1997 and 2006

4.3 Discussion

The implementation of the CPTEC-PVM in the TerraME modeling framework allows the integration of additional environmental factors in the AmazonClueINPE model. The consideration of supplementary biophysical parameters for slope and altimetry adds further possibilities to explore different factor combinations. The statistical analysis investigates the role of several variables and especially the importance of the environmental factors on land-use change processes. These statistics reveal that the environmental variables are in general not amongst the main land-use determining factors in the Brazilian Amazon. The wetnessindex, which originates from the water balance model of the CPTEC-PVM, builds an exception and shows relevant contribution especially for the land-use type temporary agriculture. The defined regression models serve as basis for the spatially-explicit AmazonClueINPE model. This dynamic modeling approach allows further exploration of the potential land-use determining factors. The objective is to distinguish the effect of the environmental variables in the allocation process of pasture and temporary agriculture changes.

The results of the AmazonClueINPE model indicate that the inclusion of the additional environmental variables does not result in major differences in deforestation patterns compared to models that do not use these variables. Analyzing the land-use types pasture and temporary agriculture it can be noticed that different regression models promote better or worse conditions for a land-use type in a specific region, depending on which variables are included. The pasture patterns can to some extent be reproduced by the models, but the temporary agriculture patterns can not be captured correctly. This statement holds for the tested models, no matter if the new environmental variables are included or not. Hence no improvement in the discrimination of pasture and temporary agriculture patterns due to the inclusion of new environmental factors can be stated.

Nevertheless the integration of slope and altimetry data and environmental variables like the wetnessindex or the humidity variable is capable of giving additional information on the formation of spatial patterns of land-use change, even though they do not show the explanatory power that other more dominant factors like distance to roads or connection to markets demonstrate.

Although the results do not promote an improvement through the inclusion of the new variables, the implementation of the CPTEC-PVM in the dynamic modeling environment can be seen as an important step, as it establishes new interaction possibilities to land-use change models like AmazonClueINPE or other models written in the TerraME modeling language.

5 Conclusion

After a short summary of the results of this thesis, the hypothesis will be evaluated, followed by a general discussion and an outlook on possible future work.

5.1 Summary

The hypothesis of this thesis is the following.

The inclusion of hydrological, slope and altimetry variables improves the ability to discriminate pasture and agriculture patterns in the Brazilian Amazon.

This hypothesis was tried to be corroborated by feeding a multi-scale LUCC model with data from a water balance model, SRTM and other data. Various combinations of regression models at two scales were tested to draw conclusions about the influence and explanatory power of each of the used factors. The results from Chapter 4 are shortly summarized below.

By following an AmazonClueINPE model test plan the statistical models were tested in a dynamical modeling approach. The preliminary regression analysis revealed that some factors like distance to roads, distance to urban areas or connection to markets determine most of the deforestation patterns. The wetnessindex has important character, especially for the land-use type temporary agriculture. At the fine scale, models lead by the distance to roads variables were used instead of the distance to urban areas variable, as these produced more reasonable spatial patterns. The connection to markets and the connection to São Paulo measure showed to be good factors to limit patterns to emerge in remote areas. Interchanging the humidity variable and the wetnessindex at the coarse scale did not reveal major changes in the model results.

For a better understanding of the role of the wetnessindex and the new biophysical variables the spatial patterns of pasture and temporary agriculture were analyzed. The pasture patterns were to some extent captured by the AmazonClueINPE model simulations, while the temporary agriculture patterns could not be reproduced. The results intensify the assumption that some sort of overestimation of land-use changes in the state of Amazonas takes place in most models. The spatial outcome of the AmazonClueINPE model also indicate that the resulting patterns are dominated by factors like accessibility to markets, distance to roads or presence of protected areas and only to a smaller degree by the slope, altimetry and other environmental variables. However the newly integrated variables influence the AmazonClueINPE model results, but their explanatory power is rather small in comparison to the dominant factors.

5.2 Hypothesis

On basis of the results and the analysis in Chapter 4 no improvement in the discrimination of pasture and temporary agriculture patterns solely on basis of the newly implemented variables could be verified. Despite this fact, the conclusion that these variables can in general not serve as auxiliary measures in LUCC models is unsuitable. But for the conducted analysis, the selected scales and the choice of potential land-use determining factors it could not be validated that the wetnessindex or the altimetry and slope variables advance simulations of land-use changes in the Brazilian Amazon.

Nevertheless integrating a combination of hydrological and biophysical data is thought to be a good and reasonable asset in the modeling approach to study land-use type conversions. Albeit these variables do not act as the main determinant factors they can support LUCC model simulations to discriminate pasture and agriculture patterns.

5.3 Discussion and Outlook

5.3.1 Statistical analysis

Several assumptions are usually made when using conventional statistical methods like multiple linear regression, e.g. data should be statistically independent and identically distributed (Aguiar, 2006; Cliff & Ord, 1981; Lesschen, Verburg, & Staal, 2005; Overmars & Veldkamp, 2003). These assumptions can not be fully satisfied by LUCC models, as land-use data usually has the tendency to be dependent, i.e. to show spatial autocorrelation (Aguiar, 2006; Overmars & Veldkamp, 2003). This spatial dependence is on one hand undesirable, but is also what gives information on spatial pattern (Aguiar, 2006; Gould, 1970; Overmars & Veldkamp, 2003).

As it is fairly common that collinearity exists in land-use analysis, it is difficult to separate the effects of each variable (Lesschen, Verburg, & Staal, 2005). As a way to reduce collinearity in this thesis a correlation analysis between the independent variables and a stepwise regression has been applied. Though this procedure might eliminate important factors, it is expected to be capable of selecting an adequate subset of variables through disregarding of non significant variables (Lesschen, Verburg, & Staal, 2005).

(Aguiar, 2006) tested a spatial lag model to simulate deforestation in the Brazilian Amazon and compared it to a multiple linear regression model. The findings were that by using the spatial lag model the resulting model allocates land-use changes mainly in previously occupied areas and thus prohibits the appearance of new patterns. Therefore the linear regression model was favored, while incorporating the diffusive nature of deforestation through scenario-dependent variables like distance to roads or connection to markets. Based on these conclusions by (Aguiar, 2006) it was decided to use a multiple linear regression model in this thesis.

5.3.2 Data Quality

The connection to markets and the connection to São Paulo measure proved to be good factors to limit patterns to emerge in remote areas, but have to be further improved as already mentioned in (Aguiar, 2006). The connection measures should also incorporate river networks. In addition to this the distinction between paved and non-paved roads is not sufficient. Quality of roads or waterways has to be incorporated as a proxy for the usability of a connection to get better measures for travel and transportation costs.

The CPTEC-PVM and its corresponding water balance model were initially developed for global purposes with a coarser resolution than used in this thesis, thus the question arises if the wetnessindex fed with detailed meteorological data is also feasible for finer resolutions or if better estimates for hydrological processes and regional climate exist for studies covering the Brazilian Amazon.

5.3.3 The wetnessindex and its relation to the connection to markets variable

In the Densely Populated Arch at the fine scale the wetnessindex and the connection to markets variable are to a high degree negative correlated (-0.70, Table 4-2). Visual analysis supports that similar patterns occur in the Arch, which raises the question, if there is some relation between these two variables. Are actors of land-use change taking environmental conditions, as represented by the wetnessindex, in their decisions into account to build stronger connections to favored places? And if so, is the wetnessindex an accurate variable to serve as a fundament for these decisions? If the statistical correlation is just a coincidence resulting from the generation of the variables or the scale of analysis or if there is a causal relation between them is an important question.

5.3.4 Regional models – local studies

Regarding (Lesschen, Verburg, & Staal, 2005) the scale of modeling has an effect on the type of pattern that will occur. (Zhang, Drake, & Wainwright, 2004) add that correlations found at one scale might not be applicable at another, as different processes are primarily dominant at different scales. Thus the question can be asked if the applied spatial scales and factors in this thesis are suitable to model deforestation processes in the study area. To simulate processes in large areas like the Brazilian Amazon with one set of regressions, assuming that the same factors are important everywhere, is an ambitious challenge. The region is probably too diverse to be represented by one model, thus intra-regional heterogeneity should be taken into account (Aguiar, 2006). Local studies and regional models both have their individual strengths, thus both types of land-use analysis approaches can strongly benefit from one another. They do not only differ in the size of the study area, but also often in the methods used to simulate land-use change processes. (Geist & Lambin, 2001) mention that a systematic comparison of local-scale case studies is a powerful tool to extract generalities on processes and causes of land-use change and provides more realistic insights than cross-national statistical analyses. The AmazonClueINPE model as applied in this thesis allows for a comprehensible way to model land-use change processes with the help of a set of factors, based on statistical relations between them, without the necessity of expert knowledge, but not leading to the explanatory power that some local studies might reveal. Being aware of the strengths and limitations of the utilized methodology is essential for correctly evaluating the results of LUCC modeling studies.

5.3.5 Outlook

To further test the AmazonClueINPE model based on the data used in this thesis, one possibility would be to adapt the land-use classes of the agricultural census 1996 to exclude the non-used agricultural areas and change the demand values based on yearly

PRODES deforestation data and the census data from 2006 to allow for quantitative comparison between the model results and the data. Besides adapting the variables, like the connection measures, further improvement could be reached by optimizing the regression models for each land-use type with the help of expert knowledge.

The implementation of the CPTEC-PVM and its corresponding water balance model in the TerraME modeling environment opens new possibilities to study the interaction between vegetation and LUCC models like the AmazonClueINPE model. Future studies could investigate the benefit of dynamic coupling of these two models.

References

Aguiar, A. P. (2006). Modeling Land Use Change in the Brazilian Amazon: Exploring Intra-Regional Heterogeneity. São José dos Campos, Brazil: INPE.

Andersen, L. E., & Reis, E. J. (1997). Deforestation, development, and government policy in the Brazilian Amazon: an econometric analysis. IPEA Report, Texto Para Discussão (Vol. 76, p. 24). Rio de Janeiro.

Andersen, L. E., Granger, C. W., Reis, E. J., Weinhold, D., & Wunder, S. (2002). The Dynamics of Deforstation and Economic Growth in the Brazilian Amazon (p. 259). Cambridge: Cambridge University Press.

Andrade, P. R., Junior, P. J., & Fook, K. D. (2005). Integration of statistics and Geographic Information Systems: the R/TerraLib case.

Aragão, L. E., Malhi, Y., Barbier, N., Lima, A., Shimabukuro, Y., Anderson, L., et al. (2008). Interactions between rainfall, deforestation and fires during recent years in the Brazilian Amazonia. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 363, 1779-85. doi: 10.1098/rstb.2007.0026.

Barbier, E. B., & Burgess, J. C. (2001). The economics of tropical deforestation. Journal of Economic Surveys, 15(3), 413-433.

Barni, P. E., Fearnside, P. M., & Graça, P. M. (2009). Deforestation and Carbon Emissions in Amazonia: Simulating the Impact of Connecting Brazil's State of Roraima to the "Arc of Deforestation" by Reconstructing the BR-319 (Manaus-Porto Velho) Highway. In XIII World Forestry Congress. Buenos Aires, Argentina.

Barreto, P., Souza Jr, C., Nogueron, R., Anderson, A., & Salomão, R. (2006). Human pressure on the Brazilian Amazon forests. World Resources Institute, Washington, DC, USA, and Imazon, Belém, Brazil.

Becker, B. (2005). Geopolítica da Amazônia. Estudos Avançados, 19(53), 71–86. SciELO Brasil.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., et al. (2007). Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology, 13, 679-706. doi: 10.1111/j.1365-2486.2006.01305.x.

Briassoulis, H. (2000). Analysis of Land Use Change: Theoretical and Modeling Approaches. Morgantown: Regional Research Institute, West Virginia University.

Carneiro, T. G. (2006). Nested-ca: a foundation for multiscale modelling of land use and land cover change.

Cliff, A. D., & Ord, J. K. (1981). Spatial Processes: Models & applications. London, UK: Pion.

Cook, K. H., & Vizy, E. K. (2008). Effects of Twenty-First-Century Climate Change on the Amazon Rain Forest. Journal of Climate, 21(3), 542-560. doi: 10.1175/2007JCLI1838.1.

Cox, P. M., Betts, R. A., Collins, M., Harris, P. P., Huntingford, C., Jones, C. D., et al. (2004). Amazonian forest dieback under climate-carbon cycle projections for the 21st century. Theoretical and Applied Climatology, 78, 137-156. doi: 10.1007/s00704-004-0049-4.

Câmara, G., Vinhas, L., Ferreira, K. R., de Queiroz, G. R., de Souza, R. C., Monteiro, A. M., et al. (2008). TerraLib: An Open Source GIS Library for Large-scale Environmental and Socio-economic Applications. In G. Brent Hall, Open Source Approaches to Spatial Data Handling (p. 247–270). Berlin: Springer.

De Koning, G. H., Veldkamp, A., Kok, K., & Bergsma, A. R. (1998). Clue: a tool for spatially explicit and scale sensitive exploration of land-use changes. In L. Lebel & D. Murdiyarso, IC-SEA report no. 5. Modelling global change impacts on tropical landscapes and biodiversity (pp. 29-40). Bogor, Indonesia.

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. doi: 10.1068/b3098.

Dorman, J. L., & Sellers, P. J. (1989). A Global Climatology of Albedo, Roughness Length and Stomatal Resistance for Atmospheric General Circulation Models as Represented by the Simple Biosphere Model (SiB). Journal of Applied Meteorology, 28(9), 833-855.

Fearnside, P. M. (2005). Deforestation in Brazilian Amazonia: history, rates and consequences. Conservation Biology, 19(3), 680-688.

Fearnside, P. M., & Graça, P. M. (2006). BR-319: Brazil’s Manaus-Porto Velho Highway and the potential impact of linking the arc of deforestation to central Amazonia. Environmental Management, 38(5), 705-716. doi: 10.1007/s00267-005-0295-y.ISSN.

Foley, J. A., Costa, M. H., Delire, C., Ramankutty, N., & Snyder, P. (2003). Green surprise? How terrestrial ecosystems could affect earth's climate. Frontiers in Ecology and the Environment, 1(1), 38–44. Eco Soc America.

GLP. (2005). GLP Science Plan and Implementation Strategy. Science. Stockholm.

Gash, J. H. (1996). Amazonian Deforestation and Climate. (J. H. Gash, C. A. Nobre, J. M. Roberts, & R. L. Victoria). Chichester: John Wiley & Sons Ltd.

Geist, H. J., & Lambin, E. F. (2001). What drives tropical deforestation? LUCC Report series (Vol. 4). Louvain-la-Neuve: LUCC International Project Office.

Goodchild, M. (2002). Issues in spatially explicit modeling. In D. Parker, T. Berger, S. M. Manson, & W. J. McConnell, Meeting the Challenge of Complexity: Proceedings of a Special Workshop on Land-Use/Land-Cover Change (pp. 24-28). Bloomington, USA: Center for the Study of Institutions, Population, and Environmental Change, Indiana University.

Gould, P. (1970). Is Statistix Inferens the Geographical Name for a Wild Goose? Economic Geography, 46, 439-448.

Heistermann, M., Müller, C., & Ronneberger, K. (2006). Land in sight? Achievements, deficits and potentials of continental to global scale land-use modeling. Agriculture, Ecosystems and Environment, 114, 141-158. doi: 10.1016/j.agee.2005.11.015.

IBGE. (1996). Censo Agropecuário 1996 (Census of Agriculture 1996). Retrieved from www.ibge.com.br/home/estatistica/economia/agropecuaria/censoagro/1995_1996/.

IBGE. (2007). Contagem da População 2007 (Population Count 2007). Retrieved from www.ibge.gov.br/home/estatistica/populacao/contagem2007/.

IBGE. (2009). Censo Agropecuário 2006 (Census of Agriculture 2006). Retrieved from www.ibge.com.br/home/estatistica/economia/agropecuaria/censoagro.

INPE. (2010). Projeto PRODES - Monitoramento da Floresta Amazônica Brasileira por Satélite. INPE. Retrieved from www.obt.inpe.br/prodes/.

Ierusalimschy, R., de Figueiredo, L. H., & Celes, W. (1996). Lua - an extensible extension language. Software: Practice & Experience, 26(6), 635–652.

Kaimowitz, D., & Angelsen, A. (1998). Economic Models of Tropical Deforestation A Review.

Kaimowitz, D., Mertens, B., Wunder, S., & Pacheco, P. (2004). Hamburger Connection Fuels Amazon Destruction. Statistics. Bogor, Indonesia.

Kok, K., Farrow, A., Veldkamp, A., & Verburg, P. H. (2001). A method and application of multi-scale validation in spatial land use models. Agriculture, Ecosystems and Environment, 85, 223-238.

Lambin, E. F. (2004). Modelling Land-Use Change. In J. Wainwright & M. Mulligan, Environmental Modelling: Finding Simplicity in Complexity (pp. 245-254). Chichester: John Wiley & Sons Ltd.

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

Lambin, E. F., Rounsevell, M. D., & Geist, H. J. (2000). Are agricultural landuse models able to predict changes in land-use intensity. Agric. Ecosyst. Environ., 82, 321-331.

Lapola, D. M., Oyama, M. D., & Nobre, C. A. (2009). Exploring the range of climate-biome projections for tropical South America: the role of CO 2 fertilization and seasonality. Global Biogeochem. Cycles, 23. doi: 10.1029/2008GB003357.

Laurance, W. F., Albernaz, A., Schroth, G., Fearnside, P. M., Bergen, S., Venticinque, E., et al. (2002). Predictors of deforestation in the Brazilian Amazon. Journal of biogeography, 29(5-6), 737–748.

Lesschen, J. P., Verburg, P. H., & Staal, S. J. (2005). Statistical methods for analysing the spatial dimension of changes in land use and farming systems. LUCC report series, 7(7). Citeseer.

Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., Nobre, C. A., et al. (2008). Climate Change, Deforestation, and the Fate of the Amazon. Science, 319(5860), 169-172. doi: 10.1126/science.1146961.

Moreira, E. (2009). Dynamic Coupling of Multiscale Land Change Models. São José dos Campos, Brazil: INPE.

Moreira, E., Costa, S., Aguiar, A. P., Câmara, G., & Carneiro, T. (2009). Dynamical coupling of multiscale land change models. Landscape Ecology, 24, 1183-1194. doi: 10.1007/s10980-009-9397-x.

Navratil, G., & Staudinger, M. (2006). Ausgleichungsrechnung I. 1040 Vienna, Austria: Institute for Geoinformation, Technical University Vienna, Gußhausstraße 27-29.

Nepstad, D. C., & Stickler, C. M. (2006). Globalization of the Amazon Soy and Beef Industries: Opportunities for Conservation. Conservation Biology, 20(6), 1595-1603. doi: 10.1111/j.1523-1739.2006.00510.x.

Nepstad, D. C., Stickler, C. M., Soares-Filho, B., & Merry, F. (2008). Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. Philosophical transactions of the Royal Society of London. Biological sciences, 363, 1737-1746. doi: 10.1098/rstb.2007.0036.

Nobre, C. A., Sellers, P. J., & Shukla, J. (1991). Amazon Deforestation and Regional Climate Change. Journal of Climate, 4.

Oliveira, R. S., Dawson, T. E., Burgess, S. S., & Nepstad, D. C. (2005). Hydraulic redistribution in three Amazonian trees. Oecologia, 145, 354-363. doi: 10.1007/s00442-005-0108-2.

Overmars, K. P., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. Ecological Modelling, 164, 257-270. doi: 10.1016/S0304-3800(03)00070-X.

Oyama, M. D. (2005). CPTEC-PVM, versão 2 (v2): inclusão de “minor changes” (Vol. 2).

Oyama, M. D., & Nobre, C. A. (2003). A new climate-vegetation equilibrium state for Tropical South America. Geophysical Research Letters, 30(23). American Geophysical Union. doi: 10.1029/2003GL018600.

Oyama, M. D., & Nobre, C. A. (2004). A simple potential vegetation model for coupling with the Simple Biosphere Model (SIB). Revista Brasileira de Meteorologia, 19(2), 203–216.

R Reference Index. (2009). R: A Language and Environment for Statistical Computing (p. 3282). The R Development Core Team. Retrieved from http://cran.r-project.org/doc/manuals/fullrefman.pdf.

SRTM Website. (2010). Shuttle Radar Topography Mission. Retrieved from http://www2.jpl.nasa.gov/srtm/.

Salazar, L. F., Nobre, C. A., & Oyama, M. D. (2007). Climate change consequences on the biome distribution in tropical South America. Geophys. Res. Lett., 34(L09708). doi: 10.1029/2007GL029695.

Schaldach, R., & Priess, J. (2008). Integrated Models of the Land System: A Review of Modelling Approaches on the Regional to Global Scale. Living Reviews in Landscape Research, 2.

Soares-Filho, B. (2002). DINAMICA - a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological Modelling, 154(3), 217-235. doi: 10.1016/S0304-3800(02)00059-5.

Soares-Filho, B., Nepstad, D. C., Curran, L. M., Cerqueira, G. C., Garcia, R. A., Ramos, C. A., et al. (2006). Modelling conservation in the Amazon basin. Nature, 440(7083), 520-3. doi: 10.1038/nature04389.

TerraME Website. (2010). TerraME: Simulation and Modelling of Terrestrial Systems. Retrieved from www.terrame.org.

TerraView Website. (2010). TerraView. Retrieved from www.dpi.inpe.br/terraview_eng.

Veldkamp, A., & Fresco, L. O. (1996). CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. Ecological Modelling, 91, 231-248.

Veldkamp, A., & Fresco, L. O. (1996). CLUE: a conceptual model to study the Conversion of Land Use and its Effects. Ecological modelling, 85, 253-270.

Veldkamp, A., Verburg, P. H., Soepboer, W., Limpiada, R., Espaldon, V., Mastura, S. S., et al. (2002). Modeling the spatial dynamics of regional land use: the CLUE-S model. Environmental Management, 30(3), 391–405. Springer. doi: 10.1007/s00267-002-2630-x.

Verburg, P. H., & Overmars, K. P. (2007). Dynamic Simulation of Land-Use Change Trajectories with the CLUE-s Model. In E. Koomen, J. Stillwell, A. Bakema, & H. J. Scholten, Modelling Land-Use Change. Progress and applications. (Volume 90., Vol. 1996, pp. 321-335). Springer. doi: 10.1007/978-1-4020-5648-2.

Verburg, P. H., De Koning, G. H., Kok, K., Veldkamp, A., & Bouma, J. (1999). A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. Ecological modelling, 116(1), 45–61. Elsevier.

Voldoire, A., & Royer, J. F. (2004). Tropical deforestation and climate variability. Climate Dynamics, (May), 857-874. doi: 10.1007/s00382-004-0423-z.

Werth, D., & Avissar, R. (2002). The local and global effects of Amazon deforestation. Journal of Geophysical Research, 107(D20). doi: 10.1029/2001JD000717.

Willmott, C. J., & Matsuura, K. (2001). Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999) (V 1.02). Retrieved from http://climate.geog.udel.edu/~climate/html_pages/download.html#ghcn_T_P2.

Willmott, C. J., Rowe, C. M., & Mintz, Y. (1985). Climatology of the terrestrial seasonal water cycle. Journal of Climatology, 5, 589-606.

Zhang, X., Drake, N. A., & Wainwright, J. (2004). Scaling Issues in Environmental Modelling. In J. Wainwright & M. Mulligan, Environmental Modelling: Finding Simplicity in Complexity (pp. 319-334). Chichester: John Wiley & Sons Ltd.

List of Figures

List of Tables