Centre for Geo-Information

Thesis Report GIRS-2005-30

Modelling Land Use Change in the Pantanal

Fritz Hellmann

August 2005

WAGENINGEN UNIVERSITY

WAGENINGEN UR

Modelling Land Use Change in the Pantanal

Fritz Hellmann

Registration number 780607-321-130

Supervisors:

Dr. Ir. Ron van Lammeren
Ir. Michiel van Eupen (Alterra)

A thesis submitted in partial fulfillment of the degree of Master of Science at Wageningen University and Research Centre, The Netherlands.

August 2005
Wageningen, The Netherlands

Thesis code number: GRS-80340
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing
Thesis Report: GIRS-2005-30

Acknowledgements

I would like to thank my supervisor Michiel van Eupen for the time and effort he has put into this study. It was always possible to ask something or discuss something with him. This study has benefited enormously from his input. I would like to thank him for his help.
I would also like to thank Ron van Lammeren for his supervision. He was always able to clearly point out into which direction this study needed to go.
In addition, I would like to thank John Stuiver. On many occasions, he was able to help me overcome serious problems in developing the model. Without his help, I would not have been able to develop a working model.
Rob Jongman, Bart Makaske, Carlos Padovani, Luc Boerboom and Ben Maathuis were always prepared to answer my questions. I would like to thank them for this.

Summary

Traditionally cattle's ranching in the Pantanal has been extensive, using natural pastures and low densities of cattle. This is relatively compatible with conservation objectives. However, recently ranchers have started to clear land and to convert it to cultivated pasture. This cultivation can have important environmental consequences. This study has focused on this land use change process.

This study aimed to prototype and to calibrate a model for the Pantanal, which can simulate future cultivation in scenario-studies. In addition, this study aimed to check the functionality of the model by running three scenarios. These three scenarios were:

1. The current trend of cultivation in the Pantanal will be extrapolated.
2. Lower 90% confidence interval of scenario 1 will be extrapolated.
3. Upper 90% confidence interval of scenario 1 will be extrapolated.

A 1st step was to study the process of cultivation in more detail. The following topics regarding cultivation were studied:

- Driving factors: A number of driving factors were identified, but five driving factors were considered most important: neighbouring cultivation, flood duration, geomorphology, elevation, and ecotope types.
- Trend: The historical increase in cultivated pasture was extrapolated to determine the expected yearly increase in cultivated pasture for the next 21 years. This extrapolation was used as input for the model during scenario 1. For scenario 2 and 3, the lower and upper 90% confidence interval of this extrapolation were determined.
- Regional differences: The rate of cultivation is higher in the eastern part of the Pantanal than in the western part
- Sub-processes: It was found that the process of cultivation was consists of three sub processes: first existing cultivated pastures expand at their edges, second new, isolated patches of cultivated pasture appear, and third a small number of cultivated pastures disappear.

A 2nd step was to design a conceptual model. The most important elements of this conceptual model are:

- Driving factors: The five driving factors that were considered most important were included in the model as driving factors.
- Regional differences: The model will take the regional differences into account.
- Sub processes: The model will include all three sub processes.

Subsequently, these decisions of the conceptual model were formalized into a working spatially explicit land use change model.

A 3rd step was to test the model. A sensitivity analysis and calibration were performed. The sensitivity analysis indicated that the model is not extremely sensitive to changes in the chosen parameters; all tested parameters have had a comparable (small) effect on the output after changing their value. The calibration indicated that the overall performance of the model is good, and that it is good enough to be used in scenario-studies.

Thus, the model was successfully prototyped and calibrated. However, running the scenarios proved to be a problem; the model could not adequately allocate the increase in cultivation as specified in the different scenarios. Probably, this is caused by the small differences in suitability for cultivation; most parts of the Pantanal are equally suitable according to the chosen criteria, and this makes allocation difficult.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

The Pantanal is a huge wetland that covers about 180.000 km^2 (Jongman et al. 2005). It spreads out across Brazil (provinces of Mato Grosso do Sul and Mato Grosso), Bolivia and Paraguay (Dolabella 2004). It is a bowl-shaped depression surrounded by highlands (Jongman et al. 2005; Dolabella 2004). The Pantanal and its highlands receive heavy rains from October to March; the Pantanal itself receives an annual rainfall of 1000-1400 ml while its highlands even receive an annual rainfall of 1500 ml. The Paraguay-river and its tributaries transport all this water through the Pantanal (Conservation International 2004). An enormous area is flooded each year because of this heavy seasonal rainfall and the small inclination of the Pantanal (Conservation International 2004).

The Pantanal provides habitat for many plant and animal species. Over 650 species of birds, 80 species of mammals, 50 species of reptiles and 250 species of fish have been reported for the area (Swarts 2000). The area is considered as one of the most biologically diverse regions in the world (WNF 2004). It is considered by the Brazilian government as a “natural heritage” and has therefore conservation priority (Seidl et al. 2001).

Although the ecological importance of the Pantanal has been recognized by the Brazilian government, 95% of it is in private hands (Seidl 2000). Cattle ranchers are the largest group of landowners and own about 80% of the Pantanal. The management of the Pantanal is therefore for a large part in the hands of cattle ranchers (Seidl et al. 2001).

1.2 Problem definition

Cattle ranching in the Pantanal has traditionally been extensive, using natural pastures and low densities of cattle (Calheiros 2004). This extensive form of ranching is considered to be relatively compatible with conservation objectives (WNF 2004). However, since the 1970's ranchers have started to clear land and to convert it to cultivated pastures. Although the economic drivers of this process are not yet clear as some authors questions its profitability, it does have environmental consequences. It is likely to have a negative impact on the biodiversity in the Pantanal (Seidl 2000; Seidl et al. 2001). In addition, it could also affect the hydrology of the area because of increasing sedimentation. This has already happened in the surrounding highlands, where cultivation has led to increased sedimentation rates in the rivers (Jongman et al. 2005).

Another issue is that recently the rivers of the Pantanal seem to have changed their behaviour and that flooding has become more frequent (Jongman et al. 2005). Ranchers suffer considerable damages from this flooding and demand a solution. However, any changes in the hydrology could seriously affect the biodiversity of the area. Regarding the enormous biodiversity value of the Pantanal, it would be irresponsible to make such changes without a good understanding of the potential consequences.

The Pantanal-Taquari project (part of the Water for Food and Ecosystems program; Jongman et al. 2005) aims to help the Brazilian government in finding sustainable solutions for these flooding problems. Several organizations (Alterra, ITC, WL Delft, etc.) cooperate in this project (Jongman et al. 2005). The project focuses on a sub region of the Pantanal, the lower Taquari.

The project’s overall objective is to develop tools that can support river management in the lower Taquari. Both a hydrological model and an ecological model have been developed.

With these models it is possible to assess the hydrological and ecological consequences of different river management options.
The project is also involved in studying land use because of its effects on hydrology and ecology. Up to now, however, little progress has been made in this direction. There is not yet a land use change model available that can be used in scenario-studies.
This study aims to prototype such a land use change model. The land use change model will be a separate model but must be able to use the output from the hydrological and the ecological model, and vice versa (the hydrological and ecological models must be able to use the output of the land use change model).

1.3 Research objectives

As stated, this study aims to prototype and validate a simulation model for the Pantanal which can explore future land use based on potential land use developments; i.e. to prototype a land use change model that can be used for scenario-studies.
If the model can be prototyped and validated, we will check the functionality of the model and explore future land use in three scenarios. These three scenarios will be:

1 The current trend of habitat conversion in the Pantanal will be extrapolated.
2 Lower 90% confidence interval of scenario 1 will be extrapolated.
3 Upper 90% confidence interval of scenario 1 will be extrapolated.

Then, we will use the output of our model (regarding these three scenarios) as input for the existing ecological model and thus make an ecological assessment.

1.4 Research questions

To achieve our research objective, this study will address a number of research questions:

1. Related to land use change in the Pantanal:
    - Which land use-trends are relevant in the Pantanal?
    - Which forces drive land use change in the Pantanal, and which indicators can adequately represent in the model?

2. Related to land use change models (in general):
    - Which types of land use change models exist and which type should be used?
    - How does this type of model work (general description)?
    - Which “hot” topics regarding this type of model could be included in a Pantanal Land Use model?
    - How have other studies implemented this type of model?

3. Related to our model:
    - What concept will I use to model land use change in the Pantanal?
    - How “robust” is this Pantanal Land Use model (sensitivity)?
    - How reliable is this Pantanal Land Use model (calibration)?

1.5 Report outlook

This report has the following outline:

- Chapter 2 will address the following research questions:
    - Which land use-trends are relevant in the Pantanal?
    - Which forces drive land use change in the Pantanal, and which indicators can adequately represent them in the model?

  First, chapter 2 will introduce the study area. Then, chapter 2 will address which land use-trends are relevant. It will be argued that the conversion of natural habitat into cultivated pasture is the most relevant process in the study area. Important aspects of this process will be discussed: i.e. it consists of three sub-processes, and differs between two sub-regions.

  Chapter 2 will also calculate the expected increase in cultivated pasture for the next years by extrapolating the historical increase of cultivated pasture. It will extrapolate three minor variations on the historical trend. In addition, it will discuss the results of these extrapolations concerning the two sub-processes and -regions mentioned above.

  The final part of chapter 2 will address the forces that drive the conversion of natural habitat into cultivated pasture. In addition, chapter 2 will discuss which driving factors can be used to represent this process.

- Chapter 3 will address the following research questions:
    - Which types of land use change models exist and which type should be used?
    - How does this type of model work (general description)?
    - Which "hot" topics regarding this type of model should be included in a Pantanal Land Use model?
    - How have other studies implemented this type of model?

  Chapter 3 will address these research questions by giving a general overview of the different types of land use change models. It will discuss the strengths and weaknesses of each type of land use change model. Based on this discussion, it will explain which type of model is considered most appropriate for modelling land use change in the Pantanal. Then, it will discuss the basics of this type of model in more detail.

  Finally, chapter 3 will discuss how other studies have implemented the chosen type of model by reviewing a number of existing land use change models.

- Chapter 4 will address the following research question:
    - What concept will I use to model land use change in the Pantanal?

  Chapter 4 will address this research question in two steps. First, it will discuss the conceptual model that was used to model land use change in the Pantanal. Secondly, it will discuss the formalization of the conceptual model.

- Chapter 5 will address the following research questions:
    - How "robust" is the Pantanal Land Use model (sensitivity)?
    - How reliable is the Pantanal Land Use model (validation)?

  Chapter 5 will address the "robustness" of the model by means of a "sensitivity analysis"; i.e. it will test if the model still produces sensible output if small changes in (1) the inputs of the model or (2) in the parameters of the model are made.

  Chapter 5 will address the reliability of the model by means of a validation. This validation will consist of modelling the historical period 1974-2001 and comparing the results of the model with the historical data. This will give an indication of the accurateness of the model.

2 Land Use Change Trends in the Pantanal

This chapter will address the research questions related to land use change in the Pantanal. These are:

- Which land use-trends are relevant in the Pantanal?
- Which forces drive land use change in the Pantanal, and which indicators could adequately represent them in the model?

The outline of this chapter is as following:
Paragraph 2.1 shortly introduces the study area. Paragraph 2.2 will explain that the most relevant land use change process in the Pantanal is the conversion of natural habitat into cultivated pasture. In the remaining paragraphs, the cultivation of natural habitat is discussed in detail. The following topics regarding cultivation are discussed:

- Driving forces: A number of forces drive the conversion of natural habitat into cultivated pasture and these are discussed. In addition, the discussion focuses on which indicators can adequately represent the driving forces (Paragraph 2.3).
- Trend: The historical increase in cultivated pasture is discussed. In addition, the historical increase in cultivated pasture is extrapolated to determine the expected increase in cultivated pasture in the future years (Paragraph 2.4).
- Regional differences: Regional differences in the cultivation of natural habitat are discussed (Paragraph 2.5).
- Sub-processes: The process of cultivation consist of a number of sub-processes and these are discussed (Paragraph 2.6).

2.1 Study area

As stated in the introduction, the Pantanal-Taquari project aims to find sustainable solutions for flooding in the Taquari in the Brazilian Pantanal (figure 2.1.1).

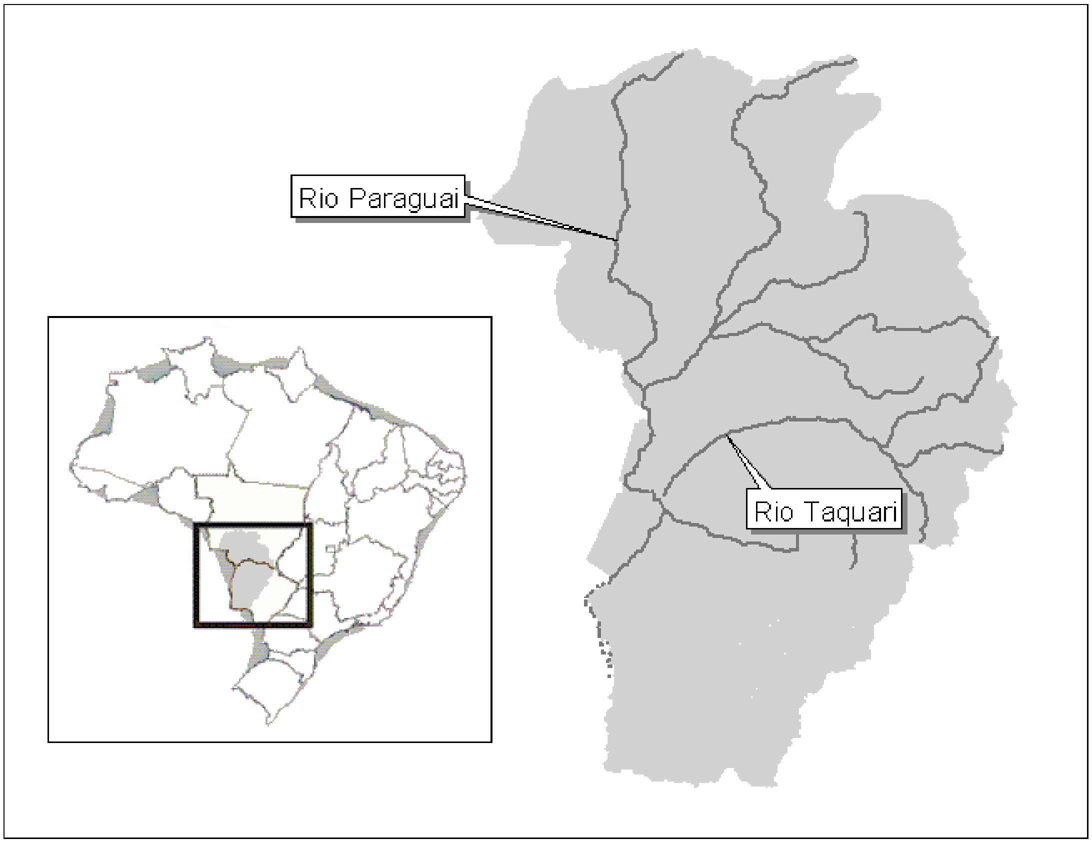


Figure 2.1.1: Location of the Pantanal in Brazil (left) and close-up of the Brazilian Pantanal (right).

Problems concerning flooding are most serious in the lower part of the Taquari and this location is therefore the focus of the Pantanal-Taquari project (Jongman et al. 2005). Since our study is a spin-off of the Pantanal-Taquari project, it also focuses on this area. The "natural" boundaries of the lower Taquari were used as much as possible as the boundaries of our study area (figure 2.1.2). The study area is 5211100 ha.

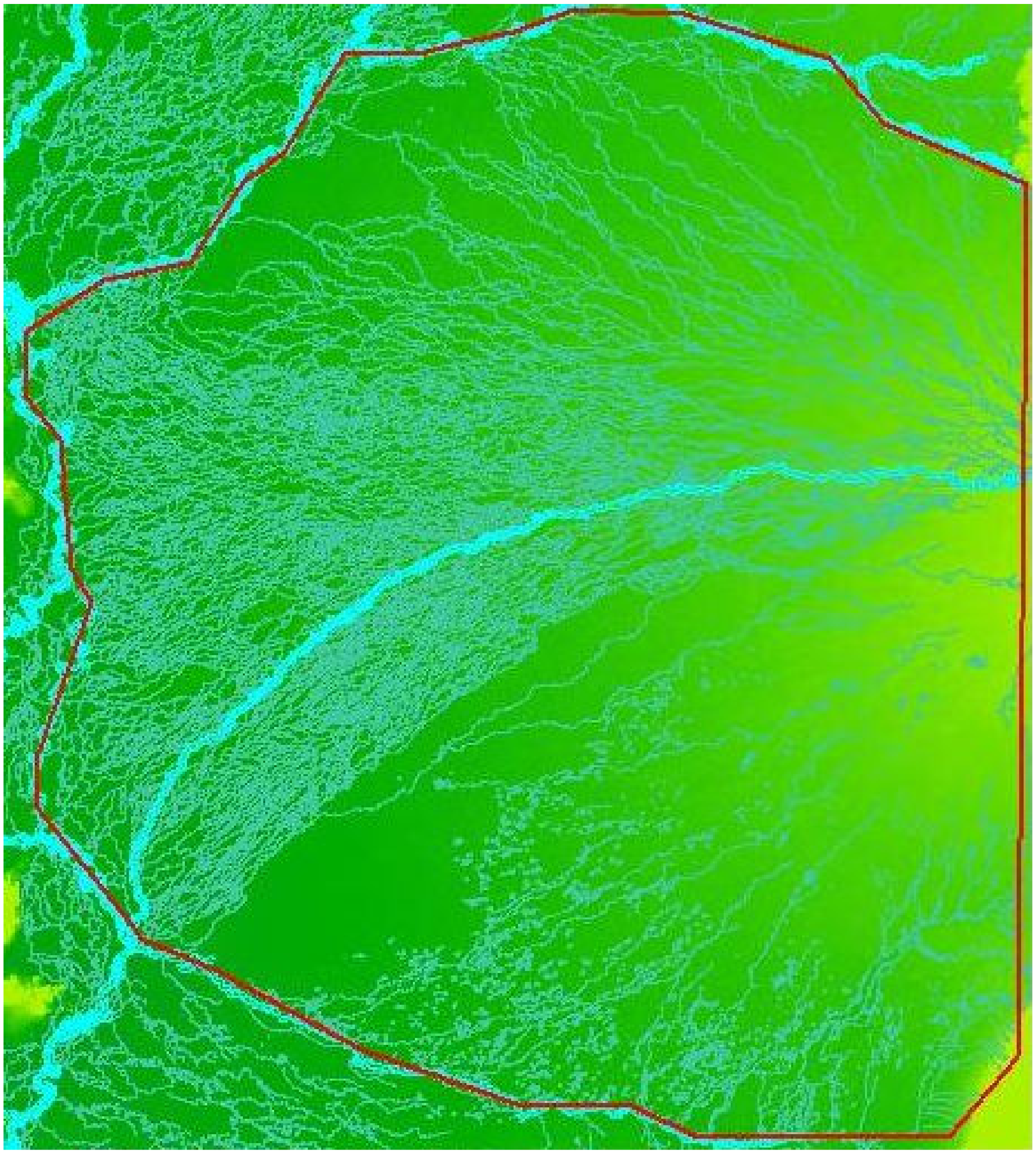


Figure 2.1.2: The lower Taquari and some additional rivers (all blue). This is the area where problems concerning flooding are most serious. It is the focus of the Pantanal-Taquari project and includes our study area (red). The study area is 5211100 ha. The figure also clearly shows areas with lower elevation (green) and with higher elevation (yellow).

2.2 Relevant land use change trends in the Pantanal

This paragraph will address the following research question:

- Which land use-trends are relevant in the Pantanal?

As stated, cattle ranching has traditionally been extensive in the Pantanal. Natural pastures are used and stocked with only one animal per 3.6 hectares (Calheiros 2004). The traditional form of cattle ranching is relatively compatible with conservation objectives (WNF 2004). However, since the 1970's ranchers have started to clear land and to convert it to cultivated pasture.

The economic motivation of ranchers to construct cultivated pasture is not yet clear. Although ranchers can keep more cattle on cultivated pasture than on natural pasture, cultivated pastures also require investments. In addition, long periods of flooding (>3 months) can severely damage cultivated pastures and render them useless. Thus, some authors question the profitability of cultivated pasture. They argue that because ranchers always have considered ranching as the only possible economic activity in the Pantanal, they tend to overlook other alternatives and do not question the profitability of this new form of ranching (Seidl 2000). Other research (Seidl et al. 2001) showed that cultivated pasture is profitable, but these researchers admitted that important information lacked in their study. They state that their study is merely an exploration into the economic incentives ranchers face (Seidl et al. 2001). In short, good information regarding the economic incentives of ranchers is very scarce and the economic motivation of ranchers to construct cultivated pasture is still obscure.

Nevertheless, the area of cultivated pasture is increasing (Seidl 2000). This has resulted in a decrease of woodland and natural pasture (Seidl 2000; Seidl et al. 2001). The higher grounds are most likely to be cultivated because they are not regularly flooded (Seidl et al. 2001). These same grounds also tend to be covered with forest because of their location; indeed has more than 13% of the woodland in the Pantanal been lost (Seidl et al. 2001). Land use maps from 1976, 1984, 1991 and 2000 clearly reflect that the conversion of natural habitat into cultivated pasture is an important trend in the Pantanal (figures 2.2.1 and 2.2.2). Local experts (W. Tomas from EMBRAPA; C. Padovani from EMBRAPA; M. van Eupen from Alterra) confirm this as well.

Besides cultivated pasture, another form of cultivation -soy plantation- has emerged. However, soy plantations cover only a tiny fraction of the land inside the Pantanal (see Seidl 2000) and they are mainly found in the surrounding highlands where they cover a huge percentage of the area (pers. comm. M. van Eupen). Therefore, we will not consider them in this study. We will assume that the conversion of natural habitat into cultivated pasture is the main land use change trend and limit our model to this. The next paragraph will discuss it extensively.

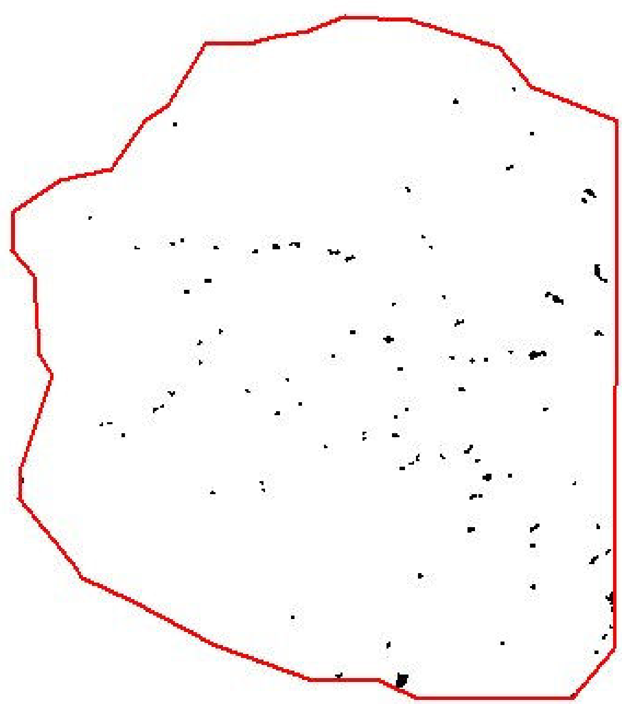


Figure 2.2.1. Cultivated pastures in study area in 1976. Cultivated pastures are black and the study area red.

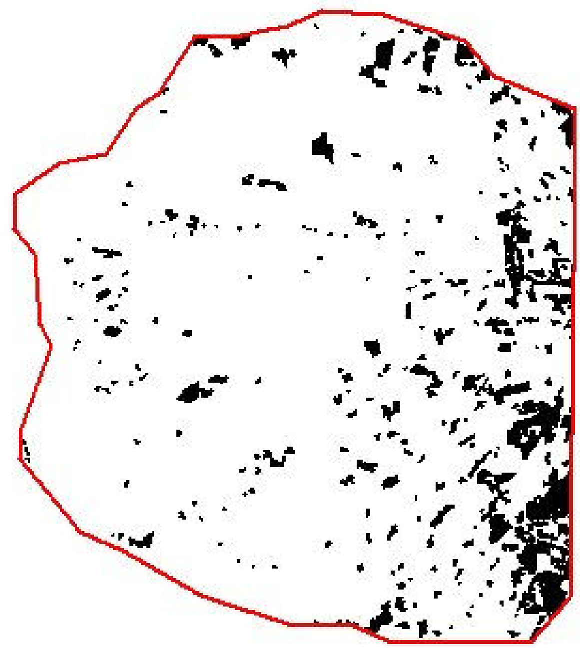


Figure 2.2.2. Cultivated pastures in study area in 2000. Cultivated pastures are black and the study area red.

2.3 Cultivated pasture: which forces drive cultivation and which indicators can represent them?

This paragraph will focus on the research question:

- Which forces drive land use change in the Pantanal, and which indicators could adequately represent them in the model?

First, this paragraph will address the forces that drive the conversion of natural habitat into cultivated pasture in the Pantanal. Then, this paragraph will discuss which indicators can be used to represent this process, and which shall ultimately be included in the model.

2.3.1 Possible driving forces

According to Bürgi et al. (2004), driving forces are the forces that cause noticeable changes in a landscape; i.e. they are the (underlying) causes of change in a landscape and as such, they drive the changes in a landscape. Bürgi et al. (2004) discerned five major groups of driving forces: socioeconomic, political, technological, natural and cultural driving forces. Socioeconomic driving forces are the economic forces that are at play, and include examples such as the increasing globalization, market economy, etc. (Bürgi et al. 2004). Political driving forces are mostly tied to socioeconomic driving forces, as most policies have socioeconomic goals. The introduction of new technologies can also have a serious effect on a landscape; these are technological driving forces. Natural driving forces are all the natural (non-human) forces that shape a landscape. Examples are climate, topography, etc. Cultural driving forces are more difficult to define, but it is clear that culture also shapes a landscape (Bürgi et al. 2004).

It is difficult to say which of these driving forces plays a role in the land use changes in the Pantanal. However, it seems apparent that socioeconomic forces play a role; globalization and intensification are influential forces in many parts of the world and are likely to have an influence in the Pantanal. Nonetheless, good information regarding the economic incentives of ranchers is very scarce and the motivations of ranchers to construct cultivated pasture are still obscure; it is therefore impossible to draw any definite conclusions on this issue.

2.3.2 Possible indicators of land use change

The former paragraph discussed the forces, which drive the conversion of natural habitat into cultivated pasture in the Pantanal. This paragraph will discuss which indicators can be used to represent this process. It starts by discussing the indicators that are generally used in land use change models, and then discusses the indicators that can be used in the Pantanal model. Hereafter, the indicators that are used to represent land use change are referred to as driving factors.

Many land use change models focus on the conversion of natural habitat into cultivated pasture (e.g. Soares-Filho et al. 2002; Mas et al. 2004; Veldkamp et al. 2001). Especially in Latin America is the expansion of pasture an important land use change trend (Angelsen & Kaimowitz 1999; Lambin et al. 2001). Most of these models include a number of biophysical variables as driving factors of land use change, such as soil, vegetation, altitude, slope, distance to forest / non-forest edge, etc. Some models have an economic orientation and include economic variables (e.g. agricultural prices, credit availability, etc.) or approximations of them (e.g., distance to roads is an approximation for market access. However, using these same driving factors without further consideration in our model is not sensible. This is for a number of reasons.

Most models have empirically derived the relations between driving factors and land use change; this means that the relation between driving factors and land use change needs to be qualified for each new study area all over again. Most of the models above also focus on deforestation and although the Pantanal has some forest cover, its main land cover consists of savannah ("cerrado"); this is an important difference. Moreover, flooding is an essential factor in the Pantanal (pers. comm. M. van Eupen; pers. comm. C. Padovani) but this factor is absent in most other models.

We have therefore consulted local experts about which variables to include as driving factors of land use change in our model. The use of expert knowledge is, together with statistical techniques (logistic regression) a common approach in simulation models to quantify the relation between land use change and its driving factors (Irwin & Geoghegan 2001). We have consulted C. Padovani, W. Tomas (EMBRAPA), M. van Eupen (Alterra) and A. Seidl (Colorado State University).

After consulting the experts above and looking at the current location of cultivated pastures in the Pantanal, we narrowed the number of variables that could be relevant as driving factors down to the following:

- Neighbouring cultivation; looking at land use maps of the Pantanal over the period 1976-1991, it is obvious that cultivation "spreads like an oil-spill" and that land is more likely to be cultivated if surrounding lands are already cultivated.
- Flooding risk; long periods of flooding (>3 months) can severely damage cultivated pastures. Converting natural habitat into cultivated pastures requires large investments and farmers will make such investments only if they can expect a high output. Thus, they would prefer areas that have a small risk of flooding.
- Elevation also seems to play an important role regarding flooding risk; a location can lie in an area that is frequently flooded (i.e. an area with a high flooding risk), but if the location itself is much higher than its surroundings -and thus save from flooding-, the location might still be attractive for cultivation.

- Soil type; some soils are too poor in nutrients to make good pastures. Such pastures could sustain only small numbers of cattle and this makes them unattractive for cultivation (pers. comm. van Eupen).
- Costs of cultivation; converting natural habitat into cultivated pastures requires large investments. However, these “costs of cultivation” differs between areas and makes certain areas more attractive for cultivation than others.
- Rancher’s wealth; converting natural habitat into cultivated pastures requires large investments. Wealthy ranchers will have more opportunities to invest in cultivation. We consider the size and stocking rate (number of cattle) of a ranch as a good approximation of the wealth of a rancher. Indeed, Seidl (2000) found that both variables correlate with habitat conversion.
- Accessibility; inaccessible regions are less suitable for cultivation because of their distance to markets. In the Pantanal, the accessibility of a location is determined by the distance to (1) a road or (2) an accessible waterway (not all rivers in the Pantanal are accessible by boat; e.g. many are too shallow).
- Ecotype; some vegetation types are more suitable for cattle than others. For instance, forests require levelling before they can be used as pasture, and many wetlands are often too wet to allow high densities of cattle. Therefore, ranchers might have a preference for ecotypes that require less input.

2.3.3 Driving factors included in the model

After many discussions with the local experts, it was decided that the following variables were the main driving factors and thus should be included in the model (ranked to importance):

- Neighbouring cultivation.
- Flooding risk.
- Elevation.
- Ecotype.

This choice might seem a bit arbitrary. However, one must remember that the process of land use change is extremely complex and that the causal relations underlying land use change are often not (fully) known (Veldkamp & Lambin 2001; Geist & Lambin 2002). Selecting the appropriate driving factors is therefore not an “exact science”, which explains the use of expert knowledge.

Therefore, we trust the expertise of our local experts and start modelling with the chosen driving factors. We calibrated and tested the model and checked if the output of the model made sense. This is discussed in detail in chapter 5.

2.4 Cultivated pasture: extrapolation of historical trend

Introduction & methodology

As stated, the area of cultivated pasture is increasing. However, how will the area of cultivated pasture increase during the next years? In other words, how will the increase in cultivated pasture develop? A way to find this out is to look at the historical trend and extrapolate it.

Land use maps are available from the years 1976, 1984, 1991 and 2000. The exact amount of cultivated pasture in these years can be estimated from these land use maps. This shows the historical growth in cultivated pasture (figure 2.4.1). A common approach in predicting future growth is to plot a trend line through historical data points and extrapolate it (Verburg[2] et al.

2004). Different types of trend lines can adequately describe landscape changes (Bürgi et al. 2004). The simplest type is a linear trend.
In this study, a linear trend could be suitable. A linear trend describes a process that increases (or decreases) at the same rate; i.e. it has a constant growth rate. If we would assume that yearly a steady amount of ranchers start to convert natural habitat into cultivated pasture, a linear trend could be appropriate in describing this process. Moreover, a linear trend seems to fit the historical data points well. Hence, in the next sub-paragraph, we will fit a linear trend line through the data points and discuss the results. The trend line will be extrapolated. This extrapolation will cover a period of 21 years (from 2000 to 2021). Extrapolation over a longer period gives too speculative results.

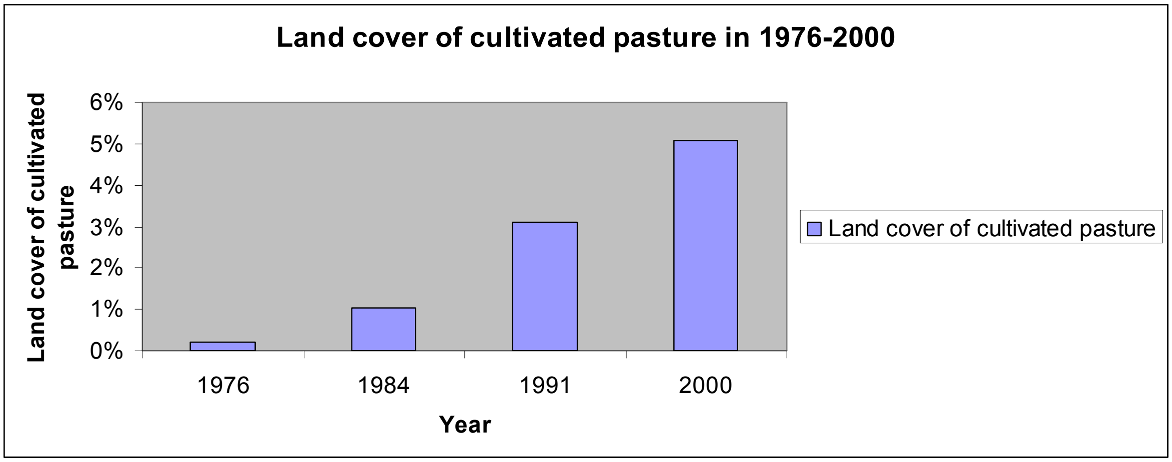


Figure 2.4.1. Historical growth of cultivated pasture. The land cover of cultivated pasture is shown.

Results
Fitting a linear trend through the data points resulted in a statistically significant trend ($p < 0.05$) (figure 2.4.2). According to this trend, the area cultivated pasture increases with 10988 ha per year, which is 0.21% of the total study area (table I). Extrapolation of this linear trend indicates that cultivated pasture will cover about 9% of the Pantanal in 2021 (figure 2.4.3). This is a realistic scenario.
In addition, we have fitted trend lines through the 90% upper confidence interval and the 90% lower confidence interval of the data points (figure 2.4.3). All linear trend lines were statistically significant ($p < 0.05$) and gave realistic results when extrapolated (figure 2.4.3).

Table 1. Expected yearly increase in cultivated pasture given in hectares and as percentage of total study area. The expected yearly increase was determined by extrapolating the historical trend.

Yearly increase in cultivation (ha)	Total area of study area (ha)	Yearly increase in cultivation (percentage of total study area)
10988 ha	5211100 ha	0.21 %

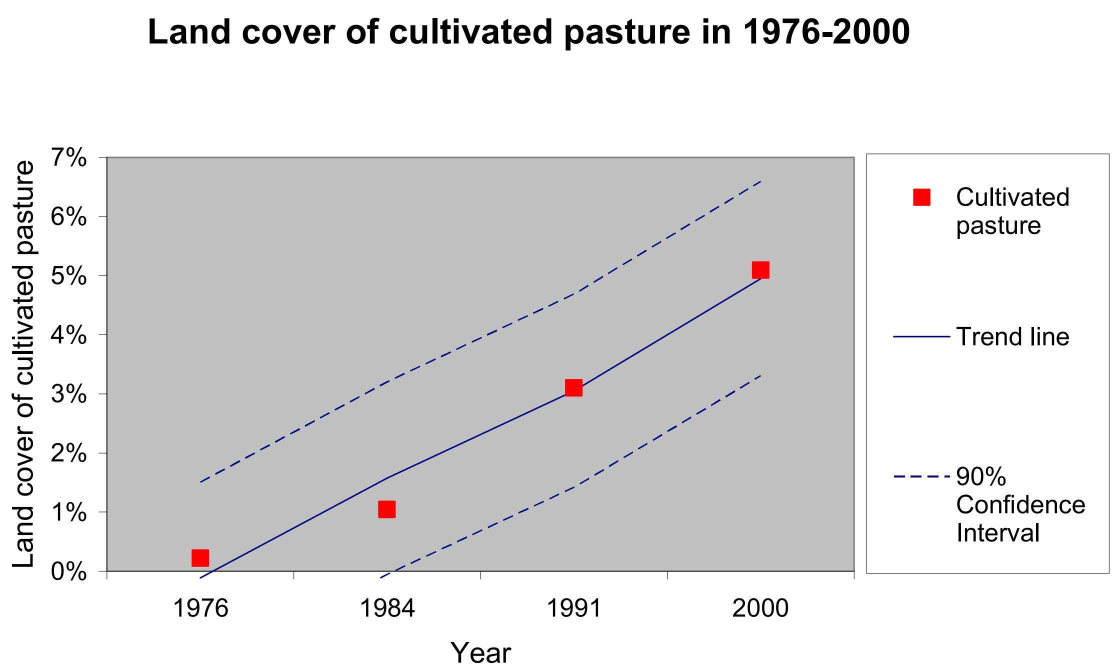


Figure 2.4.2: Area of cultivated pasture in 1976, 1984, 1991 and 2000 plus the corresponding linear trend line. In addition, linear trend lines through the 90% upper confidence interval and the 90% lower confidence interval of the data points are shown.

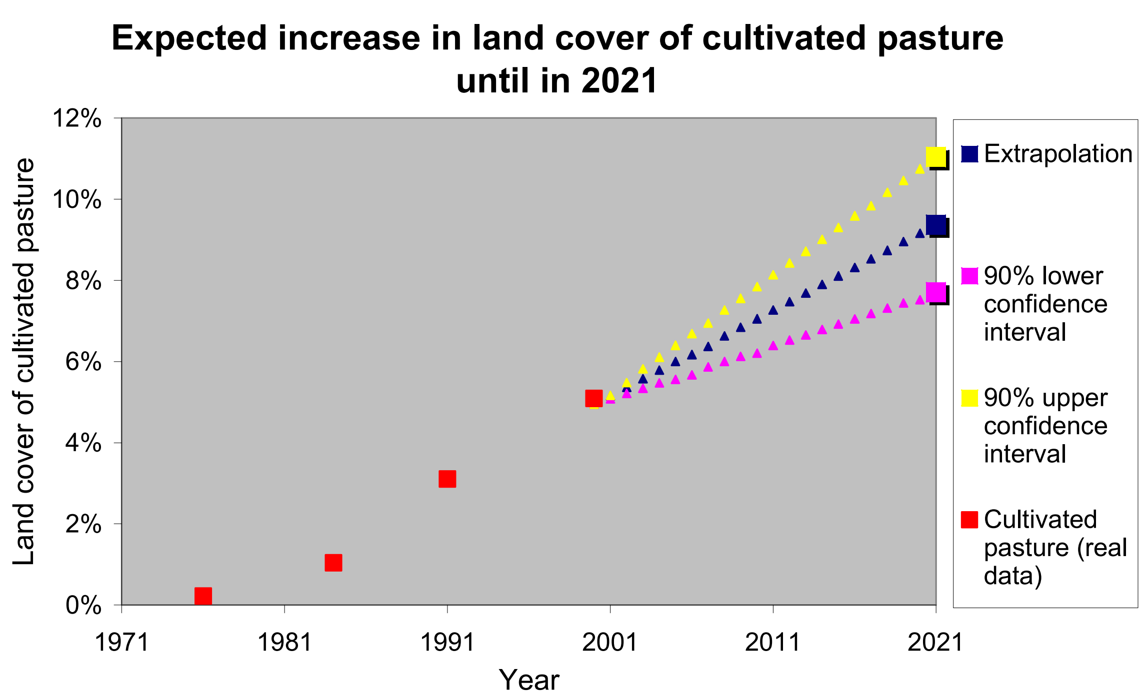


Figure 2.4.3: Extrapolation of the linear trend line through the data points. In addition, the extrapolation of the linear trend lines through respectively the 90% upper confidence interval and the 90% lower confidence interval of the data points are shown. The area of cultivated pasture in 1976, 1984, 1991 and 2000 has also been included.

Discussion
The results of the linear extrapolations are presented in figure g. This figure shows relatively small differences between the extrapolations. This suggests a relatively small confidence interval and reliable results. However, confidence intervals should only be calculated when the sample size is not too small (Campbell & Gardner 1988). This is not the case, and these results should not automatically be considered as very reliable. Moreover, the trend line was fitted through only 4 points. Determining a trend based on only 4 points is difficult. Such a trend contains a lot of uncertainty. Therefore, although all trend lines are statistically significant, they are not extremely reliable.
Based on the above, we feel that the results of these extrapolations must be viewed tentatively. This is, however, not a problem for this study. As stated in the introduction, this model aims to explore the consequences of potential land use developments. This is done by running three possible scenarios (according to the trend, lower 90% confidence interval of the trend, and upper 90% confidence interval of the trend). Scenario studies are reconnaissance studies that explore expected or even thinkable situations in the (near) future. For that reason the quantitative reliability isn't the most important item; i.e., this approach (scenario study) allows a certain degree of uncertainty. Therefore, these results can be used further on in this study as input for the model in scenario-studies.

2.5 Cultivated pasture: regional differences

As stated, the area of cultivated pasture increases each year. However, regional differences seem to exist; a general comparison of land use maps indicates that this increase differs between the western and eastern part of the study area. It seems that the area of cultivated pasture in the eastern part increases at a higher rate. This hypothesis was tested as follows:

1. The study area was divided in a western and an eastern region (figure 2.5.1). The boundary was roughly based on geo-morphological patterns and processes that already formed a "natural boundary" between both regions.
2. The area of cultivated pasture in 1991 and 2001 was determined for each region. The increase in cultivated pasture was calculated for this period.
3. The results were compared between the two regions.

Table I shows the results. In the western region, land cover of cultivated pasture increased from 1.3% in 1991 to 1.7% in 2000. In the eastern region, land cover of cultivated pasture increased from 7.7% in 1991 to 13.7% in 2000. Thus, the increase in land cover of cultivated pasture was 131% in the western region and 178% in the eastern region.

The conclusion is that the area cultivated pasture growths faster in the eastern part of our study area. There are a number of thinkable hypotheses for this.

- The eastern part of the study area is bordering a well-developed agricultural area, the Planalto. During the 1970's the government stimulated colonization of the Planalto heavily, and it has become well developed in terms of infrastructure and commerce (Jongman et al. 2005). The western part of the study area borders Bolivian and Paraguayan parts of the Pantanal, which are less developed. Consequently, the eastern part of our study area is better accessible and closer to important commercial markets; this might stimulate ranchers to construct cultivated pastures.
- The eastern part of the study area is higher and therefore suffers less from flooding. Long periods of flooding (>3 months) severely damage cultivated pastures (Jongman et al. 2005). Constructing cultivated pasture requires large investments and ranchers will make

such investments only if they can expect a high output. Thus, they will prefer areas with a small risk of flooding, which are the higher grounds.

Figure 2.5.1. Study area (red outline) divided in a western and an eastern region.

Table II. Land cover of cultivated pasture in the two regions in 1991 and 2000. The increase in area of cultivated pasture is also given; this indicates how much percent the area cultivated pasture has increased during this period.

Region	Land cover pasture '91	Land cover pasture '00	Increase in land cover
Western	1.3 %	1.7 %	131 %
Eastern	7.7 %	13.7 %	178 %

As stated, the main conclusion is that local differences in the growth rate of cultivated pasture exist. This is an important conclusion; it places the results of paragraph 2.2.4 in a different perspective. In this paragraph, the expected increase in cultivated pasture was calculated by extrapolating its historical trend. Regional differences were not considered. Instead, the expected increase in cultivated pasture was calculated for the entire study area. Since it is clear that the area cultivated pasture increases faster in the eastern part of the study area than in the western part, the results of the extrapolation should be adjusted for this. This is especially necessary since this study aims to use the results of the extrapolation as input for the model in scenario-studies; this study needs to know the expected increase in each region. Ideally, a new extrapolation should be done separately for each region to get the expected increase. However, this was not possible due to time constraints. Therefore, this study uses a different approach. The following paragraphs will explain this in detail.
The expected increase in cultivated pasture was already calculated for the entire study area. However, it is clear that cultivated pasture increases faster in the eastern part of the study area than in the western part. Thus, the expected increase should be divided or "allocated" between both regions; the increase in the western part will have to become less than calculated, whereas in the eastern part the increase will be greater than calculated. This partition will be based on an allocation formula, which is inferred from the increase of cultivated pasture in

each region during the period 1991-2000. During this period, cultivated pasture increased in the eastern region with 14924 ha and in the western region with 88806 ha. This means that the western region has accounted for 14.39% of the total, whereas the eastern region for 85.61% of the total (table II). Consequently, we assume that 14.39% of the expected yearly increase will take place in the western region and 85.61% of the expected increase will take place in the eastern region. The corresponding increase in cultivated pasture in each region is shown in figure 2.5.2.

Table III. Increase in cultivated pasture in the two regions (in ha). In addition, the part of each region in the total increase is given.

Region	Increase pasture '91-'00	Part of total increase
Western	14924 ha	14.39 %
Eastern	88806 ha	85.61 %
Combined	103730 ha	100 %

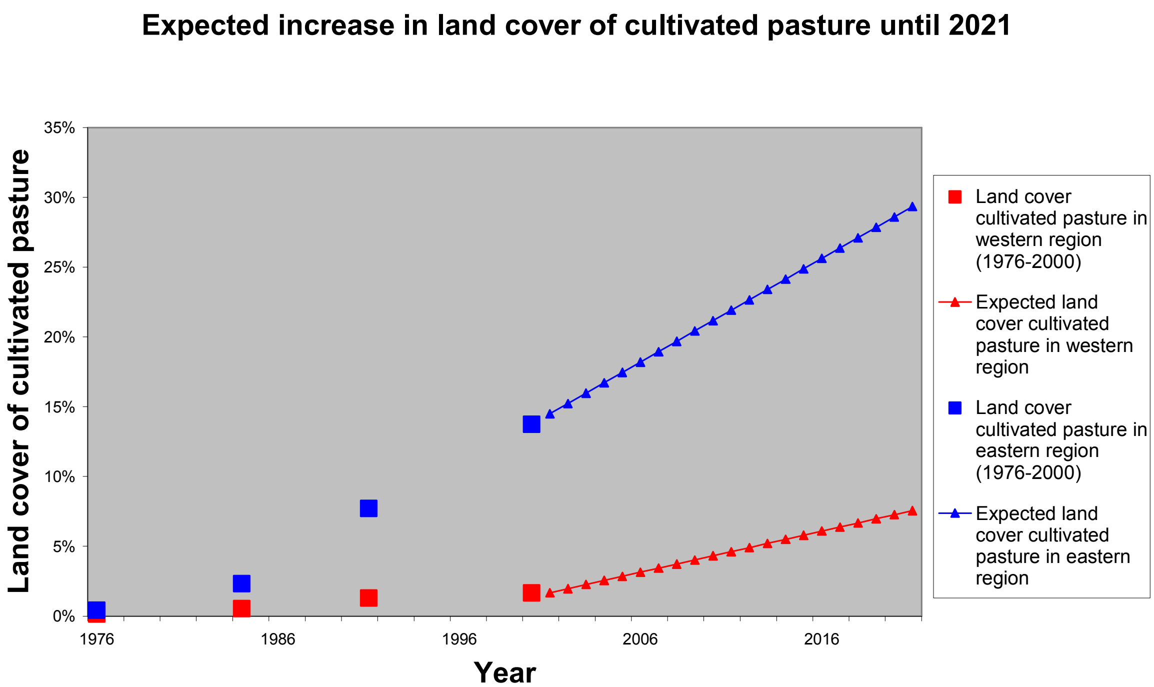


Figure 2.5.2. Expected increase in cultivated pasture in the two regions based on the calculations described in the text. In addition, historical data (1976-2000) regarding the land cover of cultivated pasture are given for both regions.

2.6 Cultivated pasture: sub-processes

As stated, the most relevant land use change process in the Pantanal is the conversion of natural habitat into cultivated pasture. At a large- scale, it is clear that the area of cultivated pasture increases each year. At a finer scale, this process is not as straightforward. A number of sub-processes can be found. These are now discussed.

"Expansion" vs. "patching" of cultivated pastures

As stated, the area cultivated pasture increases each year. This process consists of two sub-processes: "Expansion" and "Patching". "Expansion" is the process in which existing areas of cultivated pasture expand. “Patching” is the process in which new, isolated patches of cultivated pasture emerge away from existing areas of cultivated pasture. Together, these processes constitute the increase in cultivated pasture. Soares-Filho et al. (2002) did similar observations regarding deforestation; they found that deforestation in the Amazon could be divided into two separate processes: deforestation occurring along the edges of a “deforestation front” and deforestation occurring through new, isolated patches of deforestation. Wu (2002) found urban land developments also consist of “Patching” and “Expansion”-like sub processes.

The next section will determine how much the sub processes of respectively “Expansion” and “Patching” contribute to the total increase of cultivated pasture in the Pantanal. Of course, this will be determined separately for the western and eastern region, as large regional differences exist in the increase in cultivated pasture (see paragraph 3.2.2).

The increase through the process of “Expansion” will be compared with the increase through “Patching” during 1991-2000. This period was chosen for the analysis because (1) time pressure did not allow a comparison of more data sets and (2) this data was most recent and thus most representative of the current process.

The increase through the process of “Expansion” was determined by selecting cultivated pastures, which had expanded during this period. Their expansion was calculated.

The increase through the process of “Patching” was determined by selecting cultivated pastures, which did not exist in 1991 and suddenly emerged in 2000. Their area was calculated. The increase through both processes was compared separately for the western and eastern region.

Table III shows the results. In the western region, “Expansion” accounted for 14.2% of the increase, whereas “Patching” accounted for 85.8% of the total increase in the region. Apparently, the process of “Expansion” plays a much smaller role than “Patching” in this region.

In the eastern region, “Expansion” accounts for 53.4% of the increase and “Patching” for 46.6% of the increase in cultivated pasture in the region; “Expansion” and “Patching” play an equally large role in this region.

Table IV. Increase in cultivated pasture through the processes of “Expansion” and “Patching” during 1991-2000 (separately for the western and eastern region of the study area). Increase is given absolutely (in ha) and as a percentage of the total increase in the region.

	Expansion	**Patching**	**Total increase (combined)**
Western region	5065 ha (= 14.2 %)	30638 (= 85.8 %)	50681725 ha (= 100 %)
Eastern region	52206 ha (= 53.4 %)	45561 (= 46.6 %)	522108517 ha (= 100 %)

<u>Expansion of cultivated pastures</u>

As stated, the increase in cultivated pasture is achieved through two sub-processes: “Expansion” and “Patching”. “Expansion” is the process in which existing areas of cultivated pasture expand. In this section, important characteristics of the process of “Expansion” are discussed. These characteristics were determined after a short analysis.

The analysis was based on a comparison between land use maps of 1991 and 2000. In this analysis, cultivated pastures were selected which were present both in 1991 and 2000; they are the result of “Expansion”. The general characteristics of these selected cultivated pastures were determined.

The results of the analysis are:

- Most of the existing cultivated pastures have increased considerably during the period 1991-2000. In fact, their radius has expanded on average 1000 metres during 1991– 2001, which equals approximately 100 meters per year.

- "Expansion" plays mainly a role in the eastern region of the study area. Although this was already concluded in the section above, this is also clear by visually comparing the land use maps.

<u>"Patching" of cultivated pasture</u>

"Patching" is the process in which new, isolated patches of cultivated pasture emerge away from existing areas of cultivated pasture. In this section, important characteristics of the process of "Patching" are discussed. These characteristics were determined after a short analysis.

The analysis was based on a comparison between land use maps of 1991 and 2000. In this analysis, cultivated pastures were selected which did not exist in 1991 and suddenly emerged in 2000; they are the result of "Patching. The general characteristics of these selected patches of cultivated pasture were determined.

The results of the analysis are:

- Figure 2.6.1 presents a histogram of the area of the patches. This figure clearly shows that the area of the patches has quite a large variation; they range from a few hectares to more than 600 ha (figure j) and this is quite large range. Moreover, the distribution is quite skewed, which makes it difficult to determine a "mean area" of the patches. However, none of the patches has an area larger than 700 ha.
  In addition, it is striking that the patches do not have a "minimum" area and that even extremely small patches (< 0.02 ha) exist; it seemed apparent to us that ranchers will not take the trouble of moving equipment and personnel to a location to construct a cultivated pasture as small as 0.02 ha. Either this assumption is not true or a different factor plays a role. For instance, these extremely small patches could be located very close to large expanding areas of cultivated pasture. The latter are not included in this analysis because they are considered the result of the process of "Expansion". However, it makes sense that a rancher would construct a small pasture at such a location; i.e. in that case he is merely expanding an existing pasture. Further research has to show this really works.
- Figure 2.6.2 presents a histogram of the distance of the patches to the nearest cultivated pastures of "Expansion". The aim was to find if patches always emerge at a fixed distance from the cultivated pastures of "Expansion". However, this figure clearly shows that this distance is very variable; it ranges from a fraction of a kilometre to more than 40 km (figure k) and this is very large range. Again, this distribution is also quite skewed, which makes it difficult to determine a "mean distance".
- Figure 2.6.3 presents a histogram of the elevation of the patches. This figure shows that most of the patches have an elevation higher than 110 m. This makes sense, as cultivated pastures are vulnerable to flooding and elevated places tend to flood less often.
- Figure 2.6.4 presents a histogram of the land cover* of both the patches and the study area. This figure shows that some ecotypes are overrepresented in the patches; i.e. they have a higher land cover in the patches than in the study area. Mainly different types of savannah are overrepresented, whereas different types of forest and wetlands are underrepresented. This outcome is logical, as savannah is most suitable for pastures; forests require additional effort because forested areas needs levelling before they can be used as pasture and wetlands are often too wet to allow high densities of cattle.

---

* Land cover derived from Jongman et al. 2005; "current situation scenario".

- Figure 2.4.5 presents a histogram of the flood duration* of the patches. This figure shows that most of the patches have flood duration less than 2 months per year. This makes sense; cultivated pastures are vulnerable to flooding, and ranchers will tend to select locations that do not flood often.
- A short comparison of the geomorphology of the patches and the study area revealed that none of the geo-morphological classes are over-or underrepresented in the patches; i.e. all geo-morphological classes were present in the same degree in the patches as in the study area.

From these results, the conclusion can be drawn that most patches are constructed at locations that are (1) higher than 110 cm, (2) covered with savannah, and (3) are flooded less than 2 months per year. Apparently, the geo-morphological class is not important. In addition, most constructed patches are quite small (< 10 ha) but they can also be quite large (> 700 ha). A mean area of the constructed patches can not be given as the variation is too large. The latter is also true for the distance of the constructed patches to the nearest "Expansion"-cultivated pastures; most constructed patches are quite nearby them (< 1 km) but they also be quite far away (> 40 km). A mean distance can also not be given as the variation is too large.

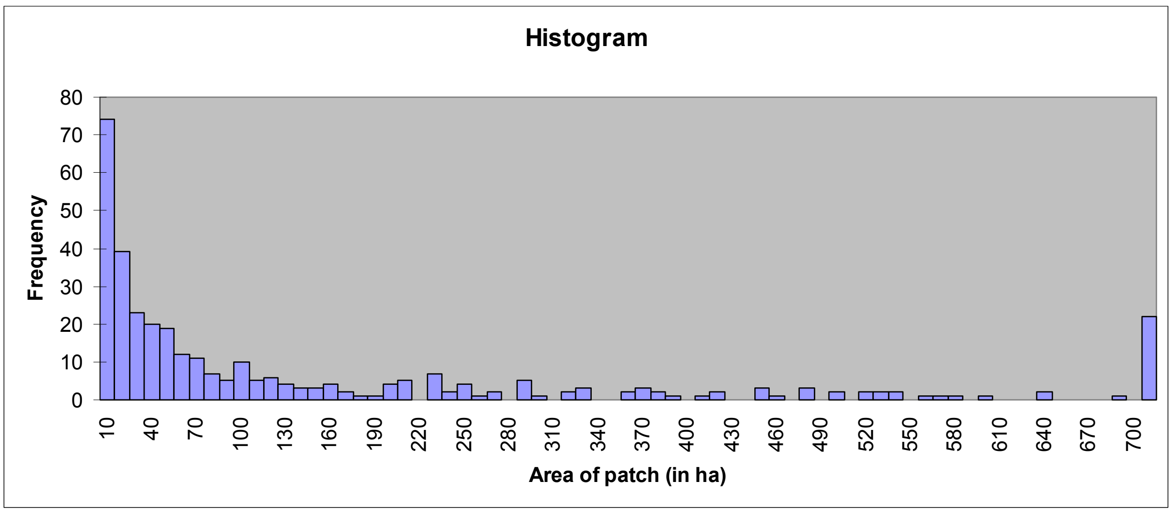


Figure 2.6.1. Histogram of the area of the patches of cultivated pasture, which emerged through the process of "Patching" during 1991-2000.

* Flood duration derived from: Jongman et al. 2005, "wet scenario".

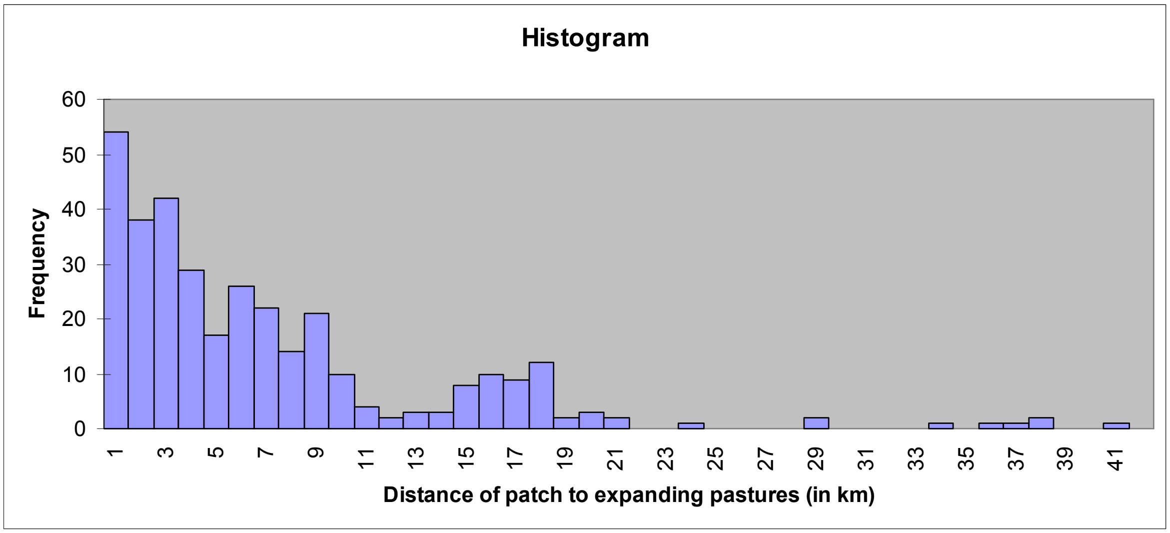


Figure 2.6.2. Histogram of the distance of the selected patches to the nearest expanding cultivated pastures.

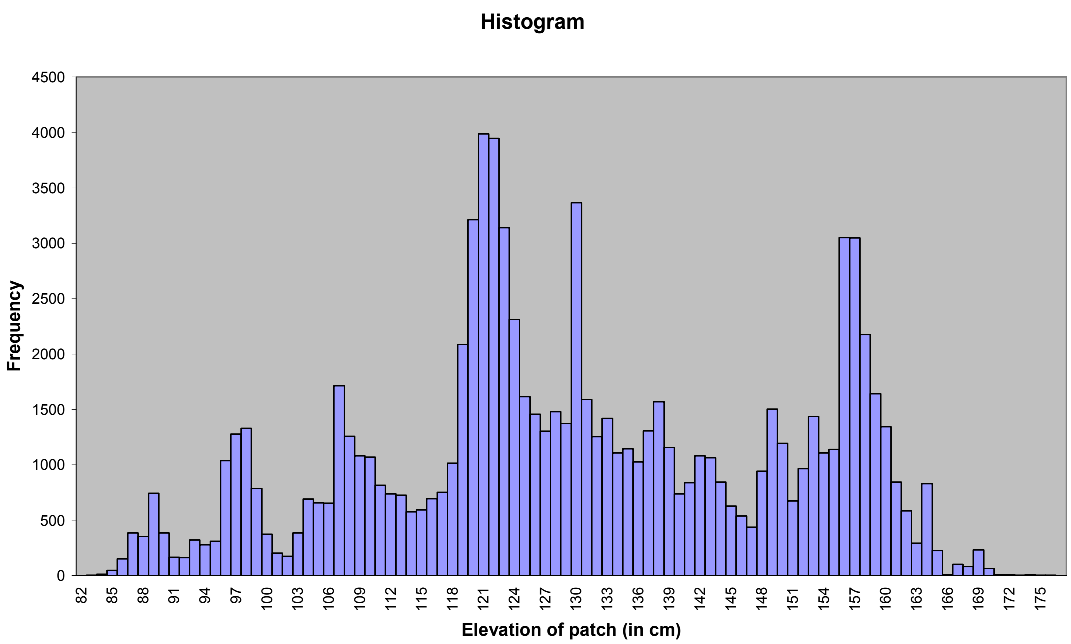


Figure 2.6.3. Elevation of patches, which emerged through the process of "Patching" during 1991-2000.

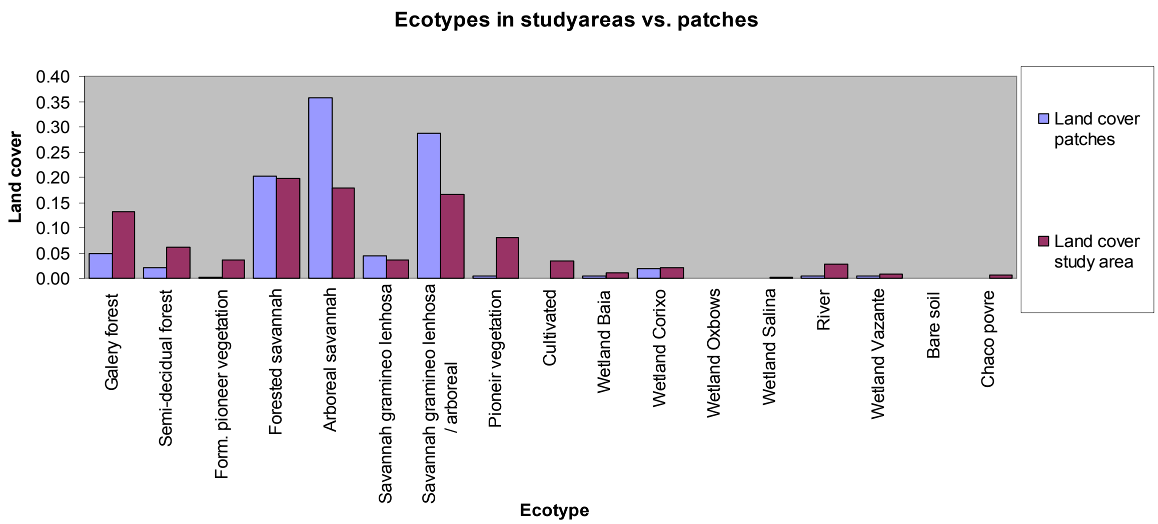


Figure 2.6.4. Land cover of the selected patches compared to land cover of the study area. Some ecotypes are overrepresented in the patches.

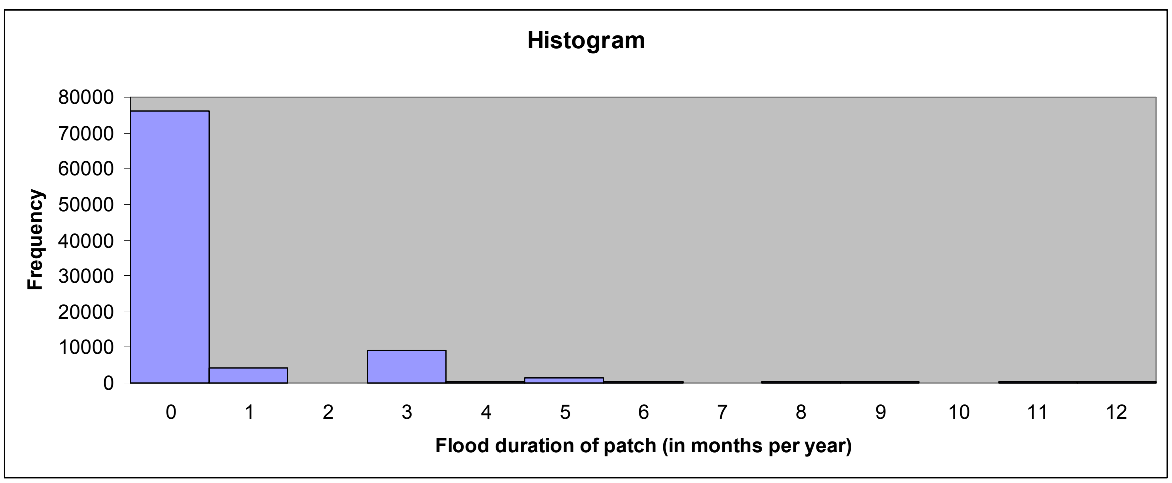


Figure 2.6.5. Flood duration of patches, which emerged through the process of "Patching" during 1991-2000.

Abandonment of cultivated pastures

Although the area cultivated pasture increases each year in the Pantanal, there is a small number of locations were cultivated pastures disappear. However, please notice that the area of cultivated pasture that is lost is many times smaller than the yearly increase in cultivated pasture.

The reason for this disappearance is not clear. The most logical explanation is that it is because of flooding; cultivated pastures become severely damaged if they are flooded more than 3 months (Jongman et al. 2005). It seems likely that such pastures are abandoned. However, this is merely a speculation. Additional research should look into this.

The disappearance of cultivated pasture might also have been the result of small digitizing mistakes; small mistakes might have been made in digitizing the 1991 land use map, and since it is unlikely that exactly the same mistakes occur twice, such areas will have been disappeared from the 2000 land use map. This can account for the disappearance of some areas of cultivated pasture. However, since such digitizing mistakes are very small, they cannot explain the disappearance of cultivated pasture entirely; it seems likely that a (small) area of cultivated pasture has really disappeared during the period 1991-2000. (*However, again please notice that the area of cultivated pasture that is lost in this way is still many times smaller than the yearly increase in cultivated pasture).

3 Land Use Change Models (in general)

This chapter will address the research questions related to land use change models in general. These are:

- Which types of land use change models exist and which type could be used?
- How does this type of model work (general description)?
- Which “hot” topics regarding this type of model could be included in a Pantanal Land Use model?
- How have other studies implemented this type of model?

3.1 Different types of land use change models

This paragraph will address the following research questions:

- Which types of land use change models exist & which type should we use?
- How does this type of model work (general description)?

This will be done by giving a general overview of different types of land use change models. The strengths and weaknesses of each type of land use change model will be discussed. Based on this discussion we will explain which type of model we consider most appropriate for modelling land use change in the Pantanal. Then, the basics of this type of model will be discussed in more detail.

3.1.1 Types of land use change models

There are several approaches to developing a spatially explicit model of land use change (Irwin and Geoghegan 2001; Veldkamp et al. 2001). A distinction can be made between spatially explicit economic models and spatially explicit non-economic models (Irwin and Geoghegan 2001). The difference between both types is that non-economic models just relate land use change to certain parameters, whereas economic models also analyze the human decisions behind land use change, which can explain why certain parameters are important (Irwin and Geoghegan 2001).

Spatially explicit non-economic models can be subdivided in three classes: simulation, estimation and a hybrid approach. Simulation models are mostly based on the cellular automata approach (Irwin and Geoghegan 2001). They mimic the spatial patterns of land use change through allocating a set of decision rules to a grid. The state of each cell of the grid is updated each time step according to decision rules, which use as input the state of neighbouring cells and the cell itself in the previous time step. A drawback of these models is that they *mimic* spatial patterns of land use change and do not mechanistically explain how these patterns have developed (Irwin and Geoghegan 2001). Typically, decision rules of simulation models are based on the user’s expert knowledge and not on statistical analyses (Verburg et al. 2002; Verburg[1] et al. 2004; Wu 2002). This lack of empirical input and quantitative understanding is often a point of criticism regarding these models (Verburg et al. 2002).

Estimation models use remotely sensed time series of data and try to relate (or fit) the observed land cover changes to explanatory variables that can be seen in the remotely sensed image. Like the simulation models above, estimation models *mimic* spatial patterns of land use change and do not mechanistically explain how these patterns have developed. Unlike simulation models, estimation models based the relations between the different explanatory variables and land use change on statistical analyses (Irwin and Geoghegan 2001).

The hybrid approach combines both approaches; it uses an estimation model to find out which variables are relevant for explaining land use changes, and then uses these variables in a simulation model that simulates land use change (Irwin and Geoghegan 2001). A good example of this hybrid approach is the CLUE-model (Irwin and Geoghegan 2001; Veldkamp et al. 2001; Verburg et al. 2002) or the LUCAS-model (Irwin and Geoghegan 2001; Hazen and Berry 1997).

3.1.2 Most appropriate model type

We propose to use a simulation model for modelling land use change in the Pantanal.
The reasons for this are:

1. An economic model requires economic data, which is only sparsely available.
2. The author is not an economist but a biologist, which makes the choice for a non-economic model more appropriate.

More specifically, we believe that a simulation model (i.e. **cellular automata or CA**) is most appropriate to model land use change in the Pantanal. The reason for this is:

- An estimation model is less adequate to incorporate neighbourhood effects into the model. This is an essential element in modelling habitat conversion in the Pantanal, as many factors influencing habitat conversion relate to neighbourhood effects (see paragraph 3.2).

3.1.3 Simulation models (CA)

As stated in paragraph 4.1.2, we propose to use a simulation model (CA) for modelling land use change in the Pantanal. So it is essential to know more about CA to use it efficiently in modelling land use change:

The building blocks of cellular automata are cells, which are located in a lattice or "grid" (Torrens 2000; Schatten 2004). Cells have different states and their state can change over time. This is done through transition rules. These rules describe the state of a cell in the next time step based on (a) the current state of the cell and (b) the current state of neighbouring cells. Transition rules are generally formulated as IF, THEN and ELSE statements and replace traditional mathematic function used in models with rule-based procedures (Torrens 2000). These rules apply to all cells of the grid and are executed simultaneously. Each execution of the rules constitutes a time step or "iteration".
The general formula, which describes the state of a cell in the CA, is therefore (Torrens 2000; Schatten 2004):

$$s_{it+1} = f(s_{it}, S^h_{it}) \qquad \textbf{[Equation 1]}$$

s_{it+1} = state of cell i at time t+1

f( ) = transition rule

s_{it} = state of cell i at time t

S^h_{it} = state of cells in neighborhood h of cell i at time t.

Transition rules (indicated as f () in equation (1)) can be divided in different classes (Schatten 2004):
Normal rules: state of cell i depends on the state of individual neighbouring cells.
Totalistic rules: state of cell i depends only on the **sum** of the states of neighbouring cells.

Legal rules: "subspecies" of totalistic rules defining that if a total zero-state is reached, no changes in state are possible any longer; i.e. all cells maintain their zero-state.
There are several classical CA neighbourhoods defined for a grid (Schatten 2004):

1. Von Neumann neighbourhood: cells directly above and below, left and right of the centre cell.
2. Moore neighbourhood: cells directly above and below, left and right of the centre cell are included AND cells diagonal to the centre cell.
3. Extended Moore neighbourhood: also includes layer of cells around Moore neighbourhood.
4. Margolus neighbourhood; mainly used to model diffusion and other processes that involve gas. We will therefore not consider this neighbourhood type.

However, researchers do not limit themselves to these definitions and often create their own neighbourhoods (Torrens 2000). This is especially useful in modelling spatial processes: most spatial processes have a global and a local component (Takeyama and Couclelis 1997; Wu 2004). Since "classical" CA only use local neighbourhood functions (see above), they are not ideal to model spatial processes with (Takeyama and Couclelis 1997; Wu 2002). Takeyama and Couclelis (1997) tried to provide a mathematical framework that could be used to describe both global and local functions. Such adaptations of classical CA's are often referred to as "relaxed CA's".

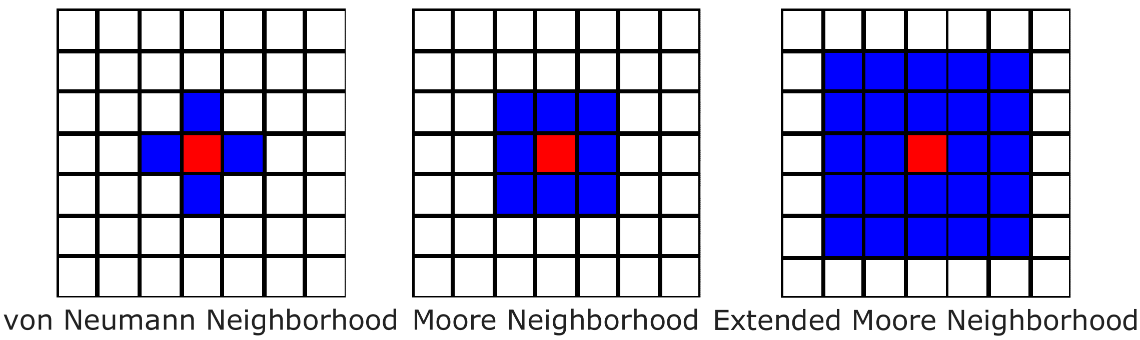


Figure 3.1.1. Several classical CA neighbourhoods. The red cell is the centre cell; the blue cells are the neighbourhood cells. The states of these cells are used to calculate the next state of the (red) centre cell according to the defined rule (from Schatten 2004).

3.2 "Hot topics" in CA

In the former paragraph we explained that we will use CA to model land use change in the Pantanal. Before starting with the actual modelling (chapter 5), we will try to answer some general questions regarding the use of CA in modelling land use change. This will be done in this paragraph and the next one.
This paragraph will address:

- Which "hot" topics" are there regarding CA & should we include any in our model?

By consulting literature we have found a number of "hot" topics" regarding CA:

  - Multi-Agent Systems (MAS)
  - Stochastically constrained CA
  - Integration of neural networks and CA

First, a general description and explanation of each of these issues will be given. The benefits and drawbacks of using one of these approaches in a land use change model will also be discussed.
Secondly, we will discuss if it is useful to incorporate any of these 'hot topics' in our land use change model.

3.2.1 Multi-Agent Systems (MAS)

Multi-Agent Systems model the behaviour of "agents" (hence the name Multi-Agent Systems). These agents have as characteristics that they are able to observe their environment or other agents and are able to respond to the observed state of their environment or other agents based on certain objectives or "goals" they have. The agents can be humans, computer programs, etc. (Balmann 2004).

In Multi-Agent Systems that model land use change, the agents are normally humans. With Multi-Agent Systems it is possible to model human decisions and actions at the individual level. Therefore, Multi-Agent Systems mimic the decision-making process of individual humans and combine (or aggregate) these decisions to predict how this will affect land use models (Ligtenberg et al. 2001; Deadman 1999).
Since human behaviour is one of the main factors shaping land use, it is considered very useful to integrate human behaviour in land use change models (Ligtenberg et al. 2001; Deadman 1999). Moreover, land use changes are often made at the level of individual humans. Multi-Agent Systems are designed to model behaviour at the individual level and are thus an appropriate tool to integrate human behaviour in land use change models.
Multi-Agent Systems can easily be combined with CA. CA is for instance used to enable agents to gather spatial information about their environment. An example of this is found in Ligtenberg et al. 2001.

3.2.2 Stochastically constrained CA

Basically, a stochastically constrained CA can best be described as a CA that has a stochastic component (pers. comm. A. Ligtenberg). The main reason for incorporating a stochastic component in a land use change model is that land use changes often seem to occur with a degree of randomness (de Nijs et al. 2004; White and Engelen 1993). Therefore, land use changes that are not completely understood and that show a high degree of randomness, are better modelled if a stochastic component is also included; i.e. the stochastic component can "account" for the observed randomness.
A stochastic component can be incorporated into CA in many ways. We will give some examples of case studies that have used different ways to incorporate a stochastic component in their CA.

The DINAMICA-model is a CA that is designed to model deforestation in the Amazon. In addition, it includes the processes of reforestation and clearance of reforested areas. Each of these processes is modelled in two stages. First, a part of the expected change is realized by the expansion or contraction of existing large areas (clusters) of one of these land use types. Then, the rest of the expected change is realized by creating new, isolated patches through a seedling mechanism. Both processes have a stochastic component. It selects the cells with the highest suitability for the desired change (based on a suitability analysis made beforehand), puts these cells organized in a data array and then selects randomly from this data array. The land use type of the selected cells is then changed (Soares-Filho et al. 2002).

The Environment Explorer-model (Leefomgevingsverkenner) is a CA that is used to model land use changes in the Netherlands (de Nijs et al. 2001; de Nijs et al. 2004). The model distinguishes 5 main land use types or "functions" which are each subdivided into smaller groups: agricultural function (other, pasture, cropland, and greenhouses), residential function (high-density, and low-density), economical function (industrial, commercial services, public services, recreation), natural function (forest, meadows, and nature) and other (fresh water, salt water, airports, borders). The model first calculates the expected changes in each land use type for the entire Netherlands based on demographical and economical data (trend lines). The model then selects the individual grid cells that are most suitable for change by means of a “transition potential”. According to transition rules, grid cells with the highest “transition potential” change. The “transition potential” contains a stochastic component, and is calculated for a cell i as follows:

$$\mathbf{TP_i = V * F1_i * F2_i * F3_i * F4_i} \qquad \textbf{[Equation 2]}$$

TP = transition potential

F1, F2, F3 and F4 = driving factors

V = stochastic variable

We are interested in variable V because this is the stochastic component of the model. Variable V is calculated by $V = 1 + (-\ln [rand])^a$, where $0 < rand < 1$ is a random function. Thus, the Environment Explorer introduces a stochastic component in the model by simply adding a stochastic function to the suitability calculation.

3.2.3 Integration of neural networks and CA

Land use changes are complex and involve many factors that influence land use not directly but in a complex interaction (Verburg et al. 2002). This results in non-linear relationships that are difficult to model (Mas et al. 2004). A possible approach to deal with this is to use neural networks (Mas et al. 2004); i.e. parameterization of a model is often difficult and when the modelled processes are very complex, it is useful to automate the parameterization through a neural network (Wu 2002).

Neural networks are computer programs that can “learn” (Mas et al. 2004). The learning process begins by giving them a number of maps with potential driving factors and land use maps. The network will try to relate the driving factors to the observed land use changes by giving weights to each of the driving factors. If the wrong weights are given to the driving factors, the output created by the network will not match the authentic land use maps. The network will then change the weights given to the driving factors and see if this makes a better match. This process will continue and the network will remember the best solution. Once the network has learnt from the example, it can be given new real input and process it according to what it learnt from the examples and thus predict future land use changes (Mas et al. 2004).

Neural networks and CA can be easily combined; the input is commonly in raster format and when the network finds a neighbourhood influence on land use change, it already has a CA component.

3.2.4 ‘Hot topics’ and our land use change model

Above we discussed three ‘hot topics’. These were Multi-Agent Systems, stochastically constrained CA and neural networks. However, would it be useful to incorporate one of these three ‘hot topics’ in our own model? We will discuss this now for each topic.

As stated, Multi-Agents Systems mimic the behaviour of relevant “actors” in the land use change process (Veldkamp 1999). In our study, ranchers are the most relevant actors; they own most of the land and they decide whether savannah will be cultivated. This was already implicitly considered in choosing the driving factors: driving factors were determined based on the notion that a rancher would consider them relevant if converting an area into cultivated pasture. Unconsciously, the decision of a rancher was mimicked and in this sense our land use change model had already taken a first step towards a Multi-Agent System. Our land use change model would become a full Multi-Agent System if the model would focus entirely on the decision making process of ranchers and would incorporate data about individual farmers that influences the decision making of a rancher (e.g., wealth of a rancher). Thus, our conclusion is that it is useful to combine our CA-model with a Multi-Agents System. However, the practical implementation of this requires (1) advanced programming skills that the author does not posses and (2) high-quality data about individual ranchers that is not available. Although it would be useful to combine our CA with a Multi-Agents System, the practical implementation of this is not possible at this point.

As stated, a stochastically constrained CA is a CA with a stochastic component. The main reason for incorporating a stochastic component in a land use change model is that land use changes often occur with a degree of randomness. Especially land use changes that are not completely understood and that show a high degree of randomness are better modelled if a stochastic component is also included; i.e. the stochastic component can "account" for the observed randomness. However, a stochastically constrained CA might not be appropriate for the Pantanal-model. We believe that the observed land use changes are understandable to a certain degree. Therefore, it seems more appropriate to model the observed land use changes without a stochastic component. However, it could be that the model turns out to fit the observed land use changes poorly (this will become clear during calibration). Then, it is clear that the land use changes were not well understood, and it might be useful to add a stochastic component after all.

Neural networks have as main advantage that they can deal more easily with non-linear relationships than other approaches which is an advantage when modelling land use changes (Mas et al. 2004). However, they also have a number of important drawbacks, which make them unsuitable for our land use change model:

- The training of a neural network requires a lot of data; a network learns by trial-and-error and this method naturally only works if enough data is available (Malczewski 2004). Unfortunately, such an amount of data is not available for our study.
- The development of a neural network requires extensive knowledge about programming (Malczewski 2004). Unfortunately, we do not have such knowledge.
- Neural networks are not transparent. They develop their own algorithms to model land use changes and are thus not user-defined; i.e. they resemble a "black box". This can be an important drawback as one of the main reasons for developing a land use change model is to make the consequences of certain land use trends more transparent. If a model is not transparent as is the case for neural networks, the usefulness of the model is affected (Malczewski 2004).

Thus, our conclusion is that it is not useful to integrate our land use change model with a neural network.

3.3 Comparison of land use change models

In the former paragraph, we explained that we will try to answer some general questions regarding the use of CA in modelling land use change. This paragraph will address:

- How did other studies implement CA?

To answer this, a number of land use change models will be reviewed: CLUE-S, Environment Explorer and DINAMICA.

In this paragraph, we will first explain why these models have been selected; i.e. discuss the selection criteria.

Then, we will elaborate for each model on:

- General structure
- Allocation model
- Order of processing by allocation model
- Land use changes modelled by allocation model
- Applications
- Implementation of allocation model:
  - Data structure
  - Cell size
  - Cell states
  - Neighbourhood
  - Driving factors
  - Weights of the driving factors
  - Time steps

Afterwards, some conclusions will be drawn.

3.3.1 Selection criteria land use change models

A criterion is that the land use change models should use CA. This is, obviously, because CA will be used in our land use change model (see 4.1.2) and we want to learn by reviewing comparable studies.
Another criterion is that the land use change models should not focus explicitly on urban growth; this study aims to model the change in cultivated pasture in a (relatively) natural area and a strict urban growth model was not thought to be an appropriate comparison. Instead, models were chosen that model land use changes such as deforestation, agricultural intensification, etc. However, models that include urban growth as well as deforestation, agricultural intensification, etc. were considered appropriate comparisons.
Another criterion is that the models should be relatively well known. Both the CLUE-S model and the Environment Explorer (see below) are well known in the Netherlands. Although the DINAMICA-model is not as well known as the other models, this model is included as well. This is because it models deforestation in the Amazon and thus has similarities with our study regarding (a) the land use change modelled (both deforestation and cultivated pasture can be seen as conversion of natural habitat) and (b) the location, namely Brazil.

3.3.2 Land use change model 1: CLUE-S

General structure:
The CLUE-S model (Conversion of Land Use and Effects at Small regional scale) is used to model land use changes at a small regional scale (Verburg et al. 2002). The CLUE-S model has of two phases: the demand-module and the allocation-module.
The demand-module calculates the total amount of change for all land use types for the entire region ("quantity of change"). This is calculated in a non-spatial manner. It is used as input for the actual allocation model.
The allocation-module determines how this expected amount of change should be divided over the region ("location of change"). The allocation-module is in fact an allocation model, and as such the actual spatially explicit land use change model. It will therefore be described in more detail:

Allocation model:
Each cell will change into the land use type for which it has the highest total probability ($TPROP_{i,k}$). The total probability of a land use type k is calculated for a cell i as follows:

$$\mathbf{TPROP_{i,k} = P_{i,k} + ELAS_k + ITER_k} \qquad \textbf{[Equation 3]}$$

$ELAS_k$ is a measure for the easiness with which a land use type can be converted into another land use type; for some land use types it is not likely that they will change over night (e.g. because of enormous investments) and this is reflected in a high $ELAS_k$. $ITER_k$ is a random parameter that is unique for each land use type. It is used to match the allocated amount of change to the demand during the allocation procedure (see below) (Verburg et al. 2002).
Some cells are not allowed to change; for instance cells with land use type water can not change (Verburg et al. 2002).
The allocation procedure starts by a first allocation or iteration during which $ITER_k$ is the same for all land use types (Verburg et al. 2002). After this first allocation or iteration, the model checks if the allocated amount of cells matches the demand. If so, the model continues with the next time step. However, if for some land use types the allocated amount of cells is smaller than the demand, the model will start a new iteration. However, the model will now increase the $ITER_k$ for these land use types. The model will continue with these iterations until the amount of allocated cells matches the demand. Then the model will continue with the next time step (Verburg et al. 2002).

Order of processing:
The model has the following order of processing (Verburg et al. 2002):

1. Calculation of demand.
2. Allocation:
    a. Suitability analysis.
    b. Allocation based on suitability analysis.
    c. Comparison of allocated amount of cells with demand; if they do not match, the allocation is iterated.
    d. After completing the former three steps, the model continues with the next time step.

Land use changes modelled by allocation model:
A range of land use changes can be modelled with CLUE-S, including deforestation, urbanization, and agricultural intensification. Competition between different land use types can even be modelled with CLUE-S (Verburg et al. 2002).

Applications:
The CLUE-S model has already been used, e.g. in the Philippines and Malaysia. During the discussion of the practical implementation of the model we will focus on the application in the Philippines; most data is available for this application. In this application the land use changes around a protected forest reserve are modelled (Verburg et al. 2002).

Implementation of allocation model:
<u>Data structure:</u>
Raster (Verburg et al. 2002).
<u>Cell size:</u>
Flexible; depends on the application for which the model is used. However, the CLUE-S model is only suitable for modelling at the small regional scale and can only deal with resolutions finer than 1km x 1km (Verburg et al. 2002). In the application of the model in the Philippines a resolution of 100mX100m was used (Veldkamp and Verburg 2004).
<u>Cell states:</u>
The CLUE-S model is quite flexible, and the cell states depend on the application for which the model is used. In the application of the model in the Philippines 5 land use types were

distinguished: forest, grassland, coconut plantation, rice field and others (e.g. mangrove and built-up areas) (Verburg et al. 2002; Soepboer 2001).

Driving factors:

As stated, the demand-module calculates the total amount of change for each land use type for the entire region ("quantity of change"). These calculations depend strongly on the scenarios that are evaluated and can vary between simple extrapolations and complex economical models (Verburg et al. 2002). Consequently, the driving factors that are considered in the demand module also vary considerably; for instance, in the application of the model in the Philippines no driving factors were considered at all. Instead, a simple linear extrapolation of historical land use trends was used to calculate the demand.

The allocation-module used the following driving factors in the application of the model in the Philippines: altitude, slope, distance to town, distance to stream, distance to road, distance to coast, distance to port, erosion vulnerability, geology and population density (neighbourhood 5X5).

Weights of the driving factors:

As stated above, the weights of the driving factors considered in the demand-module depend strongly on the scenarios being used. In the application of the model in the Philippines a simple linear extrapolation of historical land use trends was performed and no driving factors were considered.

Weights of the driving factors in the allocation-module of CLUE-S are always derived with logistical regression. In this procedure, possible driving factors are statistically related to observed land use changes (Verburg et al. 2002). The solution that explains best the observed changes is chosen. Driving factors that only have a marginal influence will be left out. The relative contributions of the individual driving factors can then be used as weights that determine the probability that a location (or cell) i will change into a land use type k:

$$\mathbf{Log\ (P_i / (1-P_i)) = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots.. \beta n X_{n,i}} \qquad \textbf{[Equation 4]}$$

P_i = probability that a cell i will change into land use type k
X_1, X_2, X_n = driving factors
β_1, β_2, β_n = weights of the driving factors

Neighbourhood function:

In the application of the model in the Philippines a 5x5 focal window was used to determine the effect of nearby population density (Verburg et al. 2002). However, in other applications different functions are used (reference).

Time period:

The period 1997-2012 is modelled. The model simulates land use change over a period of 15 years with time steps of 1 year (Soepboer 2001).

3.3.3 Land use change model 2: Environment Explorer

General structure:

The Environment Explorer has two phases that both use a model: the "macro"- and the "micro"-model (de Nijs et al. 2001).

In the first phase, the "macro-model" calculates the expected changes in each land use type for the entire Netherlands based on demographical and economical data (trend lines). This is calculated in a non-spatial manner. It is used as input for the actual allocation model.

In the second phase, the "micro-model" allocates the expected changes calculated by the macro-model to individual grid cells. The "micro-model" is in fact an allocation model, and as such the actual spatially explicit land use change model. It will therefore be described in more detail:

Allocation model:

The "micro-model" allocates the calculated expected change (increase or decrease) of each land use type. In the allocation model, each cell will change into the land use type for which it has the highest transition potential (P) (de Nijs et al. 2001). The transition potential depends on a number of factors. These are integrated into the transition potential as follows:

$$\mathbf{P_{i,k} = V * (S_{i,k})^{\sigma 1} * (A_{i,k})^{\sigma 2} * (Z_{,k})^{\sigma 3} * (N_{i,k})^{\sigma 4}} \qquad \textbf{[Equation 5]}$$

$S_{i,k}$ = suitability of a cell for each land use type

$A_{i,k}$ = accessibility of a cell for each land use type

$Z_{,k}$ = if a cell is included in one of the planning policies

$N_{i,k}$ = influence of neighbouring cells on cell i.

$V = 1 + (-\ln[\text{rand}])^a$ │ 0 < random function < 1

k = certain land use type

i = certain cell

σ1, σ2, σ3, σ4 = parameters determining if suitability, accessibility, planning policies and neighbourhood are incorporated (value=1) or not (value=0) in the transition potential

Then the allocation model allocates iteratively during each time step to each cell the land use type for which it has the highest transition potential (P) (de Nijs et al. 2001; Groothuysen et al. 2001). However, this allocation will not take place if there are already enough cells containing this land use type available in the nearby region. After iteration the model will check how many cells have been assigned each land use type. The model will compare this with the expected change or demand for each land use type. If this does not match, the model will start a new iteration and select the cells with high transition potentials that were not selected in the first iteration. The iterations will continue until the demand has been matched (Groothuysen et al. 2001).

However, cells with land use type's fresh water, salt water, airport and border will always have these land use types (de Nijs et al. 2001).

Order of processing:

The model has the following order of processing (de Nijs et al. 2001):

1. Calculation of demand.
2. Allocation:
   a. Suitability analysis.
   b. Allocation based on suitability analysis.
   c. Comparison of allocated amount of cells with demand; if they do not match, the allocation is iterated.
   d. After completing the former three steps, the model continues with the next time step.

Land use changes modelled by allocation model:

The following land use types are dynamically modelled: pasture, cropland, greenhouses, other agriculture, high-density residential, low-density residential, industrial, commercial services, public services, recreational, forest, meadows, and other nature. However, land use types pasture, cropland and other agriculture are not expected to increase and growth is not modelled for these land use types. Consequently, these land use types are expected to decrease at the expense of other land use types.

A number of land use types are not dynamically modelled and their area remains static: fresh water, salt water, airports and borders (de Nijs et al. 2001). However, their presence influences the allocation of the dynamically modelled land use types.

Applications:
The only application thus far is in the Netherlands; this is logical, as the model has specifically been designed for this. It is currently used to support Nature Outlook 2, in which the RIVM reports on future trends concerning environment and nature (de Nijs et al. 2001).

Implementation of allocation model:
<u>Data structure:</u>
Raster.
<u>Cell size:</u>
The Environment Explorer models land use change at the national scale and consequently uses a fixed resolution of 500mX500m. However, the resolution will change to 100mX100m in the coming years (pers. comm. Hilferink).
<u>Cell states:</u>
The model distinguishes 5 main land use types which are each subdivided into smaller groups: agricultural function (other, pasture, cropland, and greenhouses), residential function (high-density, and low-density), economical function (industrial, commercial services, public services, recreation), natural function (forest, meadows, and nature) and other (fresh water, salt water, airports, borders) (de Nijs et al. 2001).
This last land use type "other" is not dynamically modelled and consequently cells with this land use type will not change (see also "Land use changes modelled").
<u>Driving factors:</u>
The macro-model calculates the yearly increase or decrease in each land use type ("demand") based on economical and demographical data. These calculations are mainly economical or demographical by nature and it would go too far to discuss them in detail in this study. We therefore limit ourselves to giving an overview of the driving factors used in the macro-model in table I.
The micro-model uses the following driving factors (Groothuysen et al. 2001): current land use type, presence of railway station, driveway of highway, presence of airport (Schiphol), noise (traffic/airplanes), presence of harbour (Rotterdam), high sandy grounds, seepage, hydrological restoration, and a number of driving factors relating to greenhouses (labour costs, ground prices, precipitation, available sunlight, average wind speed, temperature). An overview is given in table I.

Table V. Effects of different driving factors on the suitability of a location (i.e. grid cell) for land use types dynamically modelled by the Environment Explorer. Minus sign (-) indicates a negative relation and a plus sign (+) indicates a positive relation. The number of plus or minus signs indicates the strength of the relation (based on Groothuysen et al. 2001).

Land use type	Current land use	Railway station	Driveway	Airplane noise	Traffic noise	Harbour	Airport	High grounds	Kwelwater	Hydrol. retoration	Greenhou-se factors
Residential	++++	+	++	- - -	-						
Industrial	++++		+	- -	+	+	+				
Commercial	++++	+++	+	- - -							
Public services	++++		+	- - -	+						
Forest	++++							+++			
Meadow	++++								+++	++	
Nature	++								++	++	
Greenhouses	++										++++
Recreation	++++			- -	- - - -						

In addition, for each location (i.e. grid cell) is checked if it is included in one of the planning policies; e.g. in some locations is building prohibited according to national planning policy (Groothuysen et al. 2001). Accessibility is also considered for each location (i.e. grid cell); a cell has to have a good connection to the road network to be suitable for certain land use types (de Nijs et al. 2001).

<u>Weights of the driving factors:</u>

The weights of the driving factors were determined based on a calibration with data from the period 1989-1993 (de Nijs et al. 2001).

<u>Neighbourhood function:</u>

The neighbourhood in the Environment Explorer consists of a concentric circle with a radius of 8 cells. The entire neighbourhood thus consists of 196 cells (de Nijs et al. 2001).

3.3.4 Land use change model 3: DINAMICA

General structure:

The DINAMICA-model is a CA that is primarily designed to model deforestation in the Amazon. The model has two phases. In the first phase, the yearly increase or decrease in each land use type ("expected change") is calculated in a non-spatial manner. It is used as input for the actual allocation model (Soares-Filho et al. 2002).

In the second phase, the calculated expected change in each land use type is allocated over the region. An allocation model is used for this; i.e. this is the actual spatially explicit land use change model. It will therefore be described in more detail:

Allocation model:

The allocation model consists of two components. In the first component, part of the expected change in a land use type is realized by the expansion or contraction of existing large areas (clusters). In the second component the rest of the expected change is realized by creating new, isolated patches through a seedling mechanism.

In both stages the model makes a probability analysis before each time step, in which it assesses the probability that a cell will change into a land use type. This probability is determined by the logistic regression already used in calculating the weights of the driving factors, but now the relative contributions of the individual driving factors to the best fit are used as weights that determine the probability that a location (or cell) will change into a land use type. Thus, formula 5 is used to calculate the probability that a cell i will change into land use type k. Cells with the highest probability are selected. The land use type of the selected cells is then changed (Soares-Filho et al. 2002).

Order of processing:
The model has the following order of processing (Soares-Filho et al. 2002):

1. Calculation of demand.
2. Allocation:
    a. Suitability analysis.
    b. Allocation based on suitability analysis.
    c. Comparison of allocated amount of cells with demand; if they do not match, the allocation is iterated.
    d. After completing the former three steps, the model continues with the next time step.

Land use changes modelled by allocation model:
The model can be used to simulate a variety of land use changes (including urbanization; see Almeida et al. 2002) but its primary focus is on deforestation (Soares-Filho 2002).

Applications:
The model has been used to model the processes described above in two Amazon regions: Guaranta and Terra Nova. In addition, the model has been used in a variety of applications, including the modelling of urbanization in a medium-sized town in the province of Sao Paolo in Brazil (Almeida et al. 2002) and the modelling of land use change along a major Amazon highway (Soares-Filho et al. 2004). However, the discussion of the practical implementation of the model we will focus on first two applications (Guaranta and Terra Nova) in the Amazon; most data is available for these. In these applications, the processes of deforestation, spontaneous re-growth on abandoned land and deforestation of re-growth are modelled for two areas in the Amazon.

Implementation of allocation model:

<u>Data structure:</u>
Raster (Soares-Filho et al. 2002).

<u>Cell size:</u>
Flexible. In the application that was discussed a resolution of 100mX100m was used (Soares-Filho et al. 2002).

<u>Cell states:</u>
DINAMICA actively models four cell states: deforestation, re-growth, secondary forest and forest. However, the DINAMICA accepts a range of cell states as input: dense rain forest, open rain forest, alluvial forest, dense savannah, open savannah, pastures in diverse stages of use and agricultural areas, young and intermediate successions. These cell states are either classified as deforestation, re-growth, secondary forest and forest, or else they are not actively modelled (Soares-Filho et al. 2002).

<u>Driving factors:</u>
As stated, the model first calculates the yearly increase or decrease in each land use type ("expected change") in a non-spatial manner. These calculations depend strongly on the scenarios that are evaluated and can vary between simple calculations and complex models (Soares Filho 2002). Consequently, the driving factors that are considered in these calculations also vary considerably. In the application of the model in Guaranta and Terra

Nova (Amazon), no driving factors were considered at all. Instead, it was assumed that the deforestation rate followed an asymptotic curve.

The allocation-module used the following driving factors in the application of the model in the Amazon: vegetation type, soil type, altitude, slope, urban attraction, distance to respectively main roads, secondary roads, river, deforestation, re-growth, and forest.

Weights of the driving factors:

As stated above, the weights of the driving factors considered in calculating the yearly increase or decrease in each land use type depend strongly on the scenarios being used. In the application of the model in the Amazon no driving factors were considered at all. Instead, it was assumed that the deforestation rate followed an asymptotic curve.

Weights of the driving factors in the allocation-module are always derived with logistical regression. As already explained during the discussion of the CLUE-S model are possible driving factors statistically related to observed land use changes in this procedure. The solution that explains best the observed changes is chosen. Driving factors that only have a marginal influence will be left out. The relative contributions of the individual driving factors will later be used as weights that determine the probability that a location (or cell) i will change into a land use type k (see "Transition rules") (Soares-Filho et al. 2002).

Neighbourhood function:

A 3 x 3 window for the Expander function and a 3 x 2 window for the Patcher function (Soares-Filho et al. 2002).

Time period:

A time period of 8 years was modeled in two phases: 1986-1991 and 1991- 1994 (Soares-Filho et al. 2002). The reason is that different processes played a role during each phase and thus different parameters had to be used for each phase: during the first phase the gold rush up north attracted many colonists and they abandoned their lands. During the second phase the gold rush declined and this resulted in a higher pressure on land. The model simulates land use change in time steps of 1 year (Soares-Filho et al. 2002).

3.3.5 Conclusions

Above three land use change models were reviewed. The following issues were discussed:

- General structure
- Allocation model
- Implementation of allocation model:
  - Data structure
  - Cell size
  - Cell states
  - Driving factors
  - Weights of the driving factors
  - Neighbourhood function
  - Time period

An overview of the results of this review is given in table II (see below). Some conclusions can be drawn from these results.

General structure:

All three models have two separate phases. In the first phase the increase or decrease in each land use type is calculated: i.e. the total number of cells that need to change during each time step. This phase is non-spatial and instead uses economical and demographical models, or simply extrapolates historical land use trends.

The second phase allocates the calculated changes in each land use type over the grid. This phase is spatial and uses CA, and is the actual spatially explicit land use change model; it is therefore discussed separately below ("Allocation model").
CLUE-S and DINAMICA put most emphasis on the second phase. They state clearly that this component depends strongly on the scenarios that are evaluated and that it can vary accordingly between simple calculations and complex models (Verburg et al. 2002; Soares Filho 2002). In the Environment Explorer there is a more elaborate framework for the first phase than in the other models.

Allocation model:
All three models have a comparable allocation model. Although the allocation-model of the DINAMICA-model is more complex than the CLUE-S and the Environment Explorer, it still has this structure as its basis:
In all three models a suitability analysis is carried out before the actual allocation. This suitability analysis selects cells that will be used in the allocation procedure. A first allocation is then made (based on the suitability analysis) and a number of the selected cells change. After this first allocation, the allocated amount of cells is compared with the expected amount of change or "demand" (calculated in the first component; see above). If this does not match, a new allocation or "iteration" is made with slightly adapted parameters. The allocated amount of cells is again compared with the demand. If they do not match, another iteration is made. This iterative process is repeated until the allocated amount of cells matches the demand.

Order of processing:
Strikingly, all three models have the same order of processing

1. Calculation of demand.
2. Allocation:
    a. Suitability analysis.
    b. Allocation based on suitability analysis.
    c. Comparison of allocated amount of cells with demand; if they do not match, the allocation is iterated.
    d. After completing the former three steps, the model continues with the next time step.
    - This order of processing is very common among land use change models. It can be considered the standard approach to the spatially explicit modelling of land use changes.

Implementation of allocation model:
<u>Data structure:</u>
The models were selected because they used CA and consequently have used a raster format.

<u>Cell size:</u>
Both Clue-S and DINAMICA use a cell size of 100m X 100m. The Environment Explorer uses a cell size of 500m X 500m (de Nijs et al. 2001). The reason for this is that when the first versions of the Environment Explorer were developed most computers could not handle a spatial resolution of 100m X 100m and computations became too slow (pers. comm. van Eupen). Of course, this has changed and it is planned that the Environment Explorer will switch to a cell size of 100m X 100m (pers. comm. Hilferink). Based on the studies we reviewed, it seems that a cell size of 100m X 100m is more or less the standard for studying land use changes at the local or regional level.

Both Clue-S and DINAMICA have flexible cell sizes; i.e. it is possible to model with cell sizes other than 100m X 100m. Environment Explorer has a fixed cell size; it is only possible to model with a cell size of 500m X 500m. However, this has turned out to be a problem: recently, it was decided that a resolution of 500m X 500m is too coarse resolution to adequately use the model. The resolution of the model will change to 100m X 100m. Since the model has a fixed cell size, this change will require considerable resources (pers. comm. van Lammeren).

Cell states:

Environment Explorer used 13 cell stated, whereas the other two models used much less cell states. This is logical as the Environment Explorer models many different land use processes. CLUE-S depends on how many land use changes are included; i.e. the more (land use change) processes are modelled, the more complex a model will be and the more cell states must be included.

Driving factors:

All three models include about 11 driving factors (although Environment Explorer slightly more). This is striking, as these models model different land use changes in completely different areas. However, there is a reason why these models do not include many more driving factors; by increasing the number of driving factors (and thus parameters) in a model, the uncertainty of a model also increases because no parameter is error free (Jørgensen 1986). Therefore, the total uncertainty in a model accumulates and becomes larger by increasing the number of parameters. This means that a model does not automatically get better by making it more complex. However, a too simple model does also not capture the complexity of real-world processes. Therefore, it is believed that a trade off exists between complexity and simplicity (Jørgensen 1986), and that good modelling means that not too few parameters are included and not too many parameters are included. This is clearly reflected in the models that were reviewed; these have more or less the same number of driving factors.

Weights of the driving factors:

Both Clue-S and DINAMICA use logistic regression to determine the weights of the driving factors. Environment Explorer uses calibration to determine the weights of the driving factors. However, both approaches use an empirical method to determine the weights of the driving factors, and thus the relation between driving factors and land use change. This is very common in land use change modelling. Let us now explain why.

According to Verburg[1] et al. (2004), three approaches exist to determine (quantitatively) the relation between driving factors and land use change. The first approach bases these relations on a good understanding of the process being studies; theories and physical laws are used to quantify these relations (Verburg[1] et al. 2004). However, this is an uncommon approach in modelling land use change because of the complexity of land use change processes and the absence of theories and physical laws regarding land use change. The second approach is to use expert knowledge to determine (quantitatively) the relation between driving factors and land use change. However, this lack of empirical input and quantitative understanding is often a point of criticism regarding these models (Verburg[1] et al. 2004). The third method uses empirical methods to determine (quantitatively) the relation between driving factors and land use change (Verburg[1] et al. 2004). Examples include logistic regression and calibration as used by the three reviewed models.

Because of these reasons, this third approach is commonly used in land use change modelling to determine the relation between driving factors and land use change. This is clearly reflected in the reviewed models, which all use an empirical method.

Neighbourhood function:

All three models use a different neighbourhood function:

- Clue-S: 5x5 focal window
- Environment Explorer: concentric circle with radius of 8 cells.
- DINAMICA: 3 x 3 window (Expander) and a 3 x 2 window (Patcher)

Clearly, these three models have used adaptations of the "classical" CA-neighbourhood. This is logical, as "classical" CA uses strict neighbourhoods (see paragraph 3.1), which are often not ideal to model spatial processes with. Therefore, many researchers do not limit themselves to the strict neighbourhood functions of "classical" CA and create their own neighbourhoods (Torrens 2000). Such adaptations of classical CA's are often referred to as "relaxed CA's" (see paragraph 3.1).

Time period:

The time period over which the three models model differs; both CLUE-S and DINAMICA model land use changes over a relatively short period of time (respectively 15 and 8 years) whereas the Environment Explorer is used to model land use changes over a longer period (40 years). The reason for this might be that CLUE-S and DINAMICA both extrapolate existing land use trends and do not actually address the underlying mechanisms of land use change; i.e. they assume that the relative influence of the driving factors will stay the same. The Environment Explorer addresses more underlying causes and thus makes fewer assumptions which allow it to model over a longer period of time.

Since our study also uses extrapolation to predict future land use changes, it is clear that we should not model over a too long period of time.

Table VI. Overview of the characteristics of the models reviewed in 4.3.2, 4.3.3, and 4.3.4. For more information; see these paragraphs.

	CLUE-S	**Environment Explorer**	**DINAMICA**
General structure	Model has 2 phases: demand- and allocation-phase.	Model has 2 phases: demand- and allocation-phase.	Model has 2 phases: demand- and allocation-phase.
Allocation model	Following steps: Demand Suitability Iterative allocation	Following steps: Demand Suitability Iterative allocation	Following steps: Demand Suitability Iterative allocation
Order of processing	1. Calculation of demand. 2. Allocation: -Suitability analysis. -Allocation -Comparison -Iteration (if necessary)	1. Calculation of demand. 2. Allocation: -Suitability analysis. -Allocation -Comparison -Iteration (if necessary)	1. Calculation of demand. 2. Allocation: -Suitability analysis. -Allocation -Comparison -Iteration (if necessary)
Land use changes modelled by allocation model	Variety: competition between land use types can also be modelled.	Variety: focus is on urbanization (residential and commercial) and natural habitat.	Variety including urbanization, but its primary focus is on deforestation.
Applications	Philippines, Malaysia, etc.	Netherlands	Amazon and other areas of Brazil
Data structure	Raster	Raster	Raster
Cell size	Flexible	Fixed: 500m X 500m	100m X 100m
Cell states	5	13	3
Driving factors*	11	More than 11.	10
Weights of driving factors	Logistic regression	Calibration	Logistic regression
Neighbourhood function	A 5x5 focal window.	Concentric circle with radius of 8 cells.	A 3 x 3 window and a 3 x 2 window.
Time period	15 years	40 years	8 years
Other remarks			Allocation module consists of 2 components: expander- and patcher-function.

* Only the driving factors in the allocation-module are given; see the individual paragraphs for more information on the driving factors used in the demand-module.

4 Building a Land Use Change Model

A number of land use change models have been reviewed in paragraph 4.3. All these models had two different phases. In the first phase, the increase or decrease in each land use type is calculated. This phase is non-spatial and instead uses economical and demographical models, or simply extrapolates historical land use trends. It is used as input for the second phase.
The second phase allocates the calculated changes in each land use type over the grid. This phase is spatial and uses CA, and is therefore the actual spatially explicit land use change model.
In this study, the expected change in cultivated pasture was already calculated in paragraph 3.1. This calculation of the expected change can be regarded as the first phase of our study. This chapter will deal with the second phase; i.e. actually developing a spatially explicit land use change allocation model.
This chapter will therefore centre on the following research question:

- What concept will I use to model land use change in the Pantanal?

This chapter will address this research question in two steps. First, it will discuss the conceptual model that was used to model land use change in the Pantanal. Secondly, it will discuss the formalization of the conceptual model.

4.1 What concept will I use to model land use change model in the Pantanal?

In paragraph 4.3, a number of case studies have been reviewed. Some issues concerning the practical implementation of CA in a land use change model have been discussed (e.g. structure, cell size, cell states, etc.). Based on the information from these case studies, we will now discuss these issues for our land use change model.
The following issues will be discussed:

- Land use changes modelled
- General overview of the model
- Data structure
- Cell size
- Cell states
- Driving factors of change
- Weights of the driving factors
- Neighbourhood function
- Time period
- Input

Land use changes modelled
As stated in paragraph 2.2, three processes are assumed to play a role in the Pantanal:

- The area of cultivated pasture increases at the edge of existing areas of cultivated pasture; i.e. existing areas of cultivated pasture have a tendency to expand.
- The area of cultivated pasture increases at new, isolated patches away from existing areas of cultivated pasture; i.e. isolated "islands" or clusters of cultivated pasture suddenly appear.

- A small number of cultivated pastures disappear. However, please notice that the area of cultivated pasture that is lost is many times smaller than the yearly increase in cultivated pasture.

All processes are included in the model.

General overview of the model

The general structure of the Pantanal-model is shown in figure 4.1.1. It has two distinct components: a dry season component, which models the first two processes and a wet season component, which simulates the third process. The model is divided into these two components because it was assumed that the first two processes (construction of cultivated pastures) would mainly occur during the favourable dry season, whereas the third process (disappearance of cultivated pastures) is more likely during the unfavourable wet season because of unexpected flooding. Therefore, it was necessary to design two separate components. Hereafter, the dry season component is referred to as the “dry season model” and the wet season component as the “wet season model”.

Since dry and wet season each last ½ year, the Pantanal-model models land use change in steps of ½ year. Modelling always begins in the dry season with the dry season model. The output of the dry season model is an ecotype map, which becomes the input of the wet season model. In return, the output of the wet season model is an ecotype map, which is the input of the dry season model during the next time step. However, in the starting year, the input of the dry season model is the ecotype map produced by the ecological model of the Pantanal-Taquari project (Jongman et al. 2005). In addition, the ecotype map produced by the wet season model could also be used as input for the ecological model of the Pantanal-Taquari project.

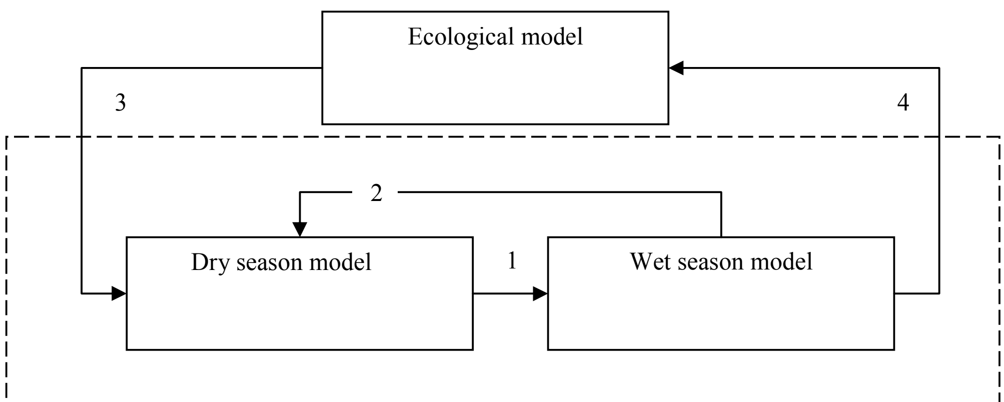


Figure 4.1.1. General structure of the Pantanal-model. The dashed line indicates the boundaries of the Pantanal model. The following actions can be discerned: (1) the dry season model produces an ecotype map, which is the input of the wet season model; (2) the wet season model produces an ecotype map, which is the input of the dry season model during the next time step. (3) In the starting year, the input of the Pantanal model is the ecotype map produced by the ecological model of the Pantanal-Taquari project. In addition, (4) the output of the wet season model could be used as input for the ecological model of the Pantanal-Taquari project.

<u>Dry season model:</u>

As stated, this component models two processes:

- “Expansion”. The area of cultivated pasture increases at the edge of existing areas of cultivated pasture; i.e. existing areas of cultivated pasture have a tendency to expand.

- "Patching". The area of cultivated pasture increases at new, isolated patches away from existing areas of cultivated pasture; i.e. isolated "islands" or clusters of cultivated pasture suddenly appear.

These processes are assumed to occur separately and simultaneously. To model them adequately, we have developed a separate function for each process. Therefore, this component has two functions, which work independently from each other but are executed simultaneously. This concept is similar to the DINAMICA model (see paragraph 4.3.4), which also was confronted with two separate processes regarding deforestation. In accordance with the DINAMICA-model, the two functions are called respectively "Expander" and "Patcher": "Expander" models the expansion of existing cultivated pasture, whereas "Patcher" models the appearance of isolated clusters of cultivated pasture. However, notice that whereas DINAMICA executes "Expander" and "Patcher" sequentially, whereas the Pantanal-model executes both functions simultaneously.
In addition, the dry season model models the increase in cultivated pasture (through the processes of "Expansion" and "Patching") separately for the eastern and western region of the study area; i.e., large differences exist in the increase in cultivated pasture between both regions and this makes it appropriate to model both regions separately. However, please notice that when modelling the increase in cultivated pasture for both regions, the model remains the same (same input, parameters, etc.) for both regions. The only parameter that is different is the expected increase in cultivated pasture or "demand". The expected increase in cultivated pasture (or "demand") was already calculated for both regions separately in paragraph 2.5. Thus, this value is used for modelling the increase in cultivated pasture for both regions.

<u>Wet season model:</u>
As stated, this component models one process:

- A small number of cultivated pastures disappear. However, please notice that the area of cultivated pasture that is lost is many times smaller than the yearly increase in cultivated pasture.

Supposedly, these locations are abandoned because they turn out to be unsuitable for cultivation (because of flooding; see paragraph 2.6). Therefore, the model assumes that locations where cultivated pasture has disappeared are not suitable for cultivated pasture; i.e. ranchers will not make a second attempt to cultivate a location if the first attempt was already a failure.

Cell states
In the Pantanal-model, three cell states exist: (1) not cultivated, (2) cultivated and (3) not available. Obviously, cell state (1) and (2) are included since this study models the conversion of natural habitat into cultivated pasture. Cell state (3) is included because some locations are not suitable for cultivated pasture. The model should not consider these locations. The model does not consider the following locations:

- Locations, which are covered by rivers/water.
- Locations that have ecotype "bare soil"; this ecotype is very unfertile and unsuitable for cultivated pasture (Jongman et al. 2005).
- Locations located outside the research area.
- Locations that were cultivated earlier but consequently abandoned; the assumption is that farmers will not make a second attempt to cultivate a location if the first attempt was already a failure.

The dry season model converts cells that are (1) not cultivated to (2) cultivated pasture, and does not consider cells with cell state 3 (not available). This is logical as the dry season model

aims to simulate the conversion of natural habitat into cultivated pasture (through the processes of "Expansion" and "Patching").
The wet season model converts cells from (2) cultivated to (3) not available, and does not consider cells with cell state 1 (not cultivated). This is logical as the wet season model aims to simulate the disappearance of cultivated pasture. As stated, the assumption was made that these abandoned pastures will not be used for cultivation any longer and this explains the transition of cells from (2) cultivated to (3) not available.

Driving factors of change

Dry season model:

The dry season model simulates the conversion of natural habitat into cultivated pasture (through the sub processes of "Expansion" and "Patching"). Therefore, the driving factors of cultivation as discussed in paragraph 2.2 should be included in the dry season model:

- Neighbouring analysis, see section neighbourhood analysis.
- Flooding duration
- Elevation
- Ecotype
- Geomorphology

At first, the weights of the driving factors will be based on "expert knowledge". Later, the model will be calibrated and the weights of the driving factors will be consequently adjusted. Initially, the following weight formula will be used:

SV = 5 * Ecotype + 5 * Elevation + 5 * Geomorphology + 5 * Flood + 5 * Neighbourhood [Equation 6]

SV = suitability value

Wet season model:

The wet season model simulates the disappearance of cultivated pasture. As stated, the assumption is that because of flooding cultivated pastures become severely damaged and are abandoned (paragraph 2.6). Therefore, the only driving factor of this process is assumed to be flood duration; according to Jongman et al. (2005), cultivated pastures are irreversibly damaged if they are flooded more than 6 months per year. The wet season model selects all cultivated cells that are flooded more than 6 months per year according to the flood duration map, and changes them into ecotype "bare soil" (which is very unfertile and unsuitable for cultivated pasture and these cells are thus no longer available for cultivation; see discussion of cell states above).

Input

Dry season model:

The dry season model has a number of geo-datasets as input. These geo-datasets contain information regarding the driving factors of cultivation that were discussed above. All these individual datasets were reclassified. This reclassification was based on the relative suitability of the attributes of the dataset. Each dataset was reclassified into the range 0 – 10000, where 0 indicates a low suitability and 10000 a high suitability. All input geo-datasets are grids because we are using a CA.

- Ecotype map: This contains information regarding land use and vegetation.
- Flood duration map: This shows how many months per year a location is flooded.
- Geo-morphological map: This shows the main geo-morphological patterns.
- DEM: This shows the elevation map of each location.

In addition, the dry season model has the following input:

- Demand: This table contains the expected increase or "demand" in cultivated pasture for each year the model is run.
- Parameters: This table contains the parameters, which contains the different parameters that are used throughout the model.

Wet season model:

The wet season model has two geo-datasets as input: flood duration and ecotype. This is logical because the wet season model simulates the disappearance of cultivated pasture and the assumption was that the only driving factor of this process is flood duration. In addition, the cell state (cultivated, not cultivated, or not available) of each location needs to be known; this can be derived from the ecotype map and this explains why the ecotype map is used as input. These geo-datasets are not reclassified. Again, all input geo-datasets are grids because the Pantanal-model uses a CA. The input geo-datasets are:

- Ecotype: This shows the ecotype of each cell.
- Flood duration: This shows how many months per year a location is flooded.

In addition to the geo-datasets, the following input is required by the model :

- Parameters: This table contains the parameters, which contains the different parameters that are used throughout the model.

Neighbourhood analysis:

The dry season model has two components with a neighbourhood function: "Expander" and "Patcher". The wet season model does not have a neighbourhood function. Thus, this section will be limited to "Expander" and "Patcher" of the dry season model.

Expander:

The neighbourhood function of "Expander" tries to simulate the "Expansion" process: existing cultivated pastures expand and increase at their edges. To simulate this, the neighbourhood function selects cells around cultivated pastures and increases their suitability value. This makes that these cells are very likely to change to cultivated pasture during the allocation. In this way, cells at the edge of cultivated pasture are very likely to be cultivated; this is in agreement with the "Expansion" process. Of course, only cells that can change to cultivated pasture (i.e., cells having cell state 1: not cultivated) are selected.
As stated, the neighbourhood function selects cells around cultivated pasture and increases their suitability value. This neighbourhood function uses a classical Moore neighbourhood. The reason for choosing this neighbourhood was that analysis showed that existing areas of cultivated pasture had expanded their radius on average 1000 meters during 1991– 2001, which equals 100 meters per year. With a resolution of 90m by 90 m, this approximates to an expansion of 1 cell per year in each direction; this equals a classical Moore neighbourhood. Thus, a classical Moore neighbourhood was chosen because it (theoretically) reflected the observed process best.

Patcher:

The neighbourhood function of "Patcher" tries to simulate the process of "Patching": new, isolated "islands" or clusters of cultivated pasture suddenly appear away from existing areas of cultivated pasture. Analysis showed that these clusters had the following characteristics:

1. Most clusters are on average located higher than 110 cm.
2. Clusters are mainly covered with savannah.
3. Most clusters are on average flooded less than 2 months per year.
4. Clusters are never larger than 700 ha.

To simulate the process of "Patching", the neighbourhood function selects cells that meet characteristic 1 until 3, and looks if "clusters" of these selected cells exist; i.e. if some of the

selected cells neighbour each other. Cells that were inside a Moore neighbourhood were considered to neighbour each other. A Moore neighbourhood was chosen because it includes all direct neighbours, which was assumed to be the most logical way to represent clusters of cells.
The neighbourhood function then calculates the area of each cluster, selects clusters smaller than 700 ha (characteristic 4) and increases the "suitability value" of cells within these clusters. In this way, clusters of cells with the same characteristics as the "real-world" clusters of the "Patching" process are given a high suitability value and are thus very likely to be cultivated during the allocation; this is in agreement with the "Patching" process.

Data structure
Paragraph 4.1.2 already stated that a CA (or simulation model) is most appropriate to model land use change in the Pantanal. Since CA is based on grid cells that dynamically interact, the data structure of the model is raster.

Cell size
All available data for our study is available in cell size 90m by 90m. A smaller cell size is thus not a possibility, whereas converting the data into a cell size larger than 90m by 90m is useless; more details will fall out and the advantage of reduced computing time is limited as most computers can handle spatial data with cell size 90m by 90m. Thus, the model will use cell size of 90m by 90m.

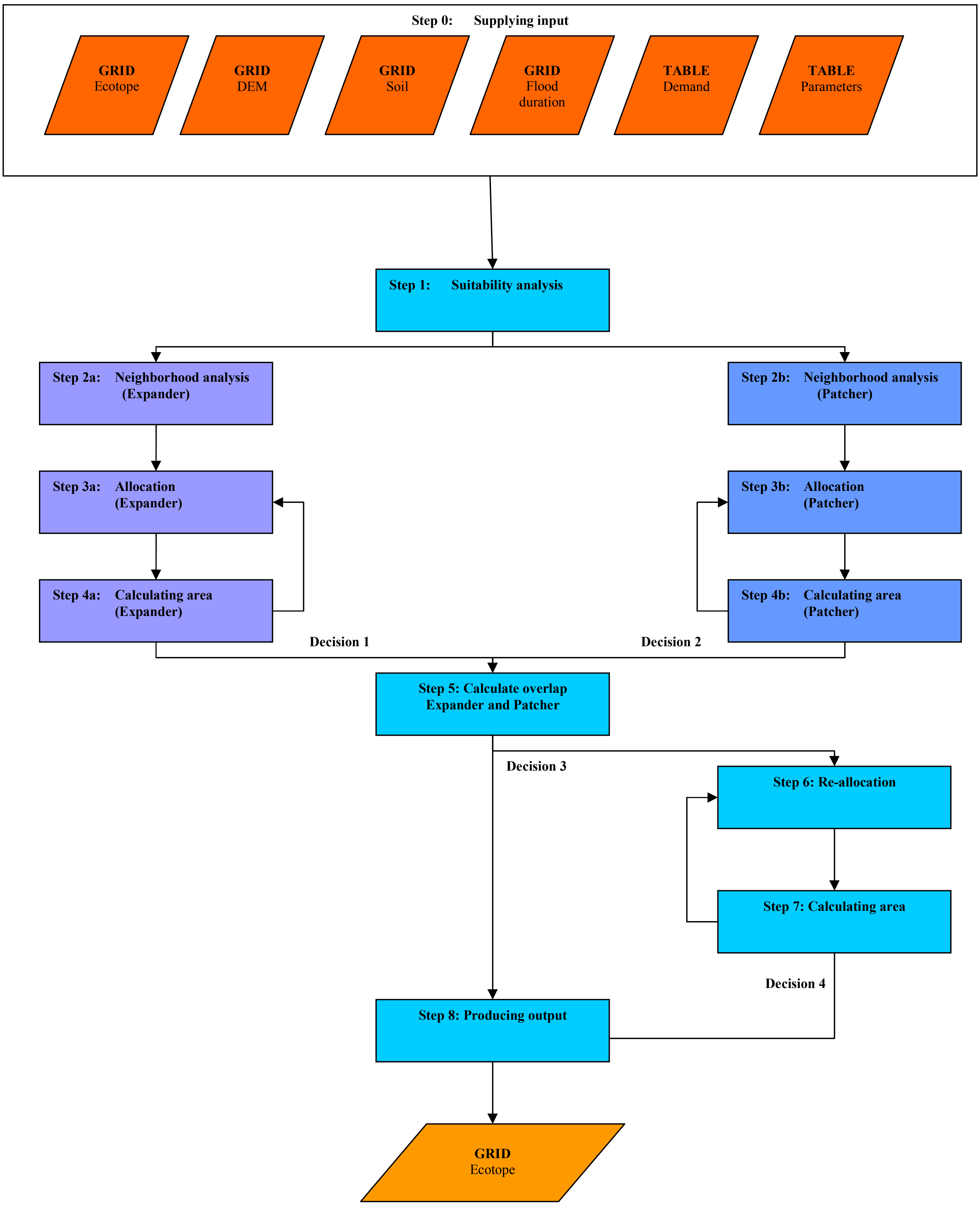
Time period
Many land use change models extrapolate current land use changes to calculate the expected change ("demand"). Our study also uses this approach. However, these studies do not address the underlying causes of land use change and do not actually address the underlying mechanisms of land use change. Therefore, they have assumed that the relative influence of the driving factors will stay the same and they can not model over a too long period (Irwin and Geoghegan 2001).
Since the Pantanal-model also uses extrapolation to calculate the "demand", it is clear that it should not model over a too long period of time. Therefore, we have chosen to model until 2021, which is a time period comparable to other land use change models that have used extrapolation (CLUE-S and DINAMICA, respectively 8 and 15 years).

4.2 Formalization of the model

As stated, the Pantanal-model has two distinct components: a dry season model and a wet season model. Both will be discussed separately.

4.2.1 Dry season model

As stated, the dry season model models the increase in cultivated pasture during the dry season. The flow of activities of the dry season model is presented in a conceptual model in figure 4.2.1. This is followed by a description of the conceptual model. Please notice that two functions of the dry season, i.e. "Expander" and "Patcher", are executed simultaneously.


Step 0: Supplying input
GRID Ecotope
GRID DEM
GRID Soil
GRID Flood duration
TABLE Demand
TABLE Parameters
Step 1: Suitability analysis
Step 2a: Neighborhood analysis (Expander)
Step 2b: Neighborhood analysis (Patcher)
Step 3a: Allocation (Expander)
Step 3b: Allocation (Patcher)
Step 4a: Calculating area (Expander)
Step 4b: Calculating area (Patcher)
Decision 1
Decision 2
Step 5: Calculate overlap Expander and Patcher
Decision 3
Step 6: Re-allocation
Step 7: Calculating area
Decision 4
Step 8: Producing output
GRID Ecotope

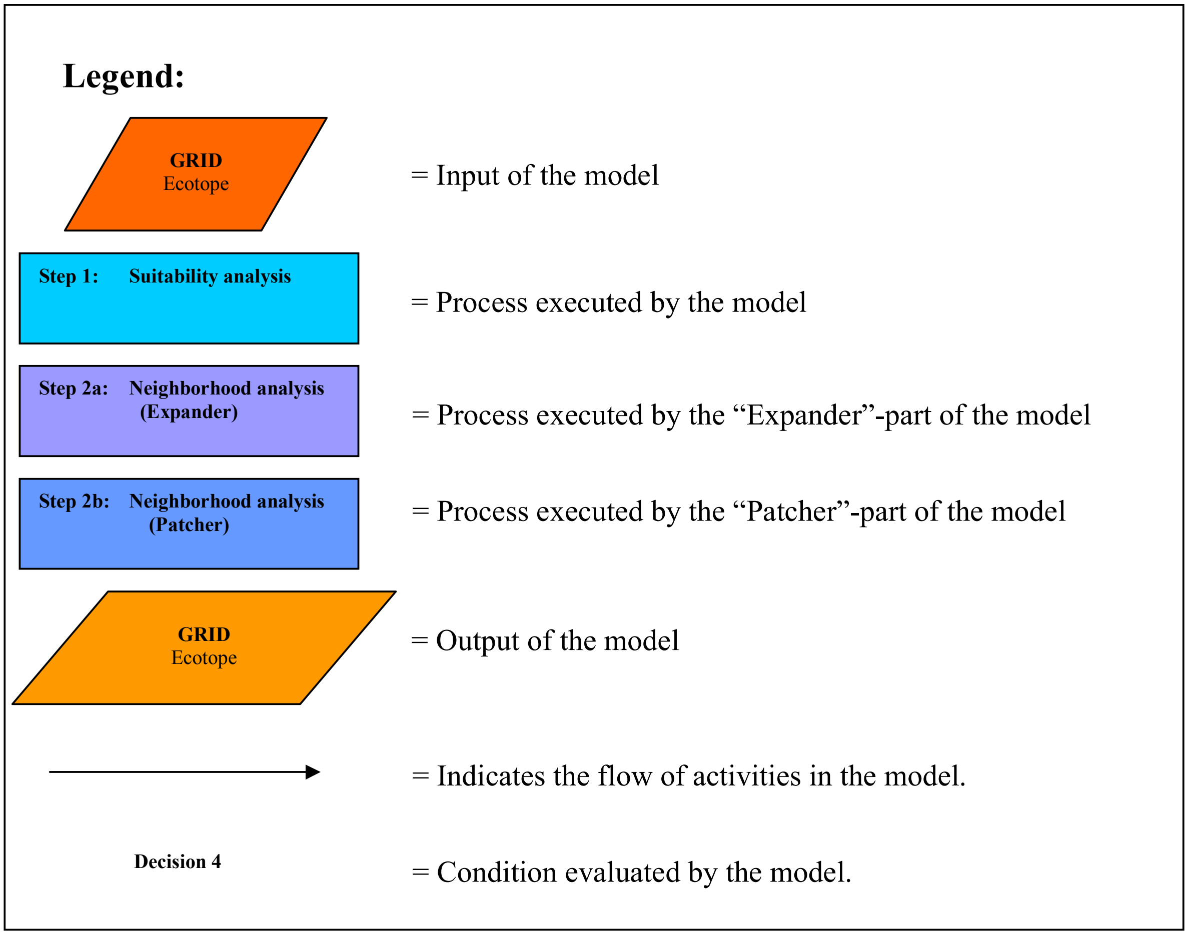


Figure 4.2.1. Flow of activities of the dry season model represented in a conceptual model.

Step 0: Supplying input
The dry season model needs input. The following input should be available in the different steps (activities) executed by the model:

- Grid: Ecotype map
  - This contains information regarding land use and vegetation.
- Grid: Flood duration map
  - This shows how many months per year a location is flooded.
- Grid: Geo-morphological map
  - This shows the main geo-morphological patterns.
- Grid: DEM
  - This shows the elevation map of each location.
- Table: Demand
  - This contains the expected increase or "demand" in cultivated pasture for each year the model is run.
- Table: Parameters
  - This contains the parameters, which contains the different parameters that are used throughout the model.

Step 1: Suitability-analysis
In this step, the model gives each cell a value, which indicates how suitable a cell is for cultivated pasture (hereafter referred to as "suitability value"). It was assumed that the following driving factors determine suitability; ecotype type, elevation, geomorphology and

flood duration. Consequently, all these driving factors are considered in calculating the "suitability value". The "suitability value" of a cell is calculated by combining its values in each dataset according to a weight-formula (equation 7). These weights were determined based on expert knowledge (see paragraph 3.2).
The assumption is that some locations are very unsuitable for cultivated pasture. The model recognizes these locations and does not consider them during the modelling; i.e., they can not change to cultivated pasture. Consequently, these locations are not given a suitability value. The following locations are considered very unsuitable and are not considered:

- Locations, which are covered by rivers/water.
- Locations that have ecotype "bare soil"; this ecotype is very unfertile and unsuitable for cultivated pasture (Jongman et al. 2005).
- Locations located outside the research area.
- Locations that were cultivated earlier but consequently abandoned; the assumption is that farmers will not make a second attempt to cultivate a location if the first attempt was already a failure.

Step 2a: Neighbourhood analysis (Expander)
During the neighbourhood analysis of "Expander", all cells in the ecotype map with cultivated pasture are selected. Then, the "suitability value" of cells within a Moore neighbourhood around the selected cells is increased with 50000.

Step 2b: Neighbourhood analysis (Patcher)

- During the neighbourhood analysis of "Patcher", a number of cells are selected. The selected cells have the following characteristics:
    - Locations higher than 110 cm.
    - Locations covered with savannah.
    - Locations flooded less than 2 months per year.

The model looks for "clusters" of selected cells; i.e. if there are selected cells that neighbour each other (using a Moore neighbourhood). The model calculates the area of each cluster, and increases the suitability value of the cells of clusters with an area smaller than 700 ha.

Step 3a: Allocation (Expander)
Model sets a threshold, and selects all cells with a "suitability value" equal or higher than the threshold. These cells will change to cultivated pasture. The threshold is equal to the "suitability value" of the cell(s) with the highest "suitability value" in the grid.

Step 3b: Allocation (Patcher)
Model sets a threshold, and selects all cells with a "suitability value" equal or higher than the threshold. These cells will change to cultivated pasture. The threshold is equal to the "suitability value" of the cell(s) with the highest "suitability value" in the grid.

Step 4a: Calculating area (Expander)
Model calculates the total area of the cells selected during the previous "Expander"-step ("3a Allocation (Expander)"). Decision 1 is then evaluated.

> ***Decision 1:***
> IF the total area of the selected cells is lower than the demand, the model returns to the previous step ("3a: Allocation (Expander)") and the threshold will be equal to the "suitability value" of the cell(s) with the second highest "suitability value" in the grid. This process will iterate, and each time the threshold will be one "suitability value" lower.

ELSE, the model continues with the next step ("5: Combining Expander and Patcher").

Step 4b: Calculating area (Patcher)
Model calculates the total area of cells selected during the previous "Patcher"-step ("3b Allocation (Patcher)"). Decision 2 is then evaluated.

Decision 2:
IF the total area of the selected cells is lower than the demand, the model returns to the previous step ("3b: Allocation (Patcher)") and the threshold will be the "suitability value" of the cell(s) with the second highest "suitability value" in the grid. This process will iterate, and each time the threshold will be one "suitability value" lower.
ELSE, the model continues with the next step ("5: Combining Expander and Patcher").

Step 5: Calculate overlap Expander & Patcher
Model compares how many cells have been allocated twice; i.e. both "Expander" and "Patcher". Decision 3 is then evaluated.

Decision 3:
IF the total area of cells that have been changed to cultivated pasture by both "Expander" and "Patcher" is equal or more than 5% of the demand, the model continues with step 6 ("Re-allocation overlapping cells").
ELSE, the model continues with step 8 ("Producing output").

Step 6: Re-allocation
Model sets a threshold, and selects all cells with a "suitability value" equal or higher than the threshold. These cells will change to cultivated pasture. The threshold is equal to the "suitability value" of the cell(s) with the highest "suitability value" in the grid.

Step 7: Calculate area
Model compares how many cells have been allocated in the previous step. Decision 4 is then evaluated.

Decision 4:
IF the total area of cells that have been allocated in the previous step is equal or more than 5% of the demand, the model returns to the previous step ("6: Re-allocation"). The threshold will be the "suitability value" of the cell(s) with the second highest "suitability value" in the grid. This process will iterate, and each time the threshold will be one "suitability value" lower.
ELSE, the model continues with step 8 ("Producing output").

Step 7: Producing output
A new ecotype map is produced. This ecotype map is an update of the input ecotype map, in which a number of cells have changed to cultivated pasture according to the allocation procedures described above.

4.2.2 Rain season model

As stated, the wet season model models the decrease in cultivated pasture during the wet season; i.e. cultivated pasture is changed back into bare soil. The flow of activities of the wet season model is presented in a conceptual model in figure 4.2.2. This is followed by a description of the conceptual model.

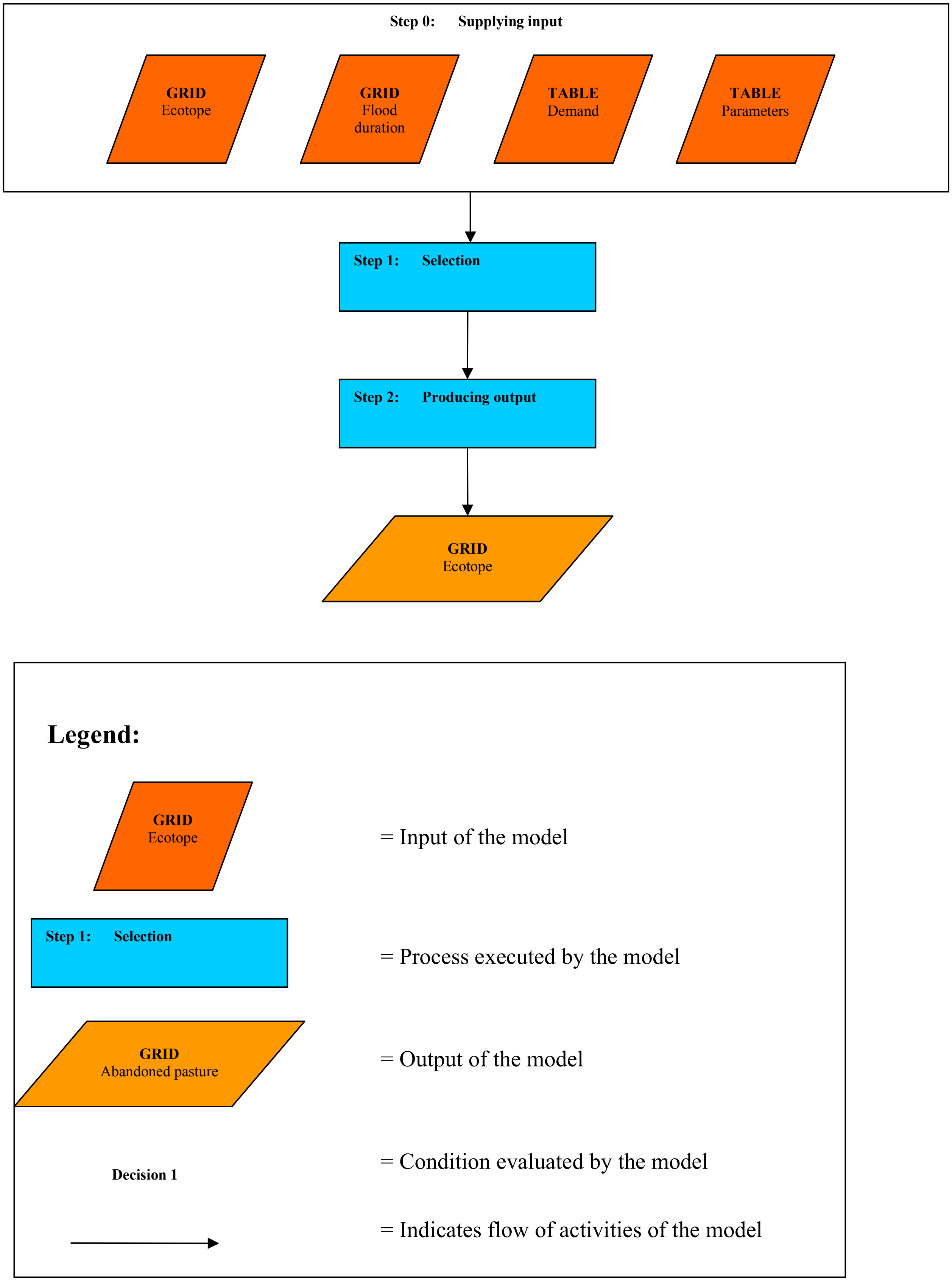


Figure 4.2.2. Flow of activities of the wet season model represented in a conceptual model.

Step 0: Supplying input
The wet season model needs input. The following input should be available in the different steps (activities) executed by the model:

- Grid: Ecotype, which shows the ecotype of each cell.
- Grid: Flood duration, which shows the duration of the floods in the dry season for each cell.
- Grid: Abandoned pasture, which shows which locations were already cultivated but consequently abandoned because cultivation was not a success. Cells in this map have either a value of "0" (not yet cultivated) or "1" (already cultivated and abandoned).
- Table: Demand, which contains the expected increase or "demand" in cultivated pasture for each year that the model is run.
- Table: Parameters, which contains the different parameters that are used throughout the model.

Step 1: Selection
In this step, the model selects all cells with cultivated pasture that have flood duration of more than 6 months per year. These cells will be changed back into ecotype type bare soil. It was assumed that the following locations are irrelevant for this process and these are not considered:

- Cells with other ecotypes than cultivated pasture.

Step 2: Producing output
A new ecotype map is produced. This ecotype map is an update of the input ecotype map, in which a number of cells have changed to bare soil according to the allocation procedure described above.
In addition, the map with abandoned pastures is updated, which shows the locations that were cultivated and consequently abandoned. All cells that have changed back to bare soil according to the allocation procedure above are assigned a value of "1". This indicates that they were cultivated and consequently abandoned. These cells will not be available for processing in the dry season model and thus will not return to cultivated pasture.

5 Test Run & Calibration

This chapter will address the following research questions related to our model:

- How "robust" is the Pantanal Land Use model (sensitivity)?
- How reliable is the Pantanal Land Use model (calibration)?

This chapter will start by describing the test runs; these are indicative for the reliability of a model.
Then, this chapter will address the "robustness" of the model by means of a "sensitivity analysis"; i.e. it will make small changes in a parameter of the model (while keeping all other parameters constant) to determine if the model is sensitive to changes in this component and if the output of the model remains (relatively) stable. This is done for a number of parameters.
Then, this chapter will address the reliability of the model by means of a calibration. This calibration will consist of modelling the historical period 1974-2001 and adjusting the model until it fits the historical data reasonably well. This will also give an indication of the accurateness of the model.

5.1 Test runs

5.1.1 First test run

Introduction:
The first test run used the exact model described in paragraph 5.2 "Formalization of the model".
Results:
The first test run produced no reasonable output but revealed two major problems:

1) The computation time of the dry season model was too long. Because of this, the computation time of the entire model for modelling a period of 21 years amounted to about 10 hours. Such a long computation time is unpractical as experimenting with the model becomes impossible. Experimenting often improves a model, because it can reveal the sensitivity of the model to extreme input values, different parameters, etc. The long computation time is therefore a serious problem.
2) The area of cultivated pasture allocated by the dry season model was very inaccurate, and did not match the calculated "demand" by a long way. Although the dry season model did allocate cultivated pasture to the grid as it was supposed to do, it allocated too much or too little; i.e. difference was often more than 200% of the "demand". The reason for this was that differences in suitability were too small, and that consequently too many cells had the same suitability value. Since the dry season model selects cells based on their suitability value, it selected too many cells simultaneously. This resulted in selecting an enormous area, and made it impossible to allocate an area as small as the calculated demand.

5.1.2 Second test run

Introduction:
During the second test run, the model was changed. This was in anticipation to the two problems revealed by the first test run: the long calculation time and the inaccurate allocation.

Improvements:
To reduce the calculation time of the model, the following was changed:

- Instead of calculating on grids, the dry season model now calculates on tables.
  During the first test run, the dry season model selected cells with a suitability value higher than the threshold and produced an output grid containing the changed cells. Then, the dry season model calculated the total area of the changed cells, compared this to the demand and adjusted the threshold if necessary. However, it does all these calculations on the grid, and grid calculations are time-consuming.
  Therefore, the dry season model now makes these calculations on the attribute table; it looks into the attribute table to see how many cells have a suitability value higher than the threshold, multiplies this count with the area of 1 cell and gets the total area of selected cells. The dry season model compares this area to the demand and adjusts the threshold if necessary. It makes an output grid if the total area of the selected cells matches the demand.
  Since this procedure creates a grid only once, it saves a lot of computation time.

To increase the accuracy of the allocation, the following was changed:

- Instead of using input geo-data with a value range of 1-100, the dry season model now uses input geo-data with a value range of 1-1000 to calculate the suitability value of each cell. In this way, relative differences in suitability between cells are "blown up"; consequently, fewer cells have the same suitability value and the model will select fewer cells simultaneously. This improvement makes it possible to allocate an area as small as the calculated demand.
  Indeed, the area allocated by the model matched the demand much better.
- Instead of calculating the suitability value with the driving factors neighbouring cultivation, vegetation, elevation, and flood risk, the dry season model calculates the suitability value with an extra driving factor, i.e. type of soil. In this way, more differences in suitability will emerge between cells, fewer cells will have the same suitability value and the model will select fewer cells simultaneously.

Results:
Running the model showed that the improvements adequately solved both problems: it reduced the computation time to about three hours, and reduced the difference between the calculated "demand" and the area allocated by the dry season model to about 1%, which is acceptable.
However, this test run revealed yet another problem: the "Patcher"-function of the dry season model does not work appropriately:

- The "Patcher"-function simulates the sudden appearance of isolated "islands" or clusters of cultivated pasture. "Patcher" simulates this by searching for clusters of extremely suitable cells and increasing their suitability value; such clusters are likely to change to cultivated pasture during the allocation procedure. Running the model showed that "Patcher" could not select such clusters of extremely suitable cells; instead, "Patcher" selected more than a third of the entire study area. As too many cells are selected, the criteria to select extremely suitable cells are not adequate and should be adjusted.

5.1.3 Third test run

Introduction:
During the third test run, the "Patcher"-function of the dry season model was changed. This was in anticipation to the second test run, which showed that "Patcher" did not function appropriately; it did not select small clusters of extremely suitable cells but selected more

than a third of the entire study area instead. As too many cells were selected, the criteria to select extremely suitable cells are not adequate and should be adjusted.

Improvements:
As stated above, “Patcher” selects too many cells. Consequently, its selection criteria are not adequate and need to be “stricter”. Originally, the selection criteria of “Patcher” were based on the characteristics of the clusters of cultivated pasture that suddenly appear away from existing areas of cultivated pasture (and which “Patcher” tries to simulate). These characteristics are:

1. Elevation: located higher than 110 cm.
2. Ecotype: covered with savannah.
3. Flood duration: flooded less than 2 months per year.
4. Area: never larger than 700 ha.

As stated, these selection criteria need to be changed. Since analysis also showed that the clusters have a topographical wetness index of 9.5 m or higher, this criterion was used to make the set of selection criteria “stricter”. The topographical wetness index represents the relative elevation by comparing the elevation of each cell with its surroundings cells (and thus indicates where water is most likely to accumulate; hence its name). Consequently, it brings out local differences in elevation. Since “Patcher” also operates on a local scale by simulating the appearance of small, isolated clusters of cultivated pasture, the topographical wetness index seems an appropriate criterion.
However, the criterion topographical wetness index was not simply added to the set of criteria listed above. Instead, the criterion elevation (i.e., located higher than 110 cm.) was replaced with the criterion topographical wetness index. The reason for this replace is that elevation and topographical wetness index are related; including both would place too much emphasis on the factor elevation.
Another improvement was that instead of a Moore neighbourhood, a Von Neumann neighbourhood was used. Originally, clusters of selected cells were identified by means of a classical Moore neighbourhood. However, a Von Neumann neighbourhood is stricter (does not consider cells that are diagonal neighbours) and this is thus another possibility to select fewer clusters of cells.

Results:
Running the model showed that the improvements adequately solved the problem: “Patcher” now selects small, isolated clusters instead of a third of the study area, and this is what “Patcher” was supposed to do.

5.1.4 Conclusions

An important conclusion that can be drawn from these test runs is that the differences in suitability are small. This was clearly illustrated by the first and third test run.
During the first test run, the dry season model could not allocate the correct amount of cultivated pasture due to the small differences in suitability; i.e., too many cells had the same suitability value and the dry season model consequently selected too many cells simultaneously.
During the third test run, the “Patcher”-function of the dry season model could not select small clusters of extremely suitable cells but instead selected more than a third of the entire study area. Again, this was due to the small differences in suitability; i.e., the differences in suitability are so small that too many cells fell within the set criteria of “Patcher”, and as a result “Patcher” selected too many cells simultaneously.
This problem can seriously affect the outcome of the model. Inevitably, a number of assumptions, parameters, etc. will deviate from reality, and their uncertainties are enhanced by these small differences in suitability. This can easily result in an unrealistic output.

However, at this point it is still difficult to make conclusions about the sensitivity of the model. Therefore, the additional analyses in the next paragraphs will have to show how sensitive and realistic the model is.

5.2 Sensitivity analysis

Introduction

The aim of a sensitivity analysis is to find out which components of a model are most sensitive (Jørgensen 1986). A sensitivity analysis is done by making changes either in (1) the input of a model or (2) in the parameters of a model and comparing the resulting changes in output. The most sensitive components of a model are those that give the largest changes in output when changed (Jørgensen 1986).
There are many ways to test the sensitivity of a CA: e.g., make changes in the resolution, neighbourhood function, number of cell states, etc. and see how this affects the output of the model. However, in this study was chosen to change the weights of the driving factors and see how these changes affect the output.

Methodology

As stated, this sensitivity analysis will change the weights of the driving factors ("elevation, flood duration, geo-morphology, ecotype, and neighbourhood" and compare the changes in output. This was done in the following way:

1. A "standard" run of the model was made. In this "standard" run, all driving factors had the same value: it was arbitrarily set at 5. This "standard" run is used as a reference.
2. A new run is made in which the weight of one of the driving factors is set at 10 and the weight of all others is kept constant at 5. The output of this run is compared with the "standard" run by determining the number of cells that overlap. This overlap is a measure of the sensitivity of this parameter.
3. A second run is made in which the weight of the same driving factor is set at 1 and the weight of all others is kept constant at 5. Again, the output of this run is compared with the "standard" run by determining the number of cells that overlap.
4. This procedure is repeated until all driving factors have been included. Then, we know what the effect is of increasing or decreasing the weight of each driving factor since this is reflected in the overlap with the "standard" run. The parameter, which has the least overlap, is the most sensitive.

Table VII. The weights of the driving factors during the sensitivity analysis. In each run, the weight of one of the driving factors is increased or decreased, while the weights of other driving factors are kept constant. In the "standard" run all driving factors have the same value because it is used as a reference.

Driving factor:	Standard run	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10
Elevation	5	10	1	5	5	5	5	5	5	5	5
Flood duration	5	5	5	10	1	5	5	5	5	5	5
Geomorphology	5	5	5	5	5	10	1	5	5	5	5
Ecotope	5	5	5	5	5	5	5	10	1	5	5
Neighborhood	5	5	5	5	5	5	5	5	5	10	1

Results

Figure 5.2.1 presents the results of the sensitivity analysis. No large differences exist in these results; i.e. all driving factors have a comparable effect on the output if they are changed. However, some minor differences exist. The driving factors "flood duration" and "neighbourhood" have the smallest effect if they are changed. The driving factors "elevation" and "geo-morphology" have the greatest effect if they are changed. The driving factor "ecotype" has almost no effect if its weight is increased, but it has the greatest effect of all driving factors if its weight is decreased.

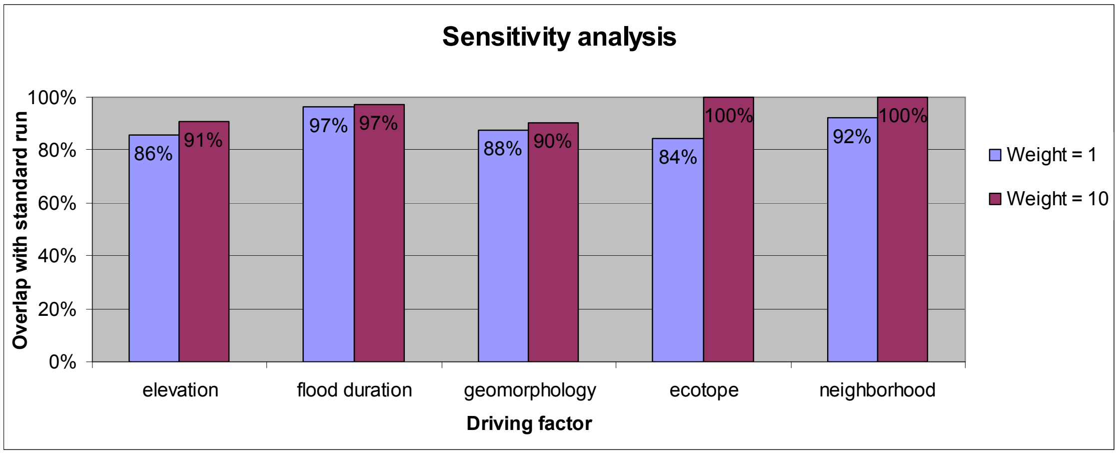


Figure 5.2.1: Results of the sensitivity analysis. In this sensitivity analysis, the weight of one of the driving factors was increased (purple) or decreased (blue), while the weights of other driving factors were kept constant; the resulting output was compared with the standard run (by giving the percentage of cells that overlaps with the standard run).

Conclusions

All driving factors have had a comparable effect on the output after changing their weights. This indicates that the correct set of driving factors was chosen. Driving factors, which do not have any effect on the output, are redundant and should be left out of the model. Alternatively, extremely sensitive driving factors are also problematic. They have a disproportionate large influence on the output, and this simplifies the model; i.e., the output becomes merely a function of this one driving factor, instead of a synergy of a number of driving factors. Fortunately, this is not the case and all driving factors have a comparable effect.

Although all driving factors have had a comparable effect on the output, some minor differences exist. Driving factors "elevation" and "geomorphology" have the greatest effect if they are changed. Driving factors "flood duration" and "neighbourhood" have the smallest effect. Driving factor "ecotype" has a sensitivity that lies between these both groups. Thus, this sensitivity analysis shows that "elevation", "geomorphology" and "ecotype" are the most sensitive parameters of the model. As a consequence, if their value would deviate only slightly from reality, this could already result in an unrealistic output. Therefore, it is essential to determine the true value of these three parameters. This can be done best by means of a calibration. The next paragraph will discuss this extensively.

5.3 Calibration and validation

Introduction

Calibration is trying to achieve the best fit between the output of a model and a real, known dataset by changing the different parameters of a model (Jørgensen 1986). Validation is testing how well a model fits a real, independent dataset (Jørgensen 1986); i.e. the output of a model is compared to a real-world dataset, which was not used to build or calibrate the model. The main difference between calibration and validation is that during calibration the model is changed to fit the data, whereas during validation the fit of the model with the data is merely determined; this gives an indication of the quality of the model. Therefore,

validation can be considered as a "quality test", and calibration as an "optimization" of the model. Normally, calibration precedes validation and the calibrated model is used for validation. However, data for validation should be entirely different from the data used in calibration to ensure that the validation remains an objective, independent test (Jørgensen 1986).
Unfortunately, this is not possible for the Pantanal model. There is only one dataset available. Since it is not possible to use the same dataset for both calibration and validation, we have chosen to use this dataset for calibration and not for validation. Validation would merely indicate the quality of the model but would not improve the model. Since the aim is to use the model in scenarios studies (see next chapter), it is better to improve the model further through calibration so that it can generate reasonable results during the scenario studies.

Methodology
To calibrate the model, the period 1976-2000 will be modelled. A land use map is available of the year 2000, and it is thus possible to compare the output of the model with this real-world data. However, there is one problem: a land use map is not available of the year 1976 and this land use map is necessary as input of the model. To solve this, an ecological model was used to make a land use map of 1976. This ecological model was developed in an earlier phase in the Pantanal-Taquari project and it generates a vegetation map of the Pantanal based on certain physical characteristics such as soil type, flooding frequency, etc. This physical data is available for 1976 and it is therefore possible to generate a vegetation map of 1976. This vegetation map of 1976 was combined with a dataset of the cultivated pastures in 1976. This dataset contains only cultivated pastures and no other land use types (forest, savannah, etc.). Therefore, it must be combined with the vegetation map of 1976 to generate a complete land use map of 1976 that can be used as input to calibrate the model.
As stated, the actual calibration of the model will consist of modelling the period 1976-2000 and comparing the output of the model with real-world data of the year 2000. Different parameters of the model will be changed to achieve the best fit between the output of a model and the dataset. During the sensitivity analysis, it became that clear that the driving factors "elevation", "geomorphology" and "ecotype" are the most sensitive parameters of the model. It is logical to calibrate the model by changing these three most sensitive parameters; i.e., if their value deviates only slightly from reality, this can already result in an unrealistic output.
The model was calibrated in the following way:

1. A series of five consecutive runs was made in which the weight of driving factor "elevation" was increased during each run, while other parameters were kept constant.
2. The resulting output was compared with the land use map of 2000 by giving the percentage of overlapping cells. The percentage of overlapping cells is considered as a measure of the fit of the model with reality.
3. The fit of the model was plotted against the weight of the selected driving factor. From this plot, it is possible to determine the optimum value of the parameter; i.e., the value of the parameter that results in the best fit.
4. This is repeated for the other very sensitive parameters of the model: "geomorphology" and "ecotype". In this way, the optimum values of these three most sensitive parameters are determined. In determining the optimum values of the driving factors, only one parameter was taken as a variable at a time. However, it is also possible to take simultaneously several parameters as variables; in this way the best fit is achieved by testing combinations of parameters. Of course, this is more realistic since often complex relations between parameters exist and this makes that combinations of parameters behave differently than parameters do separately.

However, so many combinations of parameters are possible that doing this in a systematic way requires a lot of (computation) time. This study did not have enough time to do this.

Results:
Figure 5.3.1 shows the results of the calibration of parameter "elevation". This figure clearly shows an optimum curve between the fit of the model and the weight of the parameter. The optimum value of this parameter lies at 5.

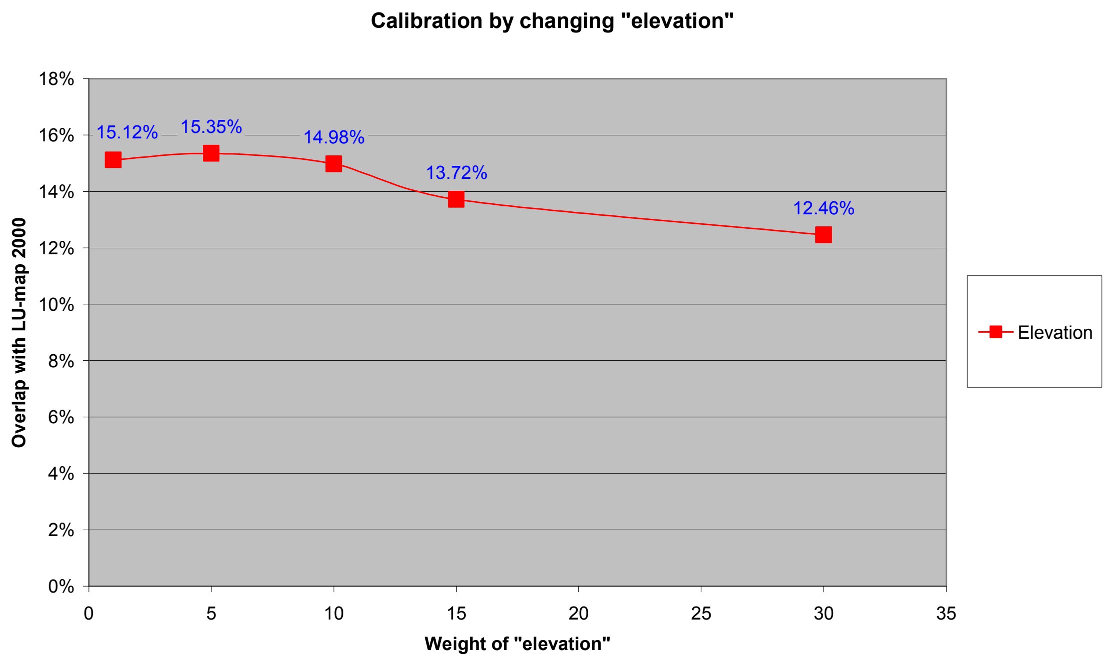


Figure 5.3.1. Results of the calibration of the driving factor "elevation". In this calibration, the weight of the driving factor "elevation" was increased during five consecutive runs, while the weights of other parameters were kept constant; the resulting output was compared with the land use map of 2000 (by giving the percentage of overlapping cells).

Figure 5.3.2 shows the results of the calibration of parameter "geomorphology". Interestingly, this figure clearly shows that there is a negative relation between the fit of the model and the weight of the parameter; i.e., if the weight increases, the fit of the model decreases. The figure even suggests that the best fit is achieved if this parameter is totally left out (i.e., value set at zero). However, it is not possible to set the value of a parameter at zero due to technical reasons; the lowest possible value is 1.
Figure 5.3.3 shows the results of the calibration of parameter "ecotype". The figure shows a "saturation" curve between the fit of the model and the weight of the parameter; i.e., the fit of the model becomes better if the weight is increased, until an optimum value is reached after which the fit remains optimal. This optimum value lies at 10.
The optimum values of the parameters "elevation", "geomorphology", and "ecotype" are given in table VII. Although the values of "flood duration" and "neighbourhood" are not determined by means of calibration, they are also included in table VII. These parameters have the standard value of 5.
The model was also run with the optimum values given in table VII. Figure 5.3.4 presents the output of this run. In figure it is possible to see which areas were allocated correctly, and which areas were allocated wrongly. Running the model with the optimum values resulted in a fit of 17 %. This seems a poor performance, but such a conclusion is not entirely correct. The model models a vast geographical area and uses a very high resolution (90m by 90m). Moreover, the aim of the Pantanal-model is to use it in scenario studies, and not to predict

with an accuracy of 90m by 90m (i.e., resolution) where exactly cultivated pasture will be constructed. Therefore, it is more important that the model is able to predict broadly in which direction land use change will develop to be useful as a management tool. Thus, it is necessary to look at a larger scale at the output of the model to say something about its usefulness as a management tool. To do this, this study scaled-up the resolution; the study area was divided in 10 km by 10 km squares. The area of cultivated pasture in each square is calculated according to the land use map of 2000. In addition, the area of cultivated pasture in each square is calculated according to the output of the model. The discrepancy between the area in reality and the area predicted by the model is a measure for the fit of the model, but in this way the results of the calibration are "scaled-up" to a more appropriate scale. Figure 5.3.5 presents the "scaled-up" results of the calibration. At this scale, it becomes clear that the model has allocated very correct in most regions. However, there are two regions in the eastern part of the Pantanal (dark blue) in which the model allocated less cultivated pasture than in reality (underestimation). There is one region in the eastern part of the Pantanal (red) in which the model allocated more cultivated pasture than in reality (overestimation).

Table VIII. Optimum values of the different driving factors determined by means of calibration. The values of "flood duration" and "neighbourhood" were not determined by means of calibration. Instead, they were given the standard value of 5.

Driving factor:	Optimum value:
Elevation	5
Geomorphology	1
Ecotype	10
Flood duration	5
Neighbourhood	5

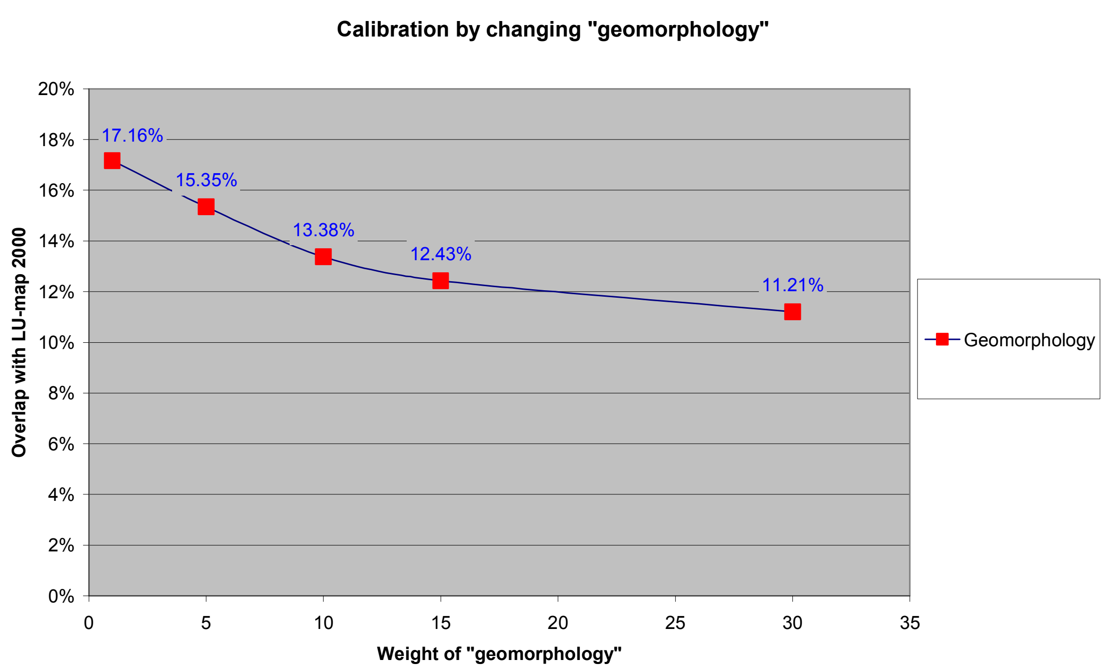


Figure 5.3.2. Results of the calibration of the driving factor "geomorphology". In this calibration, the weight of the driving factor "geomorphology" was increased during five consecutive runs, while the weights of other

parameters were kept constant; the resulting output was compared with the land use map of 2000 (by giving the percentage of overlapping cells).

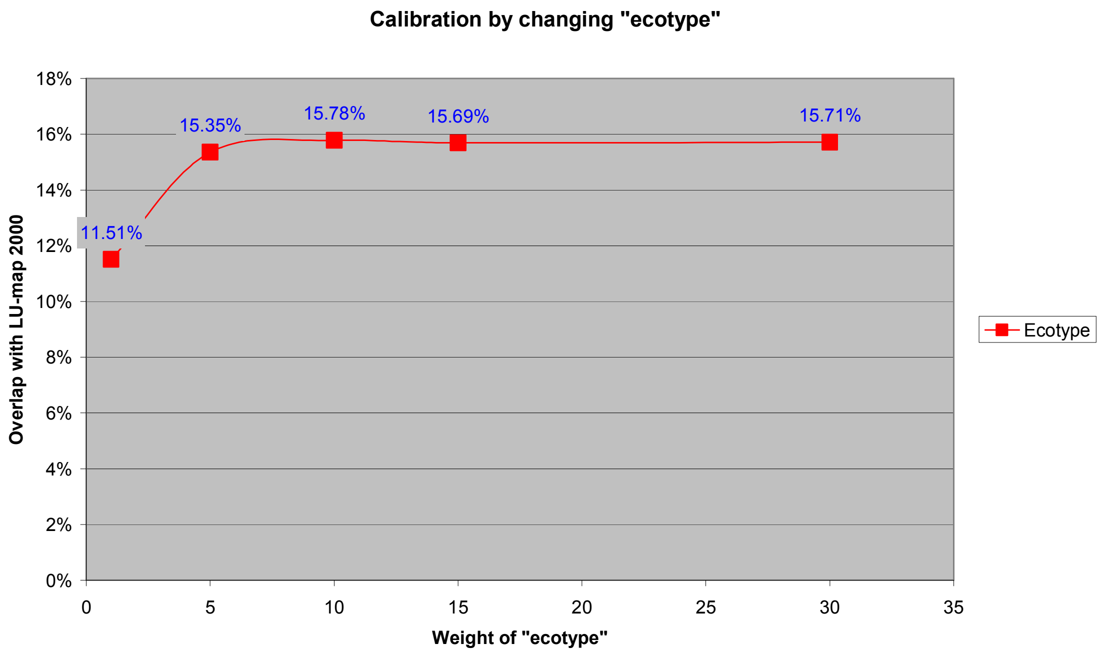


Figure 5.3.3. Results of the calibration of the driving factor "ecotype". In this calibration, the weight of the driving factor "ecotype" was increased during five consecutive runs, while the weights of other parameters were kept constant; the resulting output was compared with the land use map of 2000 (by giving the percentage of overlapping cells).

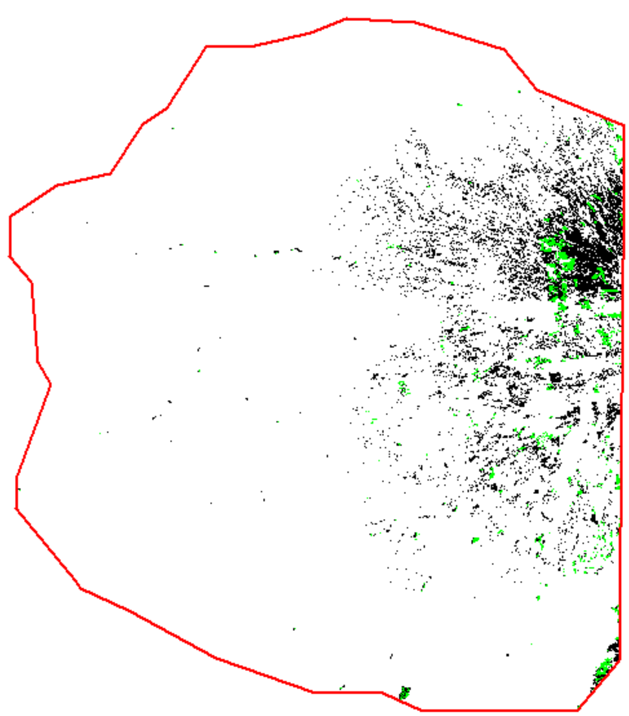


Figure 5.3.4. Output of running the model with optimum values. Red indicates the outline of the study area. Black indicates that cultivated pastures were not allocated correctly by the model (according to the land use map of 2000). Green indicates that cultivated pastures were allocated correctly by the model (according to the land use map of 2000).

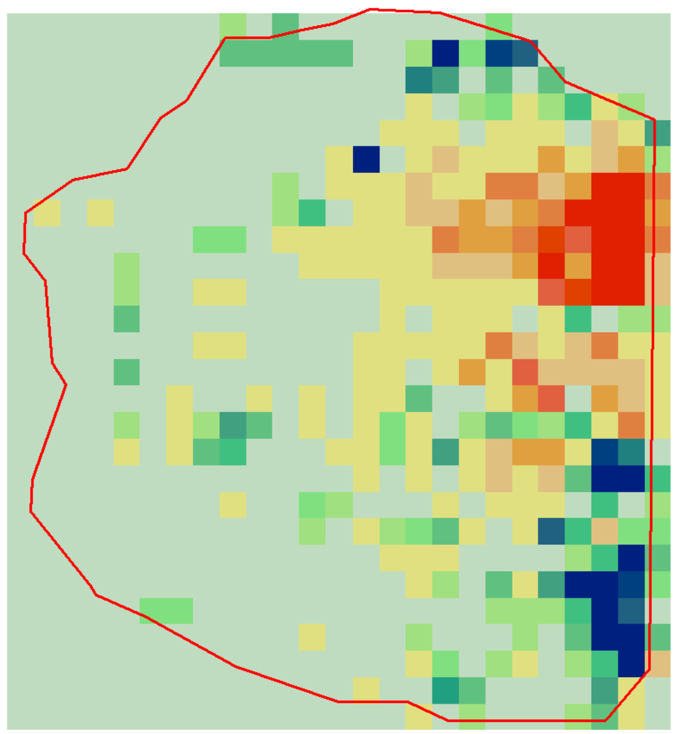


Figure 5.3.5. "Scaled-up" results of the calibration. As stated, the model was also run with the optimum values, and this figure presents the overlap of the output with the land use map of 2000 at a scale of 10 km by 10 km. Light green indicates that the area cultivated pasture allocated according to the model perfectly matches the area of cultivated pasture in reality. Dark blue indicates that the model allocated less cultivated pasture than in reality (underestimation). Red indicates that the model allocated more cultivated pasture than in reality (overestimation).

Conclusions:

Two main conclusions can be drawn from these results:

- Apparently, "geomorphology" is not a good driving factor to represent the process of cultivation in the Pantanal, and should not be used to predict the locations of change. This became clear during the calibration, which suggested that "geomorphology" is best left out of the model.
- At a fine scale, running the model with the optimum values resulted in a fit of 17 %. This fit was determined by calculating the percentage of overlapping cells with cultivated pasture that overlap between the output of the model and the land use map of 2000. Although a fit of 17% seems a poor performance, such a conclusion is not entirely correct.

  First, the input was a land use map of 1976, which was not an existing land use map but was reconstructed using an ecological model; inevitably, errors will have been produced in creating this 1976 land use map.

Secondly, the Pantanal-model is intended as a management tool. Therefore, it is merely important that the model is able to predict broadly in which direction land use change will develop, and to look at the fit of the model at a larger scale. After looking at a larger scale, it became clear that the model allocates very correct in most regions. However, it also became clear that the model underestimated the increase in cultivation in the south-eastern and north-eastern parts of the Pantanal. It is difficult to say why the model underestimated the increase in these parts, but it is striking that in these regions the infrastructure is relatively good (they are good accessible by either road or river; pers. comm. M. van Eupen). This suggests that socio-economic factors play a role.

In addition, the model overestimated the increase in cultivation in the central eastern part of the Pantanal. Although this area has not become extremely cultivated during the 1976-2000 period (which is the calibration period) and the model has clearly overestimated its cultivation for this period, it is interesting to see that in most recent years the increase in cultivation in this area is considerable; it seems that the model was right in predicting this trend but was wrong in its time period!

Still, looking at all these aspects, I feel it is safe to say that the overall performance of the model is good and that it is good enough to be used in scenario-studies.

6 Scenario-study

In this chapter, the Pantanal-model will be used to explore future land use in three scenarios. These three scenarios were discussed extensively in paragraph 2.4 and they are:

1. Extrapolation of the current trend of habitat conversion in the Pantanal (figure 2.4.3).
2. Extrapolation of the lower 90% confidence interval of scenario 1 (figure 2.4.3).
3. Extrapolation of the upper 90% confidence interval of scenario 1 (figure 2.4.3).

In this chapter, these scenarios will be run and their results will be presented and discussed.

6.1 Scenario 1

As stated, scenario 1 consists of extrapolation of the current trend of habitat conversion in the Pantanal (figure 2.4.3). The model was run with the parameters that were determined through calibration in paragraph 6.3.

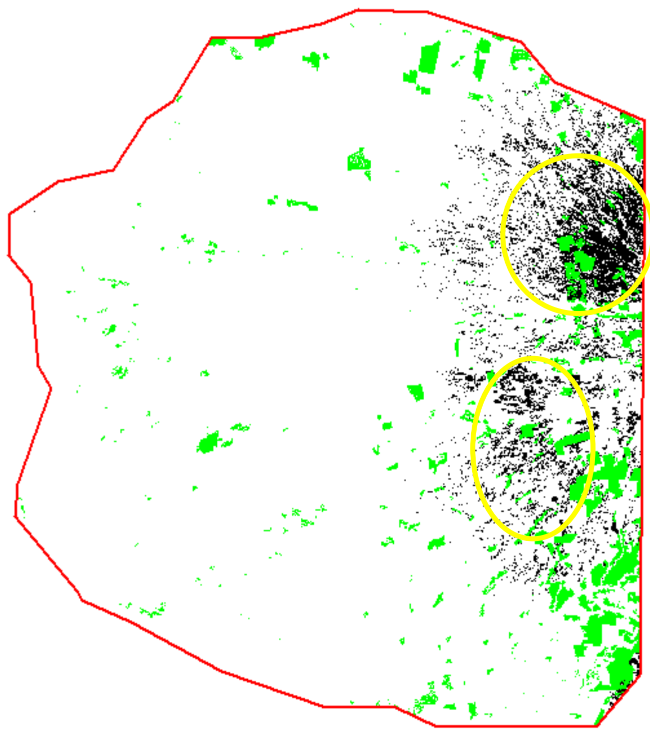


Figure 6.1.1. Output of scenario 1. Red indicates the outline of the study area. Black indicates the modelled increase in cultivated pasture in 2021. In addition, green indicates the real ("starting") situation in 2000. Yellow egg shapes indicate the regions where the increase in cultivation will be greatest.

Figure 6.1.1 presents the output of scenario 1. The greatest increase in cultivation is modelled in two regions in the central eastern part of the Pantanal (yellow). Figure 6.1.2 shows the ecotypes that will be affected by cultivation, and how many hectares of these ecotypes will be loss due to cultivation according to scenario 1. Clearly, two ecotypes suffer the largest loss: savannah arboreal and savannah gramíneo-lenhosa / arboreal.

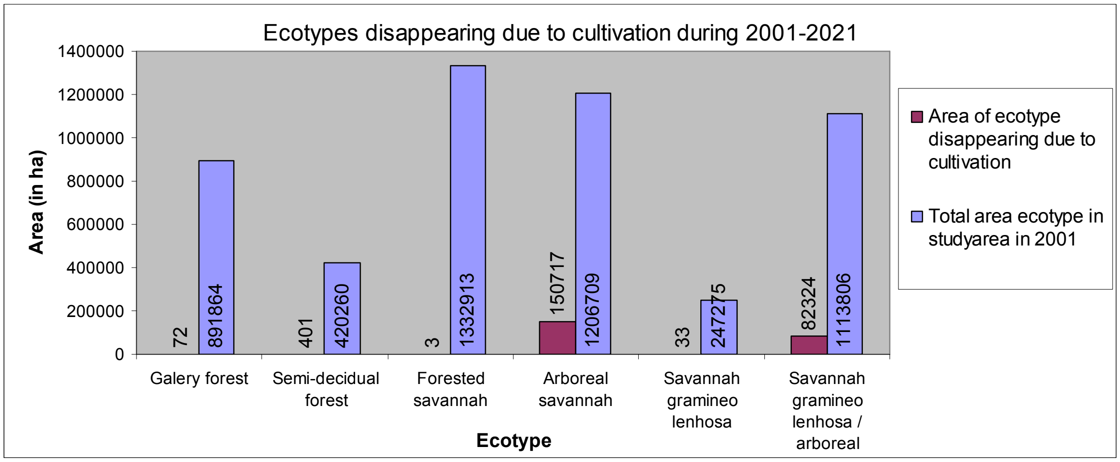


Figure 6.1.2. Area of ecotypes (in ha) that will be loss according to scenario 1 during 2001-2021.The total area of the ecotypes in 2001 (starting situation) are also given.

6.2 Scenario 2

As stated, scenario 2 consists of the lower 90% confidence interval of scenario 1 (figure 2.4.3). The model was run with the parameters that were determined through calibration in paragraph 6.3.

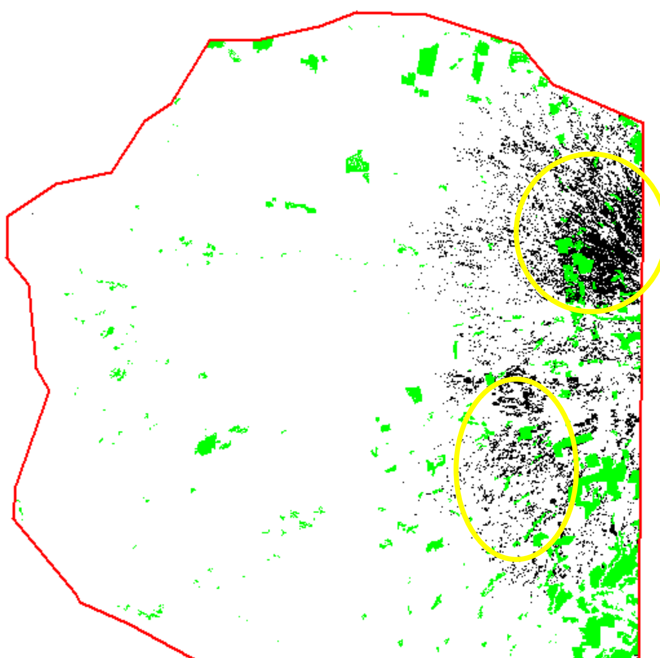


Figure 6.2.1. Output of scenario 2. Red indicates the outline of the study area. Black indicates the modelled increase in cultivated pasture in 2021. In addition, green indicates the real ("starting") situation in 2000. Yellow egg shapes indicate the regions where the increase in cultivation will be greatest.

Figure 6.2.1 presents the output of scenario 1. Again, the greatest increase in cultivation is modelled in two regions in the central eastern part of the Pantanal (yellow). However, looking closely at the output of scenario 2 revealed that it is practically identical to the output of scenario 1. Indeed, calculations show that the actual allocated area in scenario 2 is 0.44 % less than the allocated area in scenario 1. This difference should have been 10 %. Clearly, the allocation procedure failed to allocate 10 % less cultivated pasture than in scenario 1, which makes both outputs practically identical. Overlays show indeed that the differences between both outputs are minimal. Consequently, there is no use in further comparisons between the two outputs.

6.3 Scenario 3

As stated, scenario 3 consists of the upper 90% confidence interval of scenario 1 (figure 2.4.3). The model was run with the parameters that were determined through calibration in paragraph 6.3.

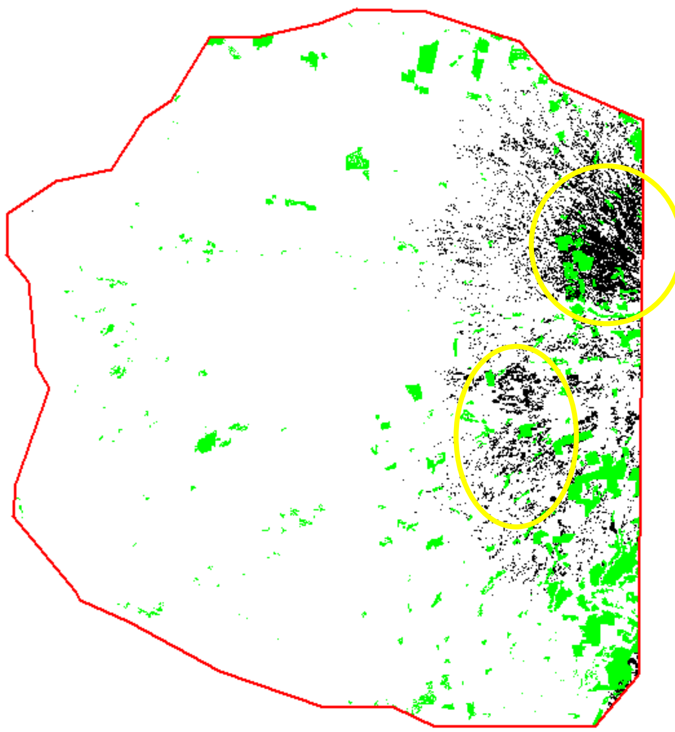


Figure 6.3.1. Output of scenario 3. Red indicates the outline of the study area. Black indicates the modelled increase in cultivated pasture in 2021. In addition, green indicates the real ("starting") situation in 2000. Yellow egg shapes indicate the regions where the increase in cultivation will be greatest.

Figure 6.2.1 presents the output of scenario 1. Again, the greatest increase in cultivation is modelled in two regions in the central eastern part of the Pantanal (yellow). However, looking closely at the output of scenario 2 revealed that it is also practically identical to the output of scenario 1. Again, calculations showed that the actual allocated area in scenario 3 is 0.57% less than the allocated area in scenario 1, whereas this difference should have been 10%. Clearly, the allocation procedure fails to allocate the demand with deviations less than 10%. Thus, both outputs are practically identical. Overlays show indeed that the differences between both outputs are minimal. Consequently, there is no use in further comparisons between the two outputs.

6.4 Conclusions

The output of scenarios 2 and 3 were practically identical to the output of scenario 1. This is strange, as the “demand” in both scenarios differed substantially from that in scenario 1. This was expected to result in a substantially different output, but this was not the case. It became clear that the model was not able to allocate the demand with a deviation less than 10%. As a result, the area actually allocated was more or less the same in all three outputs, which resulted in three identical outputs.

As stated, the model is not able to allocate demand with a deviation less than 10%. This problem in the allocation was also encountered during the test runs. During the test runs, the area of cultivated pasture allocated by the model was very inaccurate, and did not match the calculated “demand” by a long way. The reason for this was that differences in suitability were too small, and that consequently too many cells had the same suitability. The model selects cells based on their suitability, and consequently selected too many or too few cells. This made it impossible to allocate precisely the calculated demand. A number of improvements were made to “blow up” relative differences in suitability. The other test runs indicated that these improvements largely solved the problem, but apparently the allocation procedure is still not precise enough to allocate the demand with deviations smaller than 10%. However, this once again proves that the differences in suitability are very small in the Pantanal. Clearly, this problem affects the outcome of the model.

Because the three outputs are practically identical, a comparison between these outputs is not sensible. Therefore only the output of scenario 1 is now discussed.

According to scenario 1, the greatest increase in cultivation will occur in two regions in the central eastern part of the Pantanal. It is not illogical that a great increase in cultivation will occur in these two regions; the increase in cultivation in this area was already considerable in most recent years (1991-2000).

However, it is interesting to see that the model also allocated a limited increase in cultivation in the south-eastern and north-eastern parts of the Pantanal. During the calibration (over the period 1976-2000), the model also predicted limited cultivation in this area whereas in reality the increase was enormous. As stated, the model also allocated a limited increase in cultivation in these parts in the scenario study; this indicates that these areas are physic not extremely suitable for cultivation (i.e., most parameters in the model concern physic characteristics). However, in reality cultivation in this area was enormous over the period 1976-2000. Apparently, more factors played a role and made cultivation an attractive option in these areas.

7 Conclusions, Discussion and Follow-up

This chapter will look back at the research questions and objective, and discuss if they have been adequately addressed during this study.
In the introduction, the research questions were divided into three main groups:
1. Research questions related to land use change in the Pantanal.
2. Research questions related to land use change models (in general).
3. Research questions related to our model.

These three groups of research questions will be discussed separately in the next three paragraphs. In each paragraph, the results and main conclusions of all research questions within the group are repeated. In addition, these results and conclusions will be discussed, and a number of recommendations for future research regarding land use change in the Pantanal will be given.
The fourth and final paragraph of this chapter will discuss if the research objective has been adequately answered. In addition, this last paragraph will give an overall conclusion regarding this study and give suggestions for a future follow-up study.

7.1 Research questions related to land use change

Two research questions were related to land use change:
1. Which land use-trends are relevant in the Pantanal?
2. Which forces drive land use change in the Pantanal, and which indicators can adequately represent them in the model?

Which land use-trends are relevant in the Pantanal?
The most relevant land use change process in the Pantanal is the conversion of natural habitat into cultivated pasture. Historical land use maps show this clearly, but local experts (W. Tomas from EMBRAPA; C. Padovani from EMBRAPA; M. van Eupen from Alterra) and literature (Seidl 2000) confirmed this as well. During this study, a number of interesting aspects of this land use change process appeared:

- The increase in cultivated pasture is achieved either through (1) expansion of existing cultivated pastures or through (2) new, isolated cultivated pasture away from existing areas of cultivated pasture.
- Regional differences in the increase in cultivation exist; the increase in the eastern part of the Pantanal is much higher than the increase in the western part of the study area.
- Although overall the area of cultivated pasture increases, a small number of cultivated pastures disappear.

Some interesting comments can be made regarding these previous three observations.
First, the regional differences in the increase in cultivated pasture are very interesting. They indicate that particularly the eastern part of the Pantanal is especially vulnerable to cultivation. However, what makes the western part less attractive? Is this the lack of infrastructure or is this area physically less suitable for cultivation. It would be interesting to find this out, since it could have relevance for conservation: if the western part of the Pantanal is physically not very suitable for cultivation, all conservation measures should focus on the eastern part of the Pantanal that is vulnerable for cultivation.
Secondly, the disappearance of cultivated pastures is very remarkable. It would be interesting to find out what factors cause ranchers to abandon their cultivated pastures. It was already suggested that flooding might play a role, but other factors might also play a role. Knowing these factors would be useful in designing conservation measures, since they influence cultivation.

Which forces drive land use change in the Pantanal, and which indicators can adequately represent in the model?

Readily, it became clear that it would be difficult to say which driving forces play a role in the land use changes in the Pantanal. The problem is that good information regarding the economic incentives of ranchers is very scarce and that the motivations of ranchers to construct cultivated pasture are still unclear. However, it seems logical that socio-economic forces will play a role; globalization and intensification are influential forces in many parts of the world and are equally likely to have an influence in the Pantanal. Nonetheless, good information regarding the socio-economic situation of ranchers is lacking, and it is therefore impossible to draw any definite conclusions on this issue. Additional research should focus on determining the dominant socio-economic processes in the Pantanal and gathering data about these processes. More knowledge about these processes would be very helpful in modelling land use change in the Pantanal.

Because the forces that drive land use change are not known, it is also difficult to determine which indicators should be chosen to represent them in the model. Therefore, it was decided to consult local experts about which variables to include as indicators of land use change in our model. C. Padovani (EMBRAPA), W. Tomas (EMBRAPA), M. van Eupen (Alterra) and A. Seidl (Colorado State University) were consulted. After many discussions with the local experts, it was decided that the following variables were the main driving factors and thus should be included in the model:

- Neighbouring cultivation.
- Flooding risk.
- Elevation.
- Ecotype.
- Geomorphology.

However, looking back at this study, a number of comments need to be made regarding the chosen indicators:

One of the experts (C. Padovani) already doubted the relevance of the geomorphology as an indicator for cultivation. During the calibration, it became clear that the most realistic model would be achieved by leaving out geomorphology. Apparently, geomorphology is not a good indicator of cultivation. The other indicators seem to have been chosen well.

However, the model would probably benefit if more socio-economic indicators were included. During the calibration, it became clear that the model underestimated the increase in cultivation in the south-eastern and north-eastern parts of the Pantanal. It is difficult to say why the model underestimated the increase in these parts, but it is striking that in these regions the infrastructure is relatively good (they are good accessible by either road or river; pers. comm. M. van Eupen). This suggests that socio-economic factors play a role. Therefore, more research into the socio-economic factors that influence land use change seems necessary.

7.2 Research questions related to land use change models (in general)

The following research questions were related to land use change models (in general):

1. Which types of land use change models exist & which type should we use?
2. How does this type of model work (general description)?
3. Which “hot” topics are there regarding this type of model & should we include any in our model?
4. How have other studies implemented this type of model?

Which types of land use change models exist & which type should we use?
There are different ways to classify existing land use change models. However, this study chooses to follow the classification of Irwin and Geoghegan (2001).
This study used a non-economic model to simulate land use change in the Pantanal.
More specifically, a simulation model (i.e. cellular automata or CA) was used. The reason for this is:

- An estimation model is less adequate to incorporate neighbourhood effects into the model. This is an essential element in modelling habitat conversion in the Pantanal, as many factors influencing habitat conversion relate to neighbourhood effects (see paragraph 3.2).
- Estimation models and hybrid models use statistical analysis instead of expert knowledge to determine which parameters are included in the model. However, this study could make use of the expert knowledge that was already acquired during the previous phase of the Pantanal-Taquari-project (Jongman et al. 2005). Because of this readily available expert knowledge, the best approach seemed to use a simulation model.

As stated, this study used a simulation model because expert knowledge was readily available. However, a follow-up study should consider including more statistics in the model. The reason for this is that the process of cultivation in the Pantanal is still not entirely understood: especially the relation between socio-economic factors and cultivation is not yet clear. This was also demonstrated by the calibration of the model; the increase in cultivation was underestimated in those parts of the Pantanal that have the best infrastructure (i.e., access to market), and this indicates that the socio-economic aspect was underestimated. Therefore, follow-up research might want to collect more socio-economic data and try to reveal the relation between different socio-economic factors and cultivation. Of course, expert knowledge can also determine this relation, but there is still so little data regarding the socio-economic aspects of this land use change process. Therefore, it seems a good step to start gathering and producing data about it (e.g. wealth and education of ranchers, access to markets and labour, etc.); i.e., land use change research in the Pantanal might benefit from additional "facts and figures" regarding the socio-economic component.

How does this type of model work (general description)?
As stated, this study used a simulation model or CA (cellular automata) to model land use changes in the Pantanal. An extensive description of cellular automata was presented in chapter 3. This description made clear that an essential element of a CA is its neighbourhood function. Here follows a description of the neighbourhood function that was used in the Pantanal model with some additional comments.

In the Pantanal model, two neighbourhood functions were used. The reason for this is that two different processes were modelled that each required a neighbourhood.
The first neighbourhood was used to model the process in which existing cultivated pastures expand. This neighbourhood function consisted of a classical Moore neighbourhood. The reason for choosing this neighbourhood function was that analysis showed that cultivated pastures increase their radius on average with 100 m per year. With a resolution of 90m by 90 m, this approximates to an expansion of 1 cell per year in each direction; this equals a classical Moore neighbourhood. Thus, a classical Moore neighbourhood was chosen because it (theoretically) reflected the observed process best. However, it was not tested through calibration if it was indeed the best option. Therefore, it would be interesting to do a calibration and to see if this neighbourhood is indeed the best option to model the process with.
The second neighbourhood function was used to model the process in which new, isolated clusters of cultivated pasture appear away from existing areas of cultivated pasture. Since

these clusters have certain characteristics, only cells that meet these characteristics are selected. Clusters of selected cells are identified by means of a classical Von Neumann neighbourhood. Originally, a Moore neighbourhood was chosen because it includes all direct neighbours, which was assumed to be the most logical way to represent clusters of cells. However, the test runs showed that too many clusters were selected. As a consequence, a Von Neumann neighbourhood was chosen because a Von Neumann neighbourhood is stricter (does not consider cells that are diagonal neighbours) and this is thus another possibility to select fewer clusters of cells.
Other possibilities than either a Moore or Von Neumann neighbourhood were not possible with the software that was used. However, many more neighbourhood functions exist. Nevertheless, it is difficult to say which type of neighbourhood is ideal for modelling the process. A good analysis of the shape of these clusters will be necessary to determine this; i.e., analysis of the shape of these clusters should reveal which neighbourhood is best suited to model them.

Which "hot" topics are there regarding this type of model & should we include any in our model?
None "hot" topic could be included in the Pantanal model. The reasons for this are now discussed for each of these topics.

Multi-Agents Systems were not used because: (1) this would require advanced programming skills that the author does not posses and (2) high-quality data about individual ranchers that is not available.
At this point, I would like to make a suggestion for a future follow-up study; it should try to integrate the model with a Multi Agents System. I believe that the studied land use change process is pre-eminently suitable to model with a Multi Agents System because it is so strongly coupled to the decision-making process of ranchers, and that the model will be a more powerful management tool after integration with a Multi Agents System. To achieve this, the model should focus more on the decision making process of ranchers and incorporate data about individual farmers that influences the decision making of a rancher (e.g., wealth of a rancher). Of course, this would also require more knowledge and data of the socio-economic aspects of cultivation.

A stochastically constrained CA was not considered appropriate for the Pantanal-model. I believed that the observed land use changes are understood to a certain degree and that it was more appropriate to model the observed land use changes without a stochastic component. The calibration confirmed that the observed land use changes were sufficiently understood as the performance of the model was acceptable. Clearly, it was a good decision not to add a stochastic component after all.

Neural networks have a number of important drawbacks, which made them rather unsuitable for our land use change model:
- The training of a neural network requires a lot of data; a network learns by trial-and-error and this method naturally only works if enough data is available (Malczewski 2004). Unfortunately, such an amount of data is not available for our study.
- The development of a neural network requires extensive knowledge about programming (Malczewski 2004). Unfortunately, we do not have such knowledge.
- Neural networks are not transparent. They develop their own algorithms to model land use changes and are thus not user-defined; i.e. they resemble a "black box". This can be an important drawback as one of the main reasons for developing a land use change model is to make the consequences of certain land use trends more transparent. If a model is not

transparent as is the case for neural networks, the usefulness of the model is affected (Malczewski 2004).
The conclusion that it is not useful to integrate our land use change model with a neural network has not changed.

How have other studies implemented this type of model?
To answer this question, three land use change models were reviewed: CLUE-S, Environment Explorer and DINAMICA. The following issues were discussed:

- General structure
- Allocation model
- Implementation of allocation model:
    - Data structure
    - Cell size
    - Cell states
    - Driving factors
    - Weights of the driving factors
    - Neighbourhood function
    - Time period

An overview and discussion of the results can be found in table V in paragraph 3.3.5. At this point, I would like to limit the discussion to the way these results influenced this study.

First, these results influenced the general structure of this study; I found that all three reviewed models have two separate phases. In the first phase, the increase or decrease in each land use type is calculated: i.e. the total number of cells that need to change during each time step. The second phase allocates the calculated changes in each land use type over the grid. Because this approach is relatively straightforward and clear, and suited for the process that was modelled, land use change in the Pantanal was also modelled in two phases.
Moreover, the allocation procedure in all three models was more or less the same. In all three models a suitability analysis is carried out before the actual allocation. A first allocation is then made (based on the suitability analysis) and a number of the selected cells change. Then, the allocated amount of cells is compared with the "demand". If this does not match, a new allocation or "iteration" is made with slightly adapted parameters. The allocated amount of cells is again compared with the demand. If they do not match, another iteration is made. This iterative process is repeated until the allocated amount of cells matches the demand. Again, this approach is relatively straightforward and clear and it was therefore preferred over other approaches (e.g. approach used in Ruimtescanner (Schotten et al. 1997) or CLUE for national and continental level (Veldkamp and Fresco 1996)) and used in the Pantanal Land Use model.
Another important influence on this study was the approach used in the DINAMICA model. The DINAMICA-model discerned two sub processes in the process of deforestation: (1) deforestation progresses along a "deforestation front" of existing deforested areas and (2) isolated, new patches of deforestation emerge away from existing deforested areas. Both sub processes were separately modelled. Since similar sub processes could be discerned in the process of cultivation in the Pantanal, it seemed a good idea to also model these separately in the Pantanal model. This resulted in the Expander and Patcher functions described in chapter 4.

7.3 Research questions related to our model

Three research questions were related to our model:

1. What concept will I use to model land use change in the Pantanal?

2. How “robust” is our model (sensitivity)?
3. How reliable is our model (calibration)?

What concept will I use to model land use change in the Pantanal?
As stated, the most relevant land use change process in the Pantanal is the conversion of natural habitat into cultivated pasture. However, a number of interesting aspects of this land use change process appeared during this study:

1. Regional differences in the increase in cultivation exist; the increase in the eastern part of the Pantanal is much higher than the increase in the western part of the study area.
2. The increase in cultivated pasture is achieved either through (1) expansion of existing cultivated pastures (“expansion”) or through (2) new, isolated cultivated pasture away from existing areas of cultivated pasture (“patching”).
3. Although overall the area of cultivated pasture increases, a small number of cultivated pastures disappear.

All these aspects have been taken into account in the formalization of the model. Regarding the first two aspects, a number of quantitative analysis were done. Therefore, these aspects could be modelled based on detailed quantitative information. However, such quantitative analyses were not possible for the third aspect mentioned above. Thus, this aspect could only be modelled based on a number of assumptions.
An assumption was that these cultivated pastures disappear because they are so severely damaged by the flooding that they are consequently abandoned; according to Jongman et al. (2005) cultivated pastures are irreversibly damaged if they are flooded more than 6 months per year. Consequently, the only indicator of this process is assumed to be flood duration. However, the role of flooding is merely a speculation. The process might be more complex than this, and additional research should look into this.

- Another assumption was that these abandoned pastures are so heavily damaged that they can not be used as cultivated pasture any longer. Therefore, the model assumes that locations where cultivated pasture has disappeared are no longer suitable for cultivated pasture. However, little is known about this all; e.g., after which period become these locations suitable once again for cultivated pasture? Other interesting aspects regarding this process might still not be known.

Another important assumption is that the increase in cultivated pasture would mainly occur during the favourable dry season, whereas the disappearance of cultivated pastures was assumed to occur mainly during the unfavourable wet season because of (unexpected) flooding. Therefore, it was considered necessary to design two separate components; a sub model describing the increase in cultivated pasture in the dry season model and sub model describing the disappearance of some cultivated pastures in the wet season. However, cultivated pastures might also disappear during the dry season, as the exact reason for their disappearance is not yet known. In that case, all these processes (i.e., increase and disappearance of pastures) occur simultaneously and should not be modelled in two consecutively executed components, as is now the case.
Another important issue would be to find out how much cultivated pasture disappears per year. In the current model, the amount of cultivated pasture that disappears varies; in the model only those cultivated pastures disappear that have been constructed in areas that are flooded more than 6 months per year. However, there is not a predetermined demand that has to be allocated, as is the case in modelling the increase in cultivated pasture. Clearly, the process is not entirely understood and could be better modelled if more research would be done.

As stated, a number of quantitative analysis were done regarding the increase in cultivated pasture through the processes of expansion and patching as well as the regional differences in this increase. Therefore, these aspects could be modelled based on detailed quantitative information. Nevertheless, some assumptions also had to be made to model these processes. To model the regional differences in the increase in cultivation, it was assumed that the different parameters were the same for both regions; i.e., the only different parameter was the expected increase in cultivated pasture or "demand". However, it could very well be that some driving factors have a stronger influence on cultivation in the western region than in the eastern region (or vice versa). In that case, different parameters should be used for both regions. However, this study did not investigate this and assumed that the same parameters could be used for the regions.

To model the increase in cultivation through the processes of expansion and patching, some assumptions also had to be made.
A first assumption was that locations with ecotype "bare soil" are very unsuitable for cultivation. Although this ecotype is very unfertile and unsuitable for cultivated pasture (Jongman et al. 2005), it seems unlikely that no gradients exist; i.e., some "classes" of bare soil might be more suitable than others.
Another assumption was made in the neighbourhood function of Patcher. "Patcher" tries to simulate the process in which new, isolated "islands" or clusters of cultivated pasture suddenly appear. Analysis showed that these clusters had the following characteristics:

1. Most clusters are on average located higher than 110 cm.
2. Clusters are mainly covered with savannah.
3. Most clusters are on average flooded less than 2 months per year.
4. Clusters are never larger than 700 ha.

To simulate the process of "Patching", the neighbourhood function selects cells that meet characteristic 1 until 3, and looks if "clusters" of these selected cells exist and if their area is smaller than 700 ha. During the test runs, it already became clear that many cells could meet these criteria and consequently stricter criteria were chosen. Since so many cells could meet these criteria, it could very well be that these clusters are not cultivated because of these characteristics but because of other unknown factors. As stated before, the economic motivation of ranchers to construct cultivated pasture is not yet clear (Seidl 2000) and this makes it hard to determine which criteria ranchers use to determine if and where they construct cultivated pasture.

How "robust" is our model (sensitivity)?
There are some contradicting signals regarding the sensitivity of the model.
For instance, the results of the test run were worrisome. During the test run, the model did not allocate the correct amount of cultivated pasture due to the small differences in suitability; i.e., too many cells had the same suitability value and too many cells were selected simultaneously. In addition, the "Patcher"-function of the model could not select small clusters of extremely suitable cells but instead selected more than a third of the entire study area. Again, this was due to the small differences in suitability; i.e., the differences in suitability are so small that too many cells fell within the set criteria of "Patcher", and as a result "Patcher" selected too many cells simultaneously.
Also, a similar problem was encountered during the scenario study. During the scenario study, it became clear that the model was not able to allocate the demand with a deviation less than 10%. As a result, the area actually allocated in the three scenario studies was more or less the same, whereas it should have differed 10% between the scenarios. Apparently, the model was not able to allocate demand with a deviation less than 10%. This problem is similar to the problem encountered during the test runs; the differences in suitability are too small and too many cells have the same suitability value so that either too many or too few

cells are selected simultaneously. This problem can seriously affect the outcome of the model. Inevitably, a number of assumptions, parameters, etc. will deviate from reality, and their uncertainties are enhanced by these small differences in suitability. This can easily result in an unrealistic output.
However, the sensitivity analysis indicated that the model is not extremely sensitive to changes in parameters. All driving factors that were tested have had a comparable (small) effect on the output after changing their weights. This is a more positive indication regarding the sensitivity of the model.

Since only a limited number of the components have been tested in the sensitivity analysis, it is still difficult to make conclusions about the sensitivity of the model. Nevertheless, the results of the sensitivity analysis indicate that the model is not extremely sensitive, despite the small differences in suitability that were encountered during the test runs and scenario study.
A comment that can be made regarding the small differences in suitability is that mostly parameters were included that describe the physical characteristics of the area. Apparently, the differences in these physical characteristics are not very great. Including some relevant socio-economic factors might remedy this; these differences might be greater than in the physical environment.

How reliable is our model (calibration)?
During the sensitivity analysis, it already became that clear that the driving factors “elevation”, “geomorphology” and “ecotype” are the most sensitive parameters of the model. The model was calibrated by changing these three most sensitive parameters. To calibrate the model, the period 1976-2000 was modelled and the output was compared with real-world data. During the calibration, the value of one of the selected parameters was increased whilst the values of all other parameters were kept constant. Next, the fit of the output with reality was determined by calculating the percentage of overlapping cells with cultivated pasture that overlap between the output of the model and the land use map of 2000. The optimum value of each of the three parameters (elevation, geomorphology and ecotype) could be determined in this way.
Running the model with all the optimum values (which were determined in the way described above) resulted in a fit of 17 %. Although a fit of 17% seems a poor performance, such a conclusion is not entirely correct.
First, the input land use map of 1976 was not an existing land use map but was reconstructed using an ecological model; inevitably, errors will have been produced in creating this 1976 land use map.
Secondly, the Pantanal-model is intended as a management tool. Therefore, it is merely important that the model is able to predict broadly in which direction land use change will develop, and to look at the fit of the model at a larger scale. After looking at a larger scale, it became clear that the model allocates very correct in most regions. However, it also became clear that the model underestimated the increase in cultivation in the south-eastern and north-eastern parts of the Pantanal. It is difficult to say why the model underestimated the increase in these parts, but it is striking that in these regions the infrastructure is relatively good (they are good accessible by either road or river). This suggests that socio-economic factors play a greater role than expected beforehand.
Still, looking at all these aspects, I feel it is safe to say that the overall performance of the model is good and that it is good enough to be used in scenario-studies.

7.4 Follow-up

In the introductory chapter, the objective of this study was presented. The objective was formulated as follows:
This study aims to prototype and validate a simulation model for the Pantanal which can explore future land use based on potential land use developments; i.e. to prototype a land use change model that can be used for scenario-studies.
If the model can be prototyped and validated, we will check the functionality of the model and explore future land use in three scenarios. These three scenarios will be:

1 The current trend of habitat conversion in the Pantanal will be extrapolated.
2 Lower 90% confidence interval of scenario 1 will be extrapolated.
3 Upper 90% confidence interval of scenario 1 will be extrapolated.

This study has succeeded in prototyping and validating a model. However, running the scenarios proved to be a problem. During the scenario study, it became clear that the model was not able to allocate small differences in demand adequately. As a result, the outputs of all scenarios were similar. A similar problem was encountered during the test runs. All these problems were caused by the fact that the relative differences in suitability are very small; most parts of the Pantanal are equally suitable according to the chosen criteria, and this makes the allocation difficult (too many cell are selected simultaneously).
Most variables that have been included describe the physical characteristics of the Pantanal. The differences in suitability are not very great in the Pantanal according to these variables. If they have been chosen well, this means that in principle the largest part of the Pantanal is suitable for cultivation. Another option is that some important variables have been left out; including these might increase the relative differences in suitability. Suddenly it might become clearer why cultivation occurs in some regions and not in others. Since it seems logical that ranchers construct cultivated pastures out of economic motivations, I suggest that socio-economic variables might play a role. Thus, it seems useful to study further the socio-economic motivations of ranchers to construct cultivated pastures. Ideally, these results should be used to actually model the decision making process of ranchers in a Multi Agents System, which can be integrated with this simulation model.

References

Almeida, de C. M., A.M. V. Monteiro, G. Camara, B.S. Soares-Filho, G.C. Cerqueira, C.L. Pennachin, & M. Batty (2002). Empiricism and stochastics in cellular automaton modeling of urban land use dynamics. CASA Working Papers, no.42. Working paper. Centre for Advanced Spatial Analysis (UCL), London, UK. Available at URL: http://eprints.ucl.ac.uk/archive/00000252/01/Paper42.pdf

Angelsen, A. and D. Kaimowitz (1999). Rethinking the Causes of Deforestation: Lessons from Economic Models. The World Bank Observer 14(1): 73-98.

Balmann 2004. Modeling Land Use with Multiple-Agent Systems. Available at URL: http://www.zef.de/download/iaae_symp/balmann_paper.pdf

Bürgi, M., A.M. Hersperger & N. Schneeberger (2004). Driving forces of landscape change – current and new directions. Landscape Ecology 19: 857-868.

Campbell, M.J. and M.J. Gardner (1988). Calculating confidence intervals for some non-parametric analyses. British Medical Journal 296: 1454-1456.

Conservation International (2004). Available at URL: http://investigate.conservation.org/xp/IB/expeditions/pantanal/maps_facts.

Calheiros, D.F. (2004). Baia do Castelo: Local Community Participation in the Upper Paraguay River Basin, Pantanal Wetland, Mato grosso do Sul State. Case Studies on Local and Indigenous People's Involvement in Wetland Management. Commissioned for the project in response to Ramsar Recommendation 6.3. Available at URL: http://ramsar.org/cop7/cop7_doc_18.1_add.htm.

Deadman, P.J. (1999). Modelling individual behaviour and group performance in an intelligent agent-based simulation of the tragedy of the commons. Journal of Environmental Management 56: 159-172.

Dolabella, A.L.L.B. (2000). The Brazilian Pantanal: An Overview. Available at: URL http://www.pantanal.org/book.htm.

Geist, H.J. and E.F. Lambin (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. BioScience 52(2): 143-150.

Groothuysen, W., T. de Nijs, J. van Vught, C. van Vien, T. Trijssenaar & H. van Laar (2001). INterActie met provinciale ruimte; Rapportage provinciale proefprojecten LeefOmgevingsVerkenner. RIVM rapport 408505005. Available at URL: http://www.rivm.nl/bibliotheek/rapporten/408505005.pdf

Hazen, B.C. and M.W. Berry (1997). The simulation of land-cover change using a distributed computing environment. Simulation practice and Theory 5: 489-514.

Irwin, E.G. and J. Geoghegan (2001). Theory, data, methods: developing spatially explicit economic models of land use change. Agriculture, Ecosystems & Environment 85 (1-3): 7-24.

Jongman, R.H.G. (eds) (2005). Pantanal, Tools for decision making in Integrated Water Management. Partners for Water project 02.045. Water for Food and Ecosystems. Alterra Special Publication 2005/02. ISSN 1574-8227.

Jørgensen, S.E. (1986). Fundamentals of Ecological Modelling. Developments in Environmental Modelling 9. Elsevier Science Publishers B.V., Amsterdan, The Netherlands.

Lambin, E.F., B.L. Turner, H.J. Geist, S.B. Agbola, A. Angelsen, J.W. Bruce, O.T. Coomes, R. Dirzo, G. Fischer, C. Folke, P.S. George, K. Homewood, J. Imbernon, R. Leemans, X. Li, E.F. Moran, M. Mortimore, P.S. Ramakrishnan, J.F. Richards, H. Skånes, W. Steffen, G.D. Stone, U. Svedin, T.A. Veldkamp, C. Vogel and J.

Xu (2001). The causes of land-use and land-cover change: moving beyond myths. Global Environmental Change 11: 261-269.
Ligtenberg, A., A.K. Bregt & R. van Lammeren (2001). Multi-actor-based land use modelling: spatial planning using agents. Landscape and Urban Planning 56 (1-2): 21-33.
Malczewski, J. (2004). GIS-based land use suitability: a critical overview. Progress in Planning 62 (1): 3-65.
Mas, J.F., H. Puig, J.L. Palacio & A. Sosa-Lopez (2004). Modelling deforestation using GIS and artificial neural networks. Environmental Modelling & Software 19: 461-471.
Nijs, de T.C.M., R. de Niet & L. Crommentuijn (2004). Constructing land-use maps of the Netherlands in 2030. Journal of Environmental Management 72 (1-2): 35-42.
Nijs, de T., G. Engelen, R. White, H. van Delden & I. Uljee (2001). De LeefOmgevingsVerkenner; technische documentatie. RIVM rapport 408505 007/2001. Available at URL: http://www.rivm.nl/bibliotheek/rapporten/408505007.pdf
Schatten, A. (2004). Cellular Automata Tutorial. Available at URL: http://www.schatten.info/info/ca/ca.html
Schotten C.G.J., R.J. van de Velde, H.J. Scholten, W.T. Boersma, M. Hilferink, M. Ransijn, and R. Zut (1997). De Ruimtescanner, geintegreerd ruimtelijk informatiesysteem voor de simulatie van toekomstig ruimtegebruik. RIVM rapport 711901002, 32 pp.
Seidl, A. (2000). Use of Incentive Measures for Conservation and Sustainable Use of Biodiversity. Study from UNEP (2000). Available at: URL: http://www.biodiv.org/doc/case-studies/default.aspx.
Seidl, A., J. dos Santos Vila de Silvae & A. Steffens Moraes (2001). Cattle ranching and deforestation in the Brazilian Pantanal. Ecological Economics 36 (3): 413-425.
Soares-Filho, B., A. Alencar, D. Nepstad, G. Cerqueira, M. del Carmen, V. Diaz, S. Rivero, L. Solórzano & E. Voll (2004). Simulating the response of land-cover changes to road paving and governance along a major Amazon highway: the Santarém–Cuiabá corridor. Global Change Biology 10 (5): 745-764.
Soares-Filho, B.S., G.C. Cerqueira & C.L. Pennachin (2002). DINAMICA-a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological Modelling 154: 217-235.
Soepboer, W. (2001). Clue-S, The Conversion of Land Use and its Effects at Small Regional Extent; An application for Sibuyan Island, The Philippines. MSC-thesis, Laboratory of Soil Science and Geology, Environmental Sciences, Wageningen University.
Swarts, F.A. (2000). The Pantanal in the 21st Century: For the Planet's Largest Wetland, an Uncertain Future. Available at: URL: http://www.pantanal.org/book.htm.
Takeyama, M. and H. Couclelis (1997). Map dynamics: integrating cellular automata and GIS through Geo-Algebra. International Journal of Geographical Information Science 11(1): 73-91.
Torrens, M. (2000). How cellular models of urban systems work. Centre for advanced spatial analysis. University College, London, UK. Available at URL: http://www.casa.ucl.ac.uk/how_ca_work.pdf.
Veldkamp, A. and P.H. Verburg (2004). Methodology for simulating the spatial dynamics of land use change in forest fringes.
Veldkamp, A., P.H. Verburg, K. Kok, G.H.J. de Koning, J. Priess and A.R. Bergsma (2001). The need for scale sensitive approaches in spatially explicit land use change modeling. Environmental Modeling and Assessment 6: 111- 121.

Landscape Level Analysis of the Spatial and Temporal Complexity of Land-Use Change
Veldkamp, A. and E.F. Lambin (2001). Predicting land-use change. Agriculture, Ecosystems and Environment 85 (1-3): 1-6.
Veldkamp, A. (1999). Predicting land use change: Current progress in spatially explicit modelling of land use/cover change (LUCC). Presentation at workshop in Wageningen 1999. Available at URL: http://www.sciconf.igbp.kva.se/Files/Presentations/B2.2_Veldkamp.PPT#267,1, Predicting land use change: Current progress in spatially explicit modelling of land use/cover change (LUCC)
Veldkamp, A. and L.O. Fresco (1996). CLUE: a conceptual model to study the Conversion of Land Use and its effects. Ecological Modelling 85: 253-270.
Verburg[1], P.H., P.P. Schot, M.J. Dijst, and A. Veldkamp (2004). Land use change modelling: current practice and research priorities. GeoJournal 61: 309-324.
Verburg[2], P.H., A. Veldkamp, L.Willemen, K.P. Overmars and J.C. Castella (2004). Landscape Level Analysis of the Spatial and Temporal Complexity of Land-Use Change. In: Ecosystems and Land Use. DeFries, R., Asner, G. and Houghton, R.A. (eds.). Geophysical Monograph Series, Volume 153, AGU, 344 pages. (2004). Pp: 217-230.
Verburg, P.H., W.Soepboer, A. Veldkamp, R.L.V. Espaldon and S.S.A. Mastura (2002). Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model. Environmental Management 30 (3): 391-405.
White, R. and G. Engelen (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land use patterns. Environment and Planning A 25 (8): 1175-1199.
WNF (2004). Available at URL: http://www.wnf.nl/wnf/website/index.cfm/id=3609981A-B66F-4E74-8293C9BFD0E84861.
Wu, F. (2002). Calibration of stochastic cellular automata: the application to rural-urban land conversion. International Journal of Information Science 16 (8): 795-818.

Personal communication with:
- C. Padovani (EMBRAPA).
- M. van Eupen (Alterra).
- A. Ligtenberg (Wageningen University).
- R. van Lammmeren (Wageningen University).
- M. Hilferink (Object Vision).
- A. Seidl (Colorado State University).

Appendix

In this section, the data, scripts, etc. on the accompanying disc are described. All necessary information is on this disc to run effectively run the Pantanal land use change model. There are three directories on the disc: arisflow_scripts, data, and workspace. Most convenient is to save all three directories in a folder called C:\TEMP\fritz. By doing this, it is not necessary to change references in the arisflow-model.

Arisflow_scripts

This directory contains the definition files that are needed to run Arisflow/Arc_Info.

Data

This directory contains the following information:

Name: dem_new
Type: grid
Location: D:\data\dem_new
Description: This dataset is the original digital elevation model of the Pantanal.

Name: dem_100
Type: grid
Location: D:\data\dem_100
Description: This dataset is the digital elevation model of the Pantanal multiplied by a factor 100. It is used to calculate the suitability index.

Name: eup_inun12
Type: grid
Location: D:\data\eup_inun12
Description: This dataset indicates the flood duration of each location. The value indicates the number of months a location is flooded per year.

Name: learngrid_0
Type: grid
Location: D:\data\learngrid_0
Description: This dataset indicates which cells were cultivated pasture but are abandoned according to the model.

Name: pantanal_00
Type: grid
Location: D:\data\pantanal_00
Description: This dataset is the input ecotope map for the model. It is the ecotope map of the Pantanal of the year 2000. Among other things, it is used to calculate the suitability index.

Name: soil
Type: grid
Location: D:\data\soil
Description: This dataset is the geomorphological map of the Pantanal. It is used to calculate the suitability index.

Name: twi

Type: grid
Location: D:\data\twi
Description: This dataset gives the topographical wetness index for the Pantanal.

Name: studyarea
Type: shapefile
Location: D:\data\studyarea
Description: This dataset gives the outline of the study area.

Name: studyarea2
Type: shapefile
Location: D:\data\studyarea2
Description: This dataset gives the outline of the two different regions in the study area. The model models both regions separately.

Name: demand
Type: dbf
Location: D:\data\demand
Description: This dataset gives the yearly increase in cultivated pasture (i.e., scenario 1). This yearly increase in cultivated pasture is derived by extrapolating the historical increase in cultivated pasture. The model models cultivation in the two regions differently, and within each region discerns between cultivation achieved through "Expansion" or "Patching". As a consequence, there are four columns in this dbf-file: (1) yearly increase through "Expansion" in western region, (2) yearly increase through "Patching" in western region, (3) yearly increase through "Expansion" in eastern region, (4) yearly increase through "Patching" in eastern region.

Name: demand_scenario2
Type: dbf
Location: D:data\demand_scenario2
Description: This dataset is comparable to the dataset described above (demand.dbf). The difference is that this dataset gives the yearly increase in cultivated pasture according to scenario 2.

Name: demand_scenario3
Type: dbf
Location: D:data\demand_scenario3
Description: This dataset is comparable to the dataset described above (demand.dbf). The difference is that this dataset gives the yearly increase in cultivated pasture according to scenario 3.

Name: demand_rain
Type: dbf
Location: D:\data\demand_rain
Description: This dataset is not needed by the model and is redundant. However, it can not be deleted.

Name: draw_reclass
Type: dbf
Location: D:\data\draw_reclass
Description: This dataset is used in drawing the output ecotope map of the model. It specifies which colour each ecotope type will get.

Name: parameter
Type: dbf
Location: D:\data\parameter
Description: This dataset contains the different parameters that are used throughout the model. It contains the following parameters:

Vegetation	This parameter gives the weight assigned to vegetation (i.e., ecotope) in the suitability analysis.
Dem	This parameter gives the weight assigned to elevation in the suitability analysis.
Soil	This parameter gives the weight assigned to geomorphology in the suitability analysis.
Flood	This parameter gives the weight assigned to flood duration in the suitability analysis.
Cultivated	This parameter gives the weight assigned to cells that neighbour existing cultivated pasture.
Cluster	This parameter gives the weight assigned to clusters of extremely suitable cells.
Ex_area1	This parameter indicates how many cells around cultivated pasture will get a higher suitability value. This parameter is used for the western region of our study area.
Ex_area2	This parameter indicates how many cells around cultivated pasture will get a higher suitability value. This parameter is used for the eastern region of our study area.
Min_area1	This parameter indicates the minimum area of clusters of extremely suitable cells. This parameter is only used for the western region of our study area.
Min_area2	This parameter indicates the minimum area of clusters of extremely suitable cells. This parameter is only used for the eastern region of our study area.
Max_area1	This parameter indicates the maximum area of clusters of extremely suitable cells. This parameter is only used for the western region of our study area.
Max_area2	This parameter indicates the maximum area of clusters of extremely suitable cells. This parameter is only used for the eastern region of our study area.

Name: ecotopen_reclass
Type: text
Location: D:\data\ecotopen_reclass
Description: This text file is used to reclassify the ecotope map (pantanal_00), according to the suitability of each ecotope type for cultivation.

Name: flood_reclass_inun12
Type: text
Location: D:\data\flood_reclass_inun12
Description: This text file is used to reclassify the flood duration map (eup_inun12).

Name: soil_reclass
Type: text
Location: D:\data\soil_reclass
Description: This text file is used to reclassify the geomorphological map (soil), according to the suitability of each geomorphological pattern for cultivation.

Name: allocation_rain
Type: aml

Location: D:\data\allocation_rain
Description: This aml file is used to select cells with cultivated pasture which are extremely long flooded each year. These cells are changed into land use type “bare soil”.

Name: allocation_rain2
Type: aml
Location: D:\data\allocation_rain2
Description: This aml file is not needed by the model and is redundant. However, it can not be deleted.

Name: cluster2
Type: aml
Location: D:\data\cluster2
Description: This aml file is used by model_fritz_1 to select clusters of extremely suitable cells and to increase the suitability of the selected cells.

Name: draw_aml
Type: aml
Location: D:\data\draw_aml
Description: This aml file contains the necessary scripts to draw ecotope maps. It is referenced by model_fritz_1 and model_fritz_2.

Name: neighborhood
Type: aml
Location: D:\data\neighborhood
Description: This aml file is used by model_fritz_1 to determine which select existing cultivated pastures and to increase the suitability of the cells that neighbour these existing cultivated pastures.

Name: re_allocation
Type: aml
Location: D:\data\re_allocation
Description: This aml file is used to check if cells have been allocated twice (i.e., both by the Expander-function and the Patcher-function of model_fritz_1). If cells have been allocated twice, this aml allocates a new area of cultivated pasture that matches the overlap. This makes sure that per year the allocated area of cultivated pasture matches the scenario.

Workspace

This directory is an ArcInfo workspace. In addition, it contains the following models:

Name: geo_processing
Type: Arisflow flowchart
Location: D:\workspace\geo_processing
Description: This model executes the necessary geo-processing of the raw data; i.e., the raw data sets described earlier need to be processed first to run the models described below.

Name: model_fritz_1
Type: Arisflow flowchart
Location: D:\workspace\model_fritz_1
Description: This model models the increase in cultivated pasture through the processes of (1) Expansion and the (2) Patching. In this report, this model is referred to as the dry season model.

Name: model_fritz_2
Type: Arisflow flowchart
Location: D:\workspace\model_fritz_2
Description: This model models the disappearance of cultivated pasture. In this report, this model is referred to as the wet season model.

Name: pantanal_lu_model
Type: Arisflow-Commander script
Location: D:\workspace\pantanal_lu_model
Description: This script controls "model_fritz_1" and "model_fritz_2"; i.e., this script makes that these models are run in the right order, the right number of times, etc.

Name: comments_pantanal_lu_model
Type: text
Location: D:\workspace\comments_pantanal_lu_model
Description: This text file is used by the Arisflow-Commander script to write certain error messages to.