



Universiteit
Leiden
The Netherlands

Linking processes and pattern of land use change

Overmars, K.P.

Citation

Overmars, K. P. (2006, June 19). *Linking processes and pattern of land use change*. Retrieved from <https://hdl.handle.net/1887/4470>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4470>

Note: To cite this publication please use the final published version (if applicable).



5

Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land use model

Abstract

In this chapter, two research approaches to specify the relation between land use types and their explanatory factors are applied to the same modelling framework. The two approaches are used to construct land suitability maps, which are used as inputs in two model applications. The first is an inductive approach that uses regression analysis. The second applies a theoretical, actor decision framework to derive relations deductively using detailed field data. Broadly speaking, this classification coincides with the distinction between empirical and theoretical models and the distinction between deriving process from pattern and pattern from process. The two modelling approaches are illustrated by a scenario analysis for a case study in a municipality in the Philippines. Goodness-of-fit of the deductive approach in predicting current land use is slightly lower compared to the inductive approach. Resulting land use projections from the modelling exercise for the two approaches differ in 15 percent of the cells, which is caused by differences in the specification of the suitability maps. The chapter discusses the assumptions underlying the two approaches as well as the implications for the applicability of the models in policy-oriented research. The deductive approach describes processes explicitly and can therefore better handle discontinuities in land use processes. This approach allows the user to evaluate a wide range of scenarios, which can also include new land use types. The inductive approach is easily reproducible by others but cannot guarantee causality. Therefore, the inductive approach is less suitable to handle discontinuities or additional land use types, but is well able to rapidly identify hotspots of land use change. It is concluded that both approaches have their advantages and drawbacks for different purposes. Generally speaking, the inductive approach is applicable in situations with relatively small land use changes, without introduction of new land use types, whereas the deductive approach is more flexible. The choice of modelling approach should therefore be based on the research and policy questions for which it is used.

Based on: Overmars, K.P., Verburg, P.H., Veldkamp, A. 2006. Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land use model. Land Use Policy (Accepted).

5.1 Introduction

Within LUCC (land use and land cover change) research much attention has been paid to the development of models (Briassoulis, 2000; Veldkamp and Lambin, 2001; Parker *et al.*, 2003). Land use models are used as a tool to combine different aspects of the complex land use system and therefore enable researchers to study the dynamics of this system. Furthermore, land use change models are applied to evaluate scenarios to inform policy makers (Brown *et al.*, 2004; Solecki and Oliveri, 2004).

In reviewing land use models many criteria have been used to classify models: for example, whether a model is economic or non-economic, spatially explicit or not or whether the model is statistical/empirical, mathematical or rule-based (Briassoulis, 2000; Brown *et al.*, 2004; Verburg *et al.*, 2004d). Most of the current land use models have in common that they all try to combine human and natural processes, which implies the involvement of various disciplines (Couclelis, 2001). In this chapter we will use the broad distinction between deductive and inductive approaches of modelling (e.g. Laney, 2004; Overmars *et al.*, 2006 (Chapter 3)). Broadly speaking, this classification coincides with the distinction between theoretical and empirical models and the distinction between deriving patterns from process and process from patterns.

Overmars *et al.* (2006) (Chapter 3) identify six types of modelling, which vary from completely deductive to completely inductive. In this study two of these types will be used to specify the relation between land use and its explanatory factors, which will be implemented in two applications of a spatially explicit land use model in the same region. The first approach can be classified as 'unstructured factors induction'. In this approach a conceptual framework is used to define the dependent variable and the independent variables but then leave it to the procedures of statistical inference to find correlations between these variables. Theories are used to construct hypotheses about the relation between land use and its explanatory factors, but the structure of these theories is not used or tested (see Serneels and Lambin, 2001; Nelson *et al.*, 2004). The second, more deductive approach used in this chapter is called 'imposed theory'. In this approach a land use theory is specified for a real world case in terms of both structure and parameters, without any fitting to empirical data, and used to predict land use.

The two approaches to quantify the relation between driving factors and land use, resulting in a land 'suitability' estimate, will be implemented in two applications of CLUE-S, which is a dynamic land use model, to simulate scenarios of LUCC in a study area in the municipality of San Mariano in the northern part of the Philippines. The remainder of the model setting will be kept the same for the two modelling approaches to be able purely assess the effect of having different methods to specify land suitability.

The aim of this chapter is to compare the differences between the two model applications which have different specifications of land suitability as input. The difference in outcome of two model applications as well as the different assumptions underlying the two model specifications will be discussed. Furthermore, the chapter describes the implications for applicability of the approaches for different research and policy questions.

5.2 Study area and data collection

5.2.1 Study area

The study area is situated in Cagayan Valley in the northeastern part of the island Luzon in the Philippines (Figure 5.1). The study area includes 16 *barangays* (villages) in the municipality of San Mariano, in the province of Isabela, and its size is approximately 26,000 ha. It is situated between the town of San Mariano in the west and the forested mountains of the Sierra Madre mountain range in the east. The area is inhabited by approximately 16,500 people (about 3,150 households) of various ethnic groups, among whom the Ilocano, Ibanag and Ifugao, who are migrants or descendents of migrants that came to the area from the 1900s onwards, and the Kalinga and Agta, who are the indigenous inhabitants of the area. At present, the study area shows a clear land use gradient ranging from intensive agriculture, with mainly rice and yellow corn, near San Mariano to a scattered pattern of yellow corn, banana, grasses and trees to residual and primary forest in the eastern part of the study area. Before immigration started the area was completely covered with tropical lowland forest. About 76 percent of the population has farming as their main source of livelihood and another 12 percent is involved in working on other people's farms.

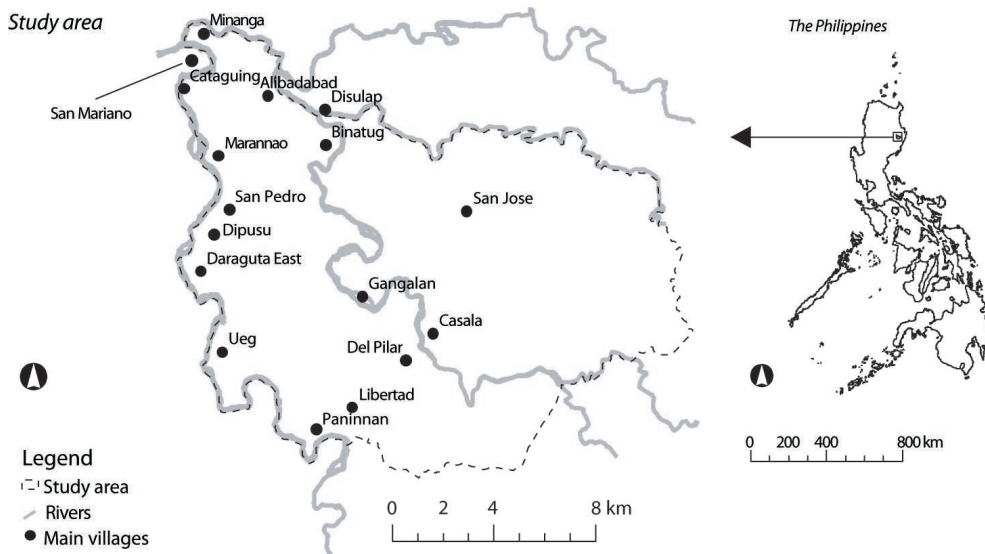


Figure 5.1: Study area

5.2.2 Data

Land use data were interpreted from two remote sensing images, a Landsat ETM+ image (<http://www Landsat.org>) from June 2001 and an ASTER image from March 2002. First, unsupervised classifications of both images were made for the study area. Second, the classes of the unsupervised classifications were recoded into land use types according to a set of 96 field observations. Finally, the land use map was constructed by combining the

classifications of the two images. In this procedure the ASTER image was first resampled from 15m resolution to the same grid as the Landsat image (30 by 30 m). Then, all land use classes of the two images were put in separate layers. In a GIS (Geographical Information System) these layers were combined, using overlay, where the delineation of the top layer overrules those of the layers underneath, in order to obtain the best fit with the field observations. This way the best elements of two images were combined and the best overall land use classification was created. To improve the identification of wet rice fields an extra satellite image from July 2001 was used. Finally, the image was resampled (aggregated) to a 50 by 50 m grid that coincides with the other data. Resampling was performed by taking the value of the original map under the centre point of the newly created grid. Classification accuracy of the land use map is 68 percent, which was calculated using an independent sample of 76 field observations (Verburg *et al.* 2004a).

In creating the land use map from remote sensing images banana plantations and low-density forest types (secondary forest) were difficult to separate from each other, because banana cultivation is quite extensive and often many trees grow in between the banana plants, which results in a similar spectral reflection as secondary forest. Therefore, a class that included both banana and secondary forest was manually divided based on field observations into a part with predominantly banana and a part with predominantly secondary forest. The western half of the area was identified as an area in which this class can be considered to contain almost exclusively extensive banana plantations. In the eastern part the same class is considered to be predominantly secondary forest. The resulting land use map is depicted in Figure 5.5A

The set of explanatory variables is based on a previous analysis (Overmars *et al.*, 2006 (Chapter 3)) and includes slope, ethnicity variables, accessibility variables, potential for rice and a reforestation policy. The slope map was derived from a 1:50,000 topographic map of the area (NAMRIA, unknown). This slope map was reclassified into five slope classes that correspond with classes in the survey held amongst farmers in the area, which was used in the deductive approach. It was not possible to obtain a map that depicts the ethnicity of the individual landowners, because no data were available that link all land managers to their individual parcels. Instead, maps of the percentage of every tribe per village were created based on census information of the National Statistics Office. The accessibility measures in this study are based on an in-depth study on accessibility in the study area (Verburg *et al.*, 2004a). The time farmers have to travel from their homes to the fields is calculated with a cost distance algorithm. In this calculation different travel speeds were attributed to different types of roads and off road and these were used to calculate minimum travel time. Transportation costs are calculated by assigning the transportation costs (to the market place in San Mariano) of the nearest village, based on the travel time calculation, to all locations. A map with the possibilities for irrigation to cultivate wet rice was constructed from a map indicating the area within 200 m distance to a creek (excluding big rivers) and a map indicating the land that can potentially be served by a NIA (National Irrigation Agency) project that was established in the area. These rules were combined to a map containing location with and locations without the possibility of cultivating rice. The final data source is a map delineating an area which is targeted by a policy called SIFMA (Socialized Industrial Forest Management Agreement), which promotes the planting of trees (DENR-CENRO, 1998). Within this policy farmers were offered 25 years of tenure rights on the condition that they plant a certain area with (fruit) trees (mainly mango, citrus and coconut). In the study area this policy was especially promoted by an NGO

Comparison of a deductive and an inductive approach

that provided free seedlings and assisted the farmers in obtaining the tenure documents (General, 1999)

Table 5.1: Descriptive statistics based on a 5 % sample of the complete dataset, $n = 5002$

Variable name	Description	Min.	Max.	Mean	St. dev.
<i>Land use variables</i>					
Wet rice	1 if wet rice, 0 otherwise	0	1	0.02	
Yellow corn	1 if yellow corn (and 10% other arable crops), 0 otherwise	0	1	0.21	
Banana	1 if banana, 0 otherwise	0	1	0.17	
Grass	1 if grass 0 otherwise	0	1	0.30	
Sec. forest	1 if secondary forest 0 otherwise	0	1	0.17	
Forest	1 if forest 0 otherwise	0	1	0.09	
Water bodies	1 if lake or river, 0 otherwise	0	1	0.02	
<i>Explanatory variables</i>					
Slope1	1 if slope < 2.5 degrees	0	1	0.22	
Slope2	1 if $2.5 \leq$ slope < 6.5 degrees	0	1	0.25	
Slope3	1 if $6.5 \leq$ slope < 12.5 degrees	0	1	0.34	
Slope4	1 if $12.5 \leq$ slope < 20.5 degrees	0	1	0.16	
Slope5	1 if slope \geq 20.5 degrees	0	1	0.04	
Dist. to small river	1 if distance to a small river < 200 m or part of NIA irrigation project	0	1	0.26	
Plot distance	Minutes walking to the plot (min.)	0	405	76.63	73.02
Transportation cost	Cost to transport a bag of corn from the house to San Mariano (pesos)	0	45	25.50	10.18
Ethnicity Ilocano	% in the barangay that is Ilocano	2.09	96.6	67.95	22.53
Ethnicity Ifugao	% in the barangay that is Ifugao	0	40.42	7.07	12.72
Ethnicity Ibanag	% in the barangay that is Ibanag	0	89.47	13.06	16.67

5.3 Methods

5.3.1 Overview

In this section the inductive and deductive approach to derive the relation between land use and its explanatory factors *i.e.* 'suitability' maps) are presented. The inductive mode using logistic regression analysis, is rather straightforward. The deductive approach use an actor decision framework. This approach is less known than the inductive approach and is therefore described in more detail. Both approaches make use of the dataset described in Table 5.1. So, differences between the results of the two approaches cannot arise from differences in the specification of variables. However, as will be explained, the two approaches differ in their model specification, for example, they use a different selection of variables from this dataset and different model parameters. Moreover, the deductive approach additionally includes variables that are constant over the area (e.g. prices and investments levels). This will result in different outcomes of the two approaches. The resulting suitability maps of the two approaches are input to two different applicat

of CLUE-S. CLUE-S is a spatially explicit and dynamic land use model, which is described below. The suitability maps produced with either the inductive or the deductive approach provides only one of the mechanisms that are responsible for land use distribution in the CLUE-S model. The other mechanisms and their inputs are modelled the same in both model applications.

Finally, we describe two scenarios that are used to illustrate the two modelling applications. One scenario is used for both model applications to compare differences. A second scenario introduces a new land use type and is only applied in the model with the deductive approach.

5.3.2 Approaches to determine the relation between land use and its explanatory factors

Inductive approach

In the inductive approach the suitability of a location for a land use type is determined in an empirical way by using logistic regression analysis. This regression model describes the relation between the occurrence of a land use type and the set of explanatory variables (Table 5.1) that are considered to influence land use allocation. The current land use is assumed to reflect the influence that these explanatory variables have exerted on the land use.

The dependent variables in the analysis are binary maps where the land use type under study has a value 1 and all other land use types have value 0. The variables that were inserted in the regression models were selected with a forward stepwise regression procedure (with probability levels of 0.01 for entry in the model and 0.02 for removal from the model). Originally, the data stems from fewer observations than a representation as grid would suggest and all cells would be considered to be observations. Therefore, a five percent sample was drawn from the original dataset of 99,863 cells to reduce spatial autocorrelation in the analysis. Sampling from a grid is a commonly used method in analysing land use patterns and will minimise spatial autocorrelation to a level that it will not affect the results (Serneels and Lambin, 2001; Stollé *et al.*, 2003). Based on the logistic regression analysis the probability of finding the land use type at each location can be determined. These probabilities are assumed to indicate the relative suitability of that location.

Deductive approach

Action-in-Context (AiC) (De Groot, 1992) is a methodology for problem-oriented research that puts activities of actors, for example land use, into context to gain insight in the cause of the activities. Based on Vayda (1983), the research sequence of the AiC methodology is to start with the actions under study, to identify the decision-making social entities directly behind these actions, and then to study the range of options available to the actors and their motivations attached to these (Verburg *et al.*, 2003). One of the elements of the AiC methodology is the 'deeper analysis', which ties the options and motivations of the primary actor to underlying cultural and structural factors. The structural framework of the deeper analysis will be used as an actor-model to study the decision-making process of farmers who are the primary land managers in the study area.

The structural framework of the deeper analysis is depicted in Figure 5.2 (De Groot, 1992) where the arrows show the direction of the causal relations. The first layer in Figure 5.2

consists of three elements: effect, action and actor. In the case of land use, effects can be degradation, biodiversity loss and greenhouse gas emissions, for example. Though, in the study the land use actions rather than the effects are the subject of study. The actors are social entities that exercise a significant decision-making capacity on the activity. In this study the actors are farmers. An example of the relations in the first layer is a farmer (actor) who grows corn (activity), leading to soil degradation (effect).

The second layer consists of 'implementable options' and 'motivations as interpreted'. Implementable options are built up from 'potential options' and 'autonomy' (layer 3). Potential options are all options the actor is aware of. Though, not all of these options can be implemented. The difference between the implementable options in the second layer and the potential options in the third layer is the difference between what the actor really can do as opposed to what the actor might do if he had the possibility. This difference is determined by the so-called autonomy of the actor. The autonomy consists of resources and restrictions, which together determine which options an actor can implement. Resources contribute to the actor's capacity to implement actions they consider. The nature of these resources can be economic, social, cognitive, environmental, moral, psychological and physical. For example, for a farmer to grow corn he needs access to land, money to buy inputs and the knowledge how to cultivate corn. Restrictions are autonomy reducing factors, like prohibitions, prescriptions and standards related to environmental licences, but also include physical restrictions.

Motivations are the aspects of the options under consideration by the actor that are normatively relevant to the actor (*i.e.* that give value to the different options). In layer two the motivational factors are specified in terms of 'advantage and appropriateness' as interpreted by the actor. These interpreted motivations are determined by 'objectified motivations' 'interpretations' (layer 3). The objectified motivations are easily quantifiable units, such as economic costs and benefits or caloric value of produced foods (Verburg *et al.*, 2003). Interpretation is shaped by the cultural and psychological opinions and ways of looking that give weight, coherence, shape and colour to the objectified motivations. Together they form the motivations as interpreted by the actor.

In the fourth layer the factors of the third level are seen as being determined by 'micro-structure', 'macrostructure' and 'interpretative frames, self image, world views', which is the cultural aspect of the actor's context. A more elaborate description of the deeper analysis can be found in De Groot (1992) and Overmars *et al.* (2006) (Chapter 3).

The options and motivations of the deeper analysis are used to construct the relations between land use and the explanatory factors in a theoretical-deductive manner as opposed to the inductive method described above. Normally, the AiC approach is applied to cases in which actors or households are the objects of study (e.g. Overmars *et al.*, 2006 (Chapter 3)). In the CLUE-S model locations, regular grid cells of 50 by 50 m, are the unit of analysis. Therefore, the options and motivations of land managers have to be converted into suitability maps. This conversion is not always straightforward (Overmars and Verburg, 2006 (Chapter 2)). The field characteristics from the deeper analysis can be easily represented in maps, because field characteristics are directly linked with locations. The influence of household characteristics on land use, as determined in the deeper analysis, is more difficult to incorporate in the suitability maps, because household data are not available in maps. Instead, the household variables are represented as aggregates at the village level. This aggregation may lead to aggregation problems, but the logistic regression analysis of the deductive approach revealed that the aggregated effects of the household variables

(ethnicity) are also present. The interpretation of those variables should be made on village level, because the relations at village level can be different from the relations at household level.

As far as the policies and restrictions are spatial these can also be directly represented in maps. Some spatial policies restrict all land use change in a certain area, like, for example the protection of a nature reserve which is implemented very strictly. Other land use policies restrict a single land use conversion, like the prohibition of the construction of houses in designated agricultural areas or permanent agriculture in the buffer zone of a nature reserve.

An important difference between the regression approach and the deeper analysis is the way potential options are modelled. In the deeper analysis spatial policies and restrictions (or lack of resources) reduce the autonomy of the land manager and therefore the number of options a farmer can implement. This approach really excludes land use type on certain locations, whereas in the regression analysis including variables can only lead to a reduced probability and not to a real exclusion of a land use type at a location due to the way the regression model is specified.

5.3.3 The CLUE-S modelling framework

In this section the CLUE-S model is described and for those aspects that are the same in both modelling approaches a specification is provided. The difference between the two model applications that will be presented is the way the suitability maps are computed. The modelling framework that is used is the most recent version of the CLUE-S model (Verburg *et al.*, 2002; Verburg and Veldkamp, 2004; Verburg *et al.*, 2004c). This modelling framework is used to integrate different mechanisms of the land use system into a spatially explicit land use model that is capable of the dynamic simulation of competition and interactions that occur in land use systems. Incorporation of these mechanisms result in model output that shows path dependency and non-linear behaviour, which characterises the land use system in real-world situations. Path dependency implies that model results of earlier modelling steps have their influence on later modelling steps. Path dependency is dependent on the specification of the incorporated mechanisms, for example land use history and conversion rules, and initial conditions (Brown *et al.*, 2005). A general way to look at models, mentioned by Couclelis (2001), is that they are frameworks for organising knowledge. This description fits very well the CLUE-S modelling framework, which integrates different aspects of the land use change process.

The CLUE-S model consists of an allocation module and a series of inputs. The allocation module is a computer program that iteratively computes land use allocation for a number of modelling steps (a detailed description of the allocation module is provided by Verburg *et al.* (2002)). The allocation module can incorporate various mechanisms that are considered to determine the distribution of land use changes in a landscape. These mechanisms are parameterised by the inputs of the model. The quantity of land use change is also imposed to the model as an input. These land use requirements impose the quantity of land use change per modelling step for every land use type in the whole study area. The future 'land claim' can be a fictitious land use scenario based on story lines, as will be used in this study, or an external modelling procedure like macro-economic modelling. Then, the allocation module allocates the aggregated land claim year by year to the cells based on the various mechanisms in an iterative way. So, the strength of the CLUE-S model is

allocate land use changes rather than modelling the quantity of change. The mechanisms responsible for the land use allocations can be divided into location characteristics and conversion characteristics. The first locational characteristic is the 'suitability', which is based on the relation between land use and a broad set of biophysical and socio-economic factors. Suitability has an important influence on the allocation of land use change in the model. The basic assumption behind this mechanism is that a location changes into a certain land use in those locations where the 'suitability' is high for that land use type. Suitabilities are represented as a map with values between 0 (low suitability) and 1 (high suitability). This is where the deductive or the inductive approach to derive the land use suitability maps is inserted.

The second location characteristic allows for the incorporation of spatial policies. The suitability map can be altered at locations that a policy applies to. The suitability can be set to zero at locations where a land use type is not allowed to change, for example in a conservation area, or the suitability value can be adjusted by a certain value in areas that are under a policy that, for example, awards subsidies for a certain land use in that area. This mechanism is not used in the applications in this study.

The third location characteristic is the neighbourhood effect (Verburg *et al.*, 2004c). Although several theories are available addressing the interaction between neighbouring land use types, for example trends in explanatory variables or spatial processes like imitation, this interaction was not studied extensively in this research. Neighbourhood effects can be included between land use types as well as within a land use type. Because the cell size of the application is smaller than the average parcel size, a small neighbourhood effect was implemented in the model for all land use types with themselves, simulating the cluster of land use into fields and parcels. The value of the neighbourhood effect was based on eight closest neighbours of each cell. In the calculation of the overall suitability to be used in the model the neighbourhood function determines 20 percent and the suitability map of the inductive and deductive approach determine the other 80 percent.

The conversion mechanisms that can be incorporated in the model are the so-called conversion elasticities and land use type specific transition sequences. Conversion elasticity can be explained as the resistance of a land use type to change location. For example, tree plantation cannot easily move to another location because the investments made to establish the tree plantation are lost when the plantation moved to another location to make room for another land use type. The conversion settings can be used to create stability in the model by assigning a large influence to the land use history (Verburg *et al.*, 2002). The conversion elasticities are implemented in the model as an additional suitability for those locations that are currently under that specific land use. The user should decide on this factor based on expert knowledge or observed behaviour in the recent past or use the factor to calibrate the model. The conversion elasticities that are incorporated in this study are estimated by the authors based on field knowledge and can be motivated as follows. Grassland is easily converted and was given a low conversion elasticity. Corn has a somewhat higher elasticity value since it is relatively easy to establish a corn field. The requirement for corn is a cleared field. For banana higher investments have to be made as it takes time before the fruits can be harvested, therefore this land use type received a higher elasticity compared to grass and corn. For rice a considerable effort has to be made to construct a rice paddy. Therefore, rice received a high elasticity. Secondary forest was given an intermediate value and forest a high value.

The transition sequence is a set of rules that determine the possible land use conversions.

Comparison of a deductive and an inductive approach

Not all land use changes are possible and many land use conversions follow a certain sequence. Sometimes these conversions include a temporal constraint. The conversion mechanisms determine to a large extent the temporal dynamics of the simulations, because they include land use history. In the model applications most land use conversions are allowed except for changes into secondary forest and forest. The only pathway allowed changes into secondary forest is through grass or banana. The idea behind this rule is that the field must be not used for five years, and thus be grassy, to become secondary (regrown) forest. Banana fields in the area are cultivated quite extensively and often trees are present. It is considered that if a banana plantation is not maintained for three years this banana plantation can become secondary forest. From secondary forest it takes another five years to grow into mature forest, which is the only pathway to mature forest. The time necessary to grow from one land use into another is estimated and might be subject for further research. Incorporating the effect as such does incorporate path dependency in the model, although it might be not the exact number of years. Banana is allowed to remain for a maximum of twenty years. After these twenty years the banana plantation has to change for at least one year. This rule is based on the lifespan of a banana plantation, which is about twenty years in the area. After these years the banana plant is not producing anymore and is replaced with an annual crop for a short period after which bananas can be replanted. The various mechanisms are combined in the allocation module. The allocation of all land use types in the case study occurs at the same time in an iterative procedure. Altogether this results in the dynamic simulation of land use competition.

5.3.4 Scenarios

Two scenarios were developed to test the models. The first scenario provides a general indication of what might happen in the research area. The principles are based on the comprehensive development plan of the municipality of San Mariano (Municipality of San Mariano and Housing and Land Use Regulatory Board, 2000). The plans and prospective indicated in this document are translated into general linear trends. The quantities of change are assessments rather than detailed calculations, but will suffice for the objective of this chapter. The second scenario introduces a new land use type.

As indicated in the planning document of the municipality, the total amount of agricultural lands will not increase substantially for two reasons. First, slope prevents agriculture to expand much further, because productivity of the steeper areas is low. Secondly, there is a necessity for more environmentally-friendly activities in these areas to prevent soil erosion and flooding and to protect natural values of the area. Forested areas will be protected and grasslands with potential for forest production will be rehabilitated and protected. To improve food security and self-sufficiency the production of rice will be promoted. Furthermore, the municipality aims at an increase in productivity per hectare to meet the necessary production and idle lands (mainly unused grasslands) should be taken into production (Municipality of San Mariano and Housing and Land Use Regulatory Board 2000). The improvement of accessibility, which is also an important municipal goal, is not taken into account in this study.

The scenario sketch above is translated to a quantitative yearly land claim (Figure 5.3 left). The land claim is the total area per land use type for every modelling step and serves as input to the CLUE-S model, which is specifically developed to allocate this land claim. The translation into real figures is fabricated by the authors based on the ideas of the munic-

pality assuming that the changes are predominantly linear. The scenarios are a projection of changes that might happen rather than strong predictions with a prospective value. In Figure 5.3 (left) the rice area expands by five percent of the 2001 rice area per year to better meet self-sufficiency in rice, which is the main staple crop of the population. The corn area increases by one percent of the 2001 area per year (allowing some agricultural intensification) and banana area reduces by one percent compared to the 2001 area per year. The latter is to represent improved production of the banana area as well as a reduced interest in this crop due to diseases. Forest area remains the same to visualise the intended conservation effort. Secondary forest will grow by two percent. All increases of land use types are at the expense of the grass area. This scenario is used in the model application using the inductive approach as well as in the application with the deductive approach. In the second scenario a new development was included (Figure 5.3 right). In a certain policy area (SIFMA) an NGO stimulated the cultivation of (fruit) trees by providing seedlings and assisting farmers to acquire a 25-year tenureship for the parcels involved. The possibility to change to fruit trees was restricted to the SIFMA area, which was identified as an area that needed reforestation. Using the deeper analysis framework an analysis was made for this new land use type, which was not yet present in the original land use map. So, this scenario could only be modelled using the deductive (AiC) approach because in the inductive approach it is not possible to make inferences for land use types that are not yet present. The information of this analysis was also converted into a suitability map. With this information fruit tree plantations were introduced in the scenario for the deductive modelling approach. The scenario starts with a newly established area of 150 ha with a 5 ha increase in the following years. The extra area for fruit trees was introduced at the expense of the grass area.

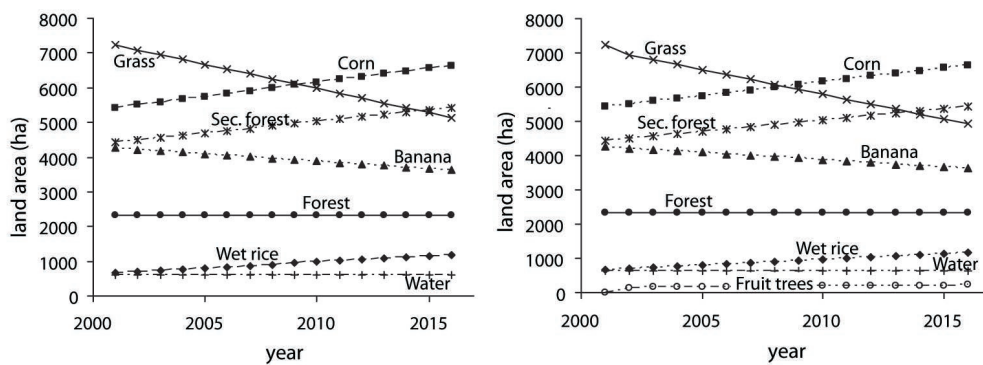


Figure 5.3: Scenario-based land claim (left) and scenario-based land claim including fruit tree plantations (right)

5.4 Results

5.4.1 Results of the inductive CLUE-S application

The results of the regression analysis that was used to determine the relation between land use and its explanatory factors are presented in Table 5.2. For rice and corn all variables

Comparison of a deductive and an inductive approach

were entered in the stepwise regression procedure. For the land use types banana, secondary forest and forest the ethnicity variables were excluded. Ethnicity was assumed to have no relation with these land use types. Table 5.2 shows which variables were actually included by the stepwise procedure together with their significance level

Table 5.2: Regression coefficients of the resulting logistic regression model

Variables	Rice	Corn	Banana	Sec. forest	Forest
Constant	-3.856***	-0.599**	-0.838***	-2.607***	-8.142***
Slope1	0.515*	1.137***	-0.655***	-0.756***	
Slope2		0.750***			
Slope3		0.399**		0.230**	0.725***
Slope4	-1.077*				1.396***
Slope5					1.551***
Creek	1.010***	-0.217*		0.313***	
Plot distance	-0.009***	-0.016***	-0.005***	0.002***	0.021***
Transportation cost		-0.026***	-0.012**	0.031***	0.084***
Ethnicity Ilocano					
Ethnicity Ifugao	0.020**				
Ethnicity Ibanag		0.011***			
ROC	0.73	0.77	0.65	0.68	0.96

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The ROC (Relative Operating Characteristic) (Swets, 1988) was used to indicate the goodness-of-fit of the regression models. The ROC summarises the performance of a logistic regression model over a range of cut-off values. The value of the ROC is defined as the area under the curve linking the relation between the proportion of true positives versus the proportion of false positives for an infinite number of cut-off values. The ROC statistic varies between 0.5 (completely random) and 1 (perfect discrimination) (see for more detail Pontius and Schneider, 2001). Rice and corn have a good model fit, banana and secondary forest had a relatively poor fit and forest had a very good fit

For grass and water no regression analyses were made. Grass is a left over land use type that is not used and is not cultivated intentionally. It just grows on locations that are not occupied with one of the other land use types and is therefore treated as a 'rest' category without any specific suitability. Water is considered to be constant over the modelling period and is therefore excluded from the modelling exercise. Examples of suitability maps for banana and secondary forest, which are constructed with the relation found in the regression analysis, are shown in Figure 5.4A and 5.4C

The resulting land use map of the inductive modelling approach after a 15 years modelling period is shown in Figure 5.5B. Under the scenario and modelling assumptions applied this model the following major trends can be identified. First, the banana area decreases the area marked with 1 and is relocated in area 2. The abandoned areas in area 1 are occupied with grass. In general, existing corn areas (like in the area near 3) expand throughout the area near places where corn was already present. Rice expands in the area near 4 and just below area 2. Secondary forest increases mostly near 5. Forest is stable in this scenario

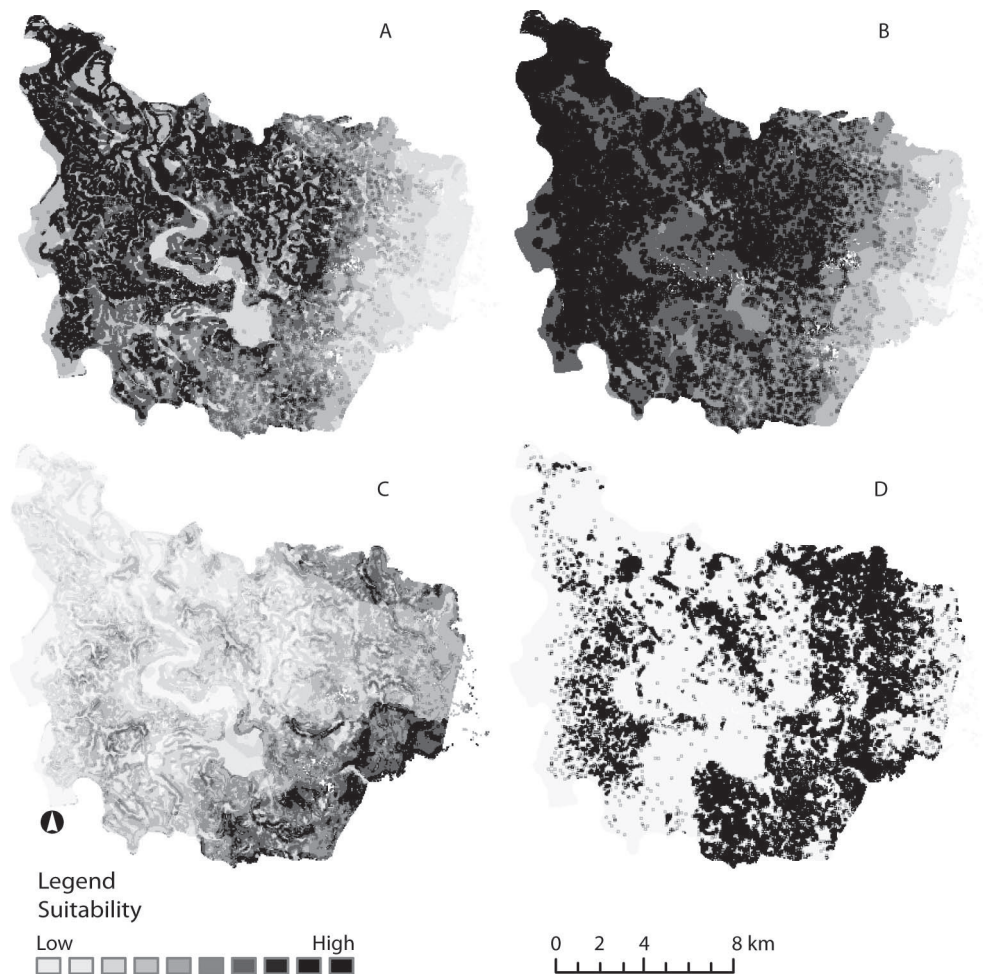


Figure 5.4: Suitability map for banana for the inductive case (A) and the deductive case (B) and suitability map for secondary forest for the inductive case (C) and the deductive case (D). The neighbourhood effect is added to these suitability maps (note that the neighbourhood effect is recalculated every time step and changes with changes in the land use). The scaling of the legends was stretched between the highest and lowest value.

5.4.2 Results of the deductive CLUE-S application

The first element of the actor decision model that is described are the motivations to cultivate a land use type, which are calculated according to Overmar *et al.* (2006) (Chapter 3). The relations between the land use and the explanatory factors are formulated in Equations 5.1 to 5.7. For every land use type this calculation is different. The parameters for this model were calculated using field observations where possible and otherwise they were based on expert knowledge and interviews with farmers. Not all parameters are provided in this chapter, but can be accessed from (Overmar *et al.*, 2006 (Chapter 3))

Comparison of a deductive and an inductive approach

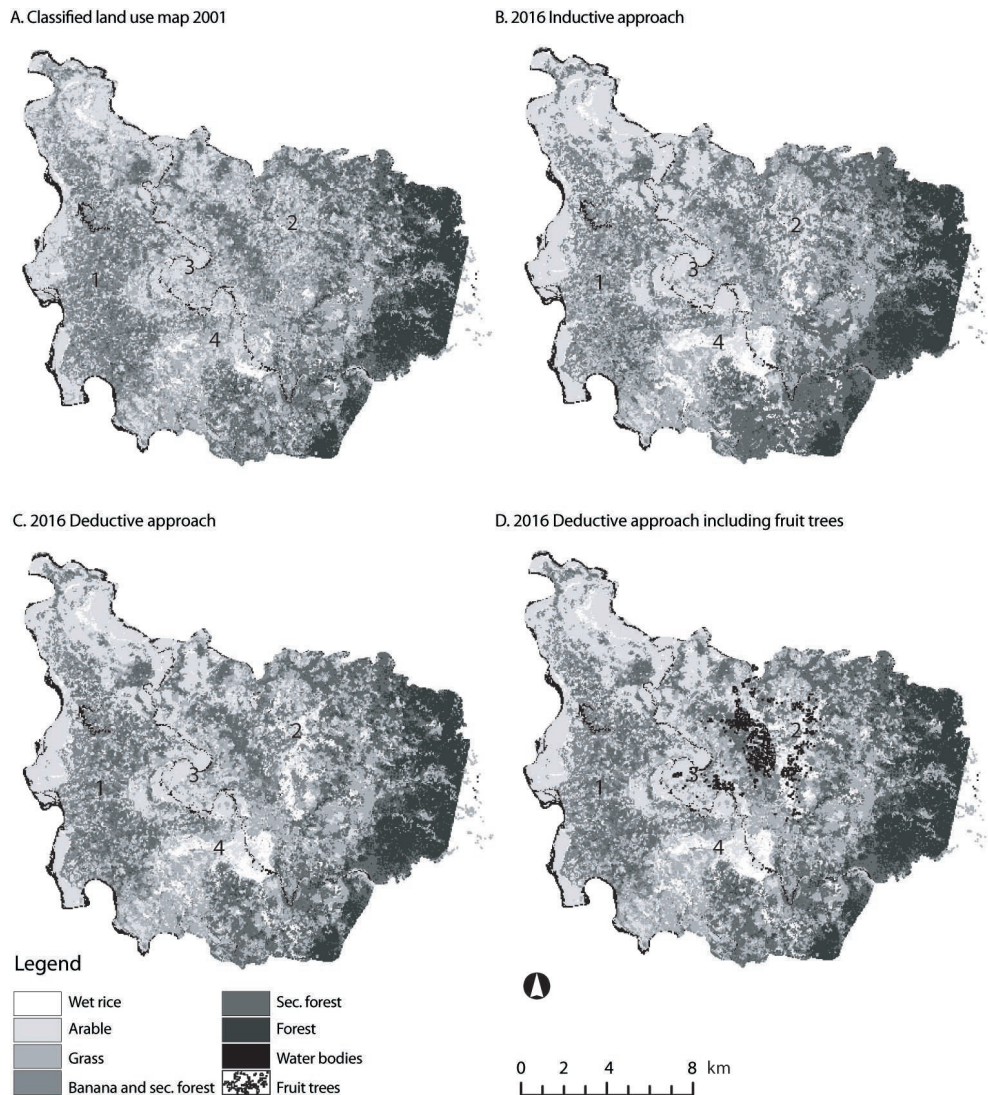


Figure 5.5: Classified land use map 2001 (A), simulated land use map in the year 2016 of the inductive modelling approach (B), the deductive modelling approach (C), and the deductive approach including fruit trees (D)

The maximum benefit is different per land use type. The slope factor applies only to corn and reflects the losses in yield due to slope. The risk is calculated from the average year loss due to typhoons, droughts, pests and diseases as reported by farmers. The costs to transport the harvested product are a combination of cost made by travelling from the field to the residence of the farmer and transportation costs to the market in San Mariano (the latter is not included for wet rice). Preferences for a specific land use type depend on the ethnicity of the household. Since the household information is not available this preference is calculated for the village as a whole. The ethnicity specific preferences are multiplied

Chapter 5

the percentages of the population that belong to an ethnicity and summed for all ethnicities. The preference for corn is considered to be higher than average for ethnicity Ibanag and low for ethnicity Ifugao. For rice the preference is high for Ilocano and Ifugao and low for Ibanag (compared to the group of other ethnicities)

$$\text{Motivations} = \text{objectified motivations} * \text{preferences} \quad (5.1)$$

$$\text{Objectified motivations (net benefit)} = (\text{max_benefit} - \text{tr_costs}) * \text{slope_fact} * (1 - \text{risk}) \quad (5.2)$$

$$\text{Max_benefit} = f(\text{CROP}) \quad (5.3)$$

$$\text{Slope_fact} = f(\text{SLOPE}, \text{CROP}) \quad (5.4)$$

$$\text{Risk} = f(\text{CROP}) \quad (5.5)$$

$$\text{Tr_costs} = f(\text{TR_COST}, \text{PLOT_DISTANCE}, \text{CROP}) \quad (5.6)$$

$$\text{Preferences} = f(\text{ETHNICITY}, \text{CROP}) \quad (5.7)$$

The implementable options are also based on Overmars *et al.* (2006) (Chapter 3). The implementable options for rice were determined by slope, which should be flat or flat to moderate, and the possibilities for irrigation, close to a creek or an irrigation facility. For corn and banana no restrictions were formulated in this analysis.

The options and motivations together form the suitability maps. The deeper analysis was used to calculate suitability for rice, corn and banana. The factors that are incorporated in the model and the land use type are indicated in Table 5.3 as well as the performance of the model, which is indicated with the ROC. These ROC values were added to compare the results with the inductive approach, although the aim of the deductive approach is not to get the best fit possible, but a good representation of the processes. This subject will be further explained in the discussion for the case of banana. The remaining land use types were modelled as having an equal suitability for all locations. For grass we made the same assumption as in the inductive approach. In contrast with the case for grass, the other two land use types without suitability analysis actually do have a use. Forest is used for (illegal) logging and for this purpose accessibility plays an important role. Secondary forest can be also used as timber or firewood and therefore accessibility may also play a role. Part of these processes is covered by the neighbourhood functions that were incorporated. The argument not to include suitabilities for forest and secondary forest in these model applications is that under the scenario forest is stable and secondary forest is increasing. The suitability for increasing (regrowing) forest and secondary forest is not related to factors that determine suitability for logging and can actually be constant for all locations.

The resulting land use map of the model application with the deductive approach after fifteen years modelling period is shown in Figure 5.5C. In this application the following major trends can be identified. First, the banana area decreases in the area marked with 1 and is relocated in area 2. The abandoned areas are occupied with grass and secondary forest. Like in the inductive approach existing corn areas expand. Rice expands in the area near 4 as well as below area 2. Secondary forest increases evenly in the study area and the forest area is stable.

In the second scenario applied to the deductive modelling approach fruit tree plantations are introduced. The suitability for this land use type (Figure 5.6) is similar to that of banana, although the general profitability of the fruit tree plantations is higher than the cultivation of banana. The most important difference with the suitability of banana is that the fruit trees are restricted to an area where the SIFMA policy applies to because in the other areas

Comparison of a deductive and an inductive approach

Table 5.3: Factors included in the deductive approach indicating if the factors are incorporated in the options or the motivations of the land managers

Variables	Rice	Corn	Banana	Sec. forest	Forest	Fruit trees
Slope1	options	motivations	-	-	-	-
Slope2	options	motivations	-	-	-	-
Slope3	-	motivations	-	-	-	-
Slope4	-	motivations	-	-	-	-
Slope5	-	motivations	-	-	-	-
Creek	options	-	-	-	-	-
Plot distance	motivations	motivations	motivations	-	-	motivations
Transportation cost	-	motivations	motivations	-	-	motivations
Ethnicity Ilocano	motivations	-	-	-	-	-
Ethnicity Ifugao	motivations	motivations	-	-	-	-
Ethnicity Ibanag	motivations	motivations	-	-	-	-
SIFMA	-	-	-	-	-	options
ROC	0.66	0.74	0.54			

initial investments are considered to be too high

The land use changes are similar to those in the deductive approach (without introduction of fruit trees). The additional area for fruit trees caused other land use types (mainly banana) to move to other areas

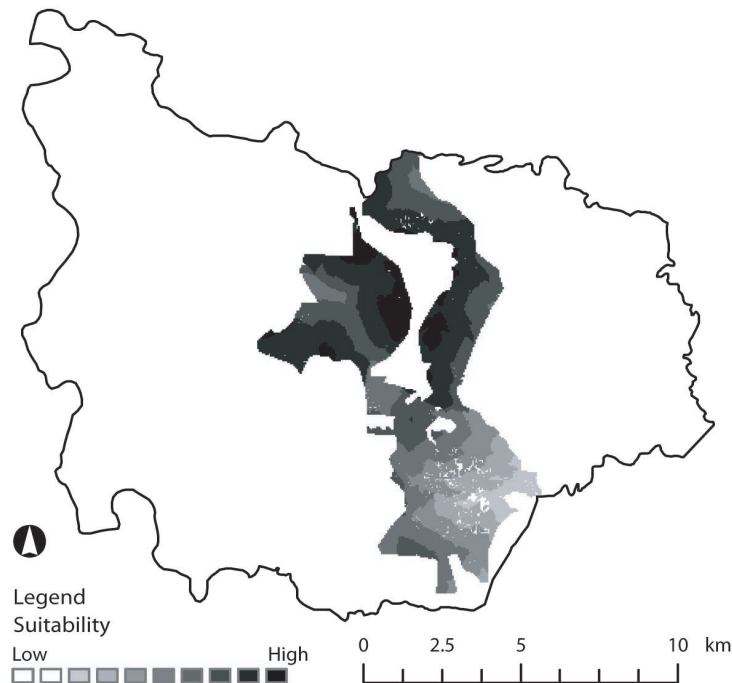


Figure 5.6: Final suitability maps for fruit trees including the restrictions

5.4.3 Comparison of the two modelling approaches

If the maps of the application using the inductive and the application using the deductive approach (Figure 5.5B and 5.5C, respectively) are compared cell by cell the maps have 85 percent in common. In total the land use changed in 25 percent of the locations after 15 years. From these changes 54 percent were exactly the same changes in both modelling approaches. If the comparison is made within larger windows (Costanza, 1989), allowing for differences in location within the window, the similarity increases (Figure 5.5).

A general observation from these scenario studies is that if the grassy areas are used both the agricultural as well as the forested areas can be improved. Agriculture is the main source of livelihood in the area and the forest can sustain the ecological function of the area. Grassland on the contrary does not contribute to production and neither has much ecological value.

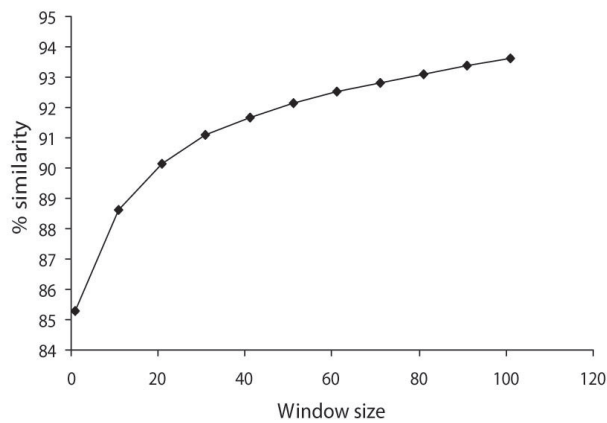


Figure 5.5. Similarity between the inductive and deductive modelling approaches with varying window size based on their 2016 maps

5.5 Discussion and conclusions

At first instance the final results of the inductive and deductive modelling approaches look quite similar (Figure 5.5B and 5.5C). If only the changes are considered 54 percent of the changes are at different locations. These differences between the maps of the two applications are caused by differences in the suitability maps, which were constructed in different ways. After all, the other inputs and model settings (i.e. the neighbourhood effect and the conversion mechanisms) are completely identical, so these cannot cause the differences. The deductive and inductive approach to create suitability maps for a land use type can vary because of different variables, different relations between the variables and different parameters. Some suitability maps show only local differences. These differences in suitability maps translate into differences in land use allocation at small distances, which disappear when the similarity is compared with bigger windows. A good example of local differences is the suitability of banana. In the inductive case (Figure 5.4A) slope was included while in the deductive approach (Figure 5.4B) slope was not included. As can be

Comparison of a deductive and an inductive approach

seen the general pattern, caused by plot distance and transportation cost (and having the same sign), is similar, but the slope introduces differences at small distances. The suitability maps of secondary forest show more differences. The inductive approach (Figure 5.4) detected correlation with a series of factors while in the deductive approach (Figure 5.4) no factors were included and therefore the suitability map shows only the neighbourhood effect. These differences in suitability are reflected in the differences in the resulting land use maps even if larger windows are used in the evaluation procedure.

So, the inductive and deductive approach to specify the suitability map do not always lead to the same results, because the inductive approach is based on correlation while the deductive approach is based on processes that were observed by the authors and/or described by the respondents. In other studies suitability from both approaches may be more different leading to even more differences in the suitability maps of the two approaches.

Even more important is that the research paradigms of the two approaches are different which has its implications for the interpretation and the use of the modelling approaches. The structural difference between the two approaches used in this study is that the deductive approach attempts to describe causality while the inductive approach to land use analysis reveals associations rather than causal relations (Serneels and Lambin, 2001; McConnell, 2002; Verburg *et al.*, 2003; Verburg *et al.*, 2004d).

In the inductive approach the current land use pattern is assumed to reflect the process of land use in the past. The result of these processes, the land use pattern, is described by the regression model using correlations between the land use and its explanatory variables. The processes themselves are not described and, therefore, changes in the processes and their effect on the suitability of a location for a land use type cannot be incorporated in the modelling of future scenarios. So, using the inductive approach, the assumption that has to be made for the modelling exercise is that the processes that determine land use do not change. This approach is described by a study by Kok and Winograd (2002) where modelling of scenarios with and without the impacts of Hurricane Mitch (Honduras) results in the same land use map after ten years under the assumption that the relations between land use and its drivers was re-established after a few years. It may be true that relations do not change at short time scales, but at larger time scales different factors may become important and sudden events, like a change in political system may cause dramatic change in behaviour. In the models presented, however, the behavioural rules are assumed to be constant. Besides this, no new land use types can be introduced, since the relation of this new land use type with the explanatory factors cannot be determined statistically. Even if the regression analysis was able to describe processes the assumption has to be made that the land use system is in equilibrium with the explanatory factors. Analysing a system that is not in equilibrium may lead to possible error in the description of the processes.

The deductive AiC approach, on the other hand, describes the processes explicitly. Therefore, changes in the processes that determine land use can be incorporated in the construction of the suitability maps, which enables the introduction of discontinuities and new land use types in the scenarios that are modelled. All these issues have their consequences on the type of scenarios that can be simulated with the modelling approach. A case with discontinuity was demonstrated to some extent by Kok and Veldkamp (2000), who used a rule-based suitability map for a new land use type to enable the incorporation of this land use type, like was done in this chapter with a sound theoretical framework. The other suitability maps in the study by Kok and Veldkamp stem from regression analysis.

A more technical difference between the two approaches is that with an inductive approach the regression analysis determines the relation between the current pattern of the land use types and its explanatory factors, whereas the AiC approach determines the potential suitability of the land use types. In the regression analysis the occurrence of a land use type, which serves as the dependent variable in the regression analysis, is not independent of the other land use types. This dependency has its consequences for the applicability of the modelling approach. To illustrate this consider the example of the bananas in this study. In general, bananas are located on 'second best' locations, because the best locations are cultivated with corn, which is (potentially) more profitable. For example, in this study bananas are correlated with high slopes, which is due to the fact that on the flat part corn is preferred, not because bananas grow better on steep slopes. In the inductive approach the calculated suitability for banana is high at these second best locations. If a large change in banana area would occur, for example when suddenly all corn would disappear and the banana area would expand fast (large changes may happen for example through large price changes or diseases), these new banana areas would first be allocated on areas with a high banana probability, which in fact are the second best locations. In reality the new banana would first appear where the suitability for banana is optimal. The deductive approach would allocate these where the potential suitability for banana is high. Generally speaking, the inductive approach to specify the suitability map in CLUE-S is applicable in situations with relatively small land use changes, without introduction of new land use types, whereas the deductive approach to specifying the suitability map in CLUE-S approach is more flexible in this respect.

The advantage of the empirical approach is that the procedure of the regression analysis is straightforward and easy to reproduce. Limitations of the empirical approach are that many regression models have a restricted specification of the relation between variables (e.g. linear, log-linear). Though, increasingly, statistical tools are introduced that can capture the structure, and therefore also part of the processes, of the land use system. For example multilevel models (Pan and Bilson, 2005) can incorporate a hierarchical structure and autoregressive models (Overman *et al.*, 2003) can capture spatial processes.

The AiC analysis used in the deductive approach depends on the skills and interests of the researcher. Therefore, the AiC analysis is less reproducible than the inductive analysis. The land use system does not have to be in equilibrium because the processes are observed directly rather than derived from the current land use pattern. Finally, the deductive (AiC) approach does not constrain the specification of the mathematical relations between factors in any way, giving more flexibility to the model.

In this respect it is regrettable that the household information and the ownership relations with the land were not spatially available for the study area. The distribution of the parcels and their ownership is an important determinant of the observed land use pattern. By not incorporating this structure the model has the tendency to allocate the land use according to the smooth patterns of the suitability maps, while the observed land use pattern shows a more irregular pattern due to land ownership. The AiC analysis of Overman *et al.* (2006) (Chapter 3), from where the deductive approach is derived, is based on a household survey and could have been easily incorporated if this information was available.

The differences described above have their implications for the applicability of the models to answer questions in research and policy-making. To have some foothold to assess the use of the two modelling approaches for research and/or policy Couclelis (2001) provides

Comparison of a deductive and an inductive approach

some qualifications for both: Besides that both must be built on good science, use good data, and should answer good questions, research models should have a higher degree of scientific rigor and should contribute original theoretical insights or technical innovation. Policy models should preferably be used, verified and validated often and should be transparent and manipulable and should include key policy variables.

As far as the qualifications for research are concerned both approaches are quite similar. The main difference is that the approaches stem from two completely different research paradigms. With respect to the policy issues the two approaches do show important differences. First of all, the inductive approach is more transparent and the CLUE-S model using this approach is validated for several cases (Kok *et al.*, 2001; Pontius *et al.*, 2005). The AiC approach is dependent on the judgement of the researcher and is therefore less transparent and reproducible. Secondly, the deductive (AiC) approach is more flexible (manipulable) than the inductive (regression) approach, which has to stick to more rigid model definitions. Concerning the inclusion of key policy variables the deductive approach has the advantage of the explicit description of parameters and relations between variables. Another advantage that adds to that is that it can include variables like market prices and investments. These variables are constant throughout the study area and can therefore not be included in regression analysis. They are included in the AiC analysis and can therefore be used to study the effect of changes in price (through for example subsidy policy) or changes in technology, which can be important policy variables. So, the deductive method has more options to analyse the effects of policies, which are often implemented at the macro-level. Potentially, this approach would also enable the modelling of the amount of land use changes and therefore the possibility to make the model more dynamic.

The scale to which both approaches can be applied is different. In principle the inductive approach can be applied to any scale (*i.e.* resolution and extent). However, the amount of detail and knowledge about the decision-making structure of actors involved that can be incorporated is limited. The deductive approach as presented in this study relies on detailed information about the land managers. To incorporate this information in a spatially explicit model the resolution should be comparable with the size of the decision units of the actors. Aggregating these units to larger grid cells would lead to aggregation problems. So, the deductive approach should preferably be applied to the watershed level using a fine resolution.

Currently, many efforts in land use modelling have adopted the multi-agent modelling approach (e.g. Parker *et al.*, 2003), which is an agent-based approach in which actors communicate and interact. The deductive approach in this chapter is not a multi-agent model. However, the model can be regarded as an agent-based. It specifies the decision of farmers in various circumstances, but without communication and interaction and without other actors than farmers involved. By using an actor-decision model to specify land suitability, decisions of the land manager were given a more prominent role in the modelling approach than with a statistical approach. The deductive approach provides more process-information than the inductive approach although the representation of the actors involved is simplified to one representative actor.

Both the deductive and the inductive approaches have their own origin and research paradigms and their own advantages and disadvantages as pointed out in this final section. Within the scope of this study, no qualification of the models was presented that was based on validation of the simulated land use maps. This would not have provided many new

Chapter 5

insights because the resulting maps were quite similar in this case study. A more important conclusion is that the research question and the nature of the case that is studied determine which approach is most suitable to use. The deductive approach can better handle discontinuities in land use processes and can therefore evaluate a wide range of scenarios, which can also include new land use types. The inductive approach is easily reproducible and well able rapidly to identify hotspots of land use change. The deductive approach is better suited for smaller study areas, but needs fieldwork to implement. The inductive approach can be applied more quickly in larger areas if basic data are available.

