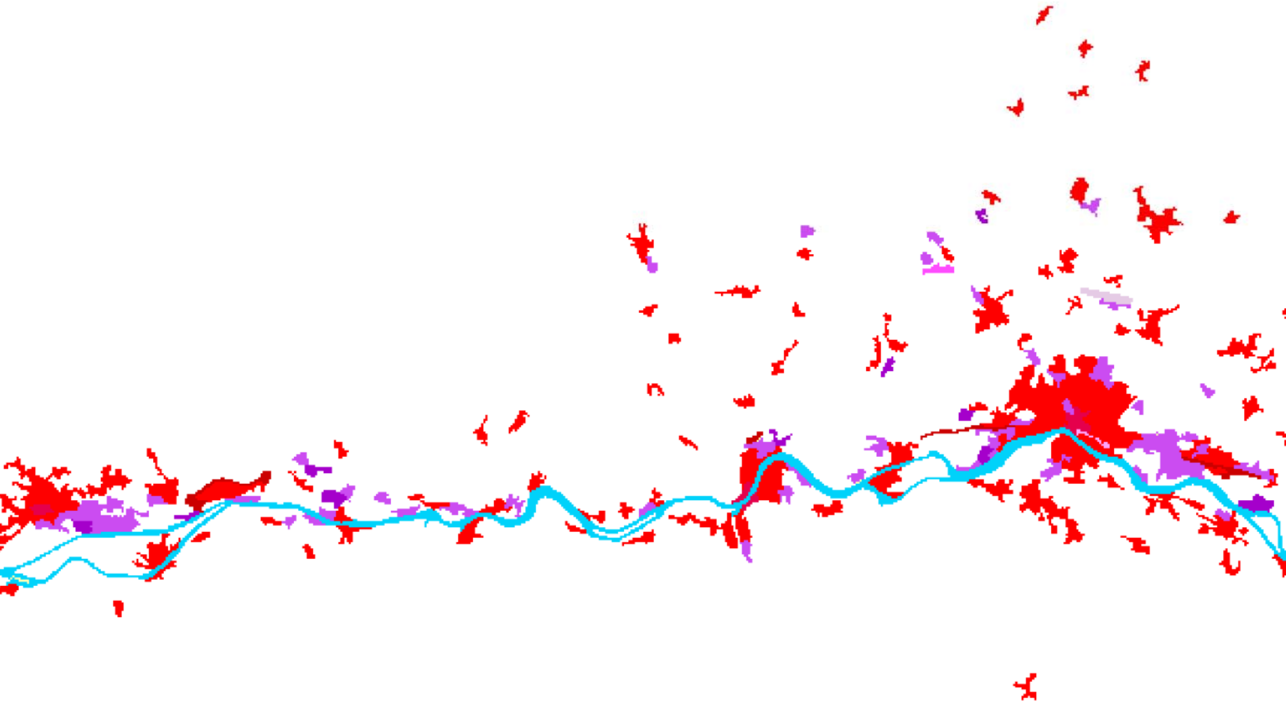
Calibration and validation of land-use models



Jasper van Vliet

Calibration and validation of land-use models

Jasper van Vliet

Thesis committee

Promotor

Prof. Dr. Ir. A.K. Bregt
Professor of Geo-Information Science
Wageningen University

Other members

Prof. Dr. H.B.J. Leemans, Wageningen University
Prof. Dr. D.G. Brown, University of Michigan, USA
Prof. Dr. D. Haase, Humboldt University of Berlin, Germany
Dr. Ir. S.J.C. Janssen, Alterra, Wageningen

This research was conducted under the auspices of the C.T. de Wit Graduate School for Production Ecology & Resource Conservation (PE&RC).

Calibration and validation of land-use models

Jasper van Vliet

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 21 January 2013
at 11 a.m. in the Aula.

Jasper van Vliet
Calibration and validation of land-use models
162 pages
Ph.D. thesis, Wageningen University, Wageningen, The Netherlands (2013)
With a summaries in English and Dutch
ISBN: 978-94-6173-443-3

Contents

1. Introduction

1.1 Land-use modelling

Land use is changing constantly. For example, some areas experience rapid urbanisation, while other areas face deforestation or agricultural intensification. These changes in itself can have important consequences such as congestion or loss of biodiversity, but also an increase or decrease in food production. For practical, ethical and financial reasons it is not possible to study these land uses in controlled experiments (Janssen and Ostrom 2006). Therefore, models are the predominant tools to study land-use changes.

Over the last two decades a number of land-use models have been presented that explicitly allocate land-use changes on a map using a simulation approach. Their results are represented as land-use maps, typically a regular grid with the cell state indicating the predominant land use at a location. The simulation approach indicates that the land-use model simulates changes over time instead of a static situation (Hartmann, 1996). In that sense these models are fundamentally different from classical economic land-use models that yield static equilibrium situations only (Anas et al., 1998; Albrecht, 2005).

The development of a land-use model is a process that involves several steps. These steps are schematically shown in Figure 1.1, which is a modified version of comparable figures available for computer simulation (Sargent, 1998) and hydrological modelling (Refsgaard and Henriksen 2004).

The **problem entity** is the process or phenomenon that is the actual topic of research. This problem entity can involve land-use change in general, but often it is more specific, such as the influence of land-use policies, or particular landuse change processes like desertification, urbanisation, or deforestation.

A **conceptual model** is a description of this problem entity, either verbal, or in terms of equations and relations. Basically, it is the modellers' perception of reality, which follows from a thorough analysis of this problem entity. A conceptual model inherently is a simplification of reality. Conceptual modelling is therefore a selection of those processes, components, and relations that are required to study the problem entity. Basically, it is a hypothesis about the problem entity. Conceptual validation assesses whether the underlying theories and assumptions are appropriate, whether the conceptual model is logic and whether causal relations are reasonable for the intended purpose of the model. This assessment can include checks for linearity or statistical properties, but also the application of Occam's razor, which indicates that a model should be as simple as possible, but not simpler (Jakeman et al., 2006). This does not indicate that models should be very simple, as land-use changes are typically complex processes; it only means that a model should not contain unnecessary elements (Clarke, 2004).

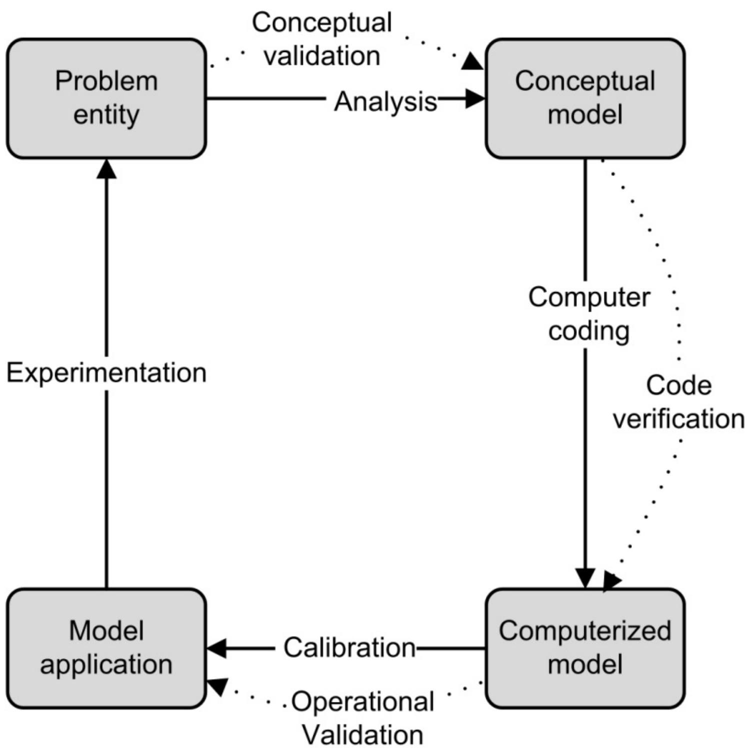


Figure 1.1: *Model development cycle. Solid lines indicate activities for model development or use, while dotted lines indicate model testing activities.*

The **computerised model** is the computerized representation of the conceptual model. The computerized model is typically available as a generic model in the sense that the relations and equations from the conceptual model are implemented but the values of model parameters are not yet defined and no data is included. Examples of such generic land-use models are Sleuth, (Clarke et al., 1996), CLUE (Verburg and Overmars, 2009) and Metronamica (Van Delden and Hurkens, 2011). The implementation of the conceptual model into computerized code can be verified, as it is possible to exactly check each component from the conceptual model and verify if it is implemented correctly (Sargent, 1998). Therefore, verification essentially is a software engineering challenge. For example the implementation of a random number generator can be very important in models that are sensitive to extreme events.

A **model application** is an application of the computerised model that includes the data and parameters for a particular case study. The process of adjusting the parameters to improve model results is called calibration. Calibration of land-use models is by definition site specific, because case study areas face different types of land-use change caused by different drivers (Silva and Clarke, 2002). Additionally, data for different applications is typically not strictly comparable in terms of its spatial, temporal or thematic resolution. The process of assessing the quality of a model result is called operational validation. A calibrated and validated model application is ready for use in experimentation. This experimentation relates back to the original problem entity: for example, the model application can be used to study a particular type of land-use change process, or to analyse specific land-use policies.

Calibration, operational validation and experimentation are highly dependent on the available data. Consequently, data quality can limit the possibilities and influence the outcome of all three processes. Hence, before any modelling exercise takes place, data quality should be considered. After all, a model cannot be better than the quality of the data that is used as input.

The steps in model development are sequentially described here. In reality however, iteration is a key aspect of model development and model validation (Balci, 1997; Jakeman et al., 2006). For example, during calibration it might be found that the conceptual model needs adjustments to improve the model (Rykiel, 1996).

1.2 What is valid?

Several authors have advocated that in fact some models cannot be validated. For example Konikow and Bredehoeft (1992) state that groundwater models cannot be validated, while Oreskes et al. (1994) point at problems in the validation of numerical models in the earth sciences. Their arguments are mostly applicable to land use models. The main issues raised are the following:

- It is not possible to demonstrate the truth of any proposition, except for a closed system. Take for example the proposition that an increase in population would require an increase in agricultural area to produce food. Now real world data shows an increase in population, but this is not matched by an increase in agricultural area. It turns out that some technological improvement made it possible to produce more food on the same amount of land. While the proposition was valid under the initial circumstances, it was not closed for technological developments. In general, land-use changes, like most other real world systems are not closed systems.

- A second problem is the scalability of non-additive systems. For example a land-use model partitions space in cells, and therefore model parameters typically represent drivers at the scale of a cell. However, in reality, changes on that scale might be the result of processes that act on a smaller scale, such as the behaviour of inhabitants. The relation between the scale of the process and the scale of the representation of the process is often unknown. For some processes observations are available at the appropriate scale, but in practice this is almost certainly not the case for all properties.
- Land-use models are typically numerical models that approximate the real land-use system. Solutions of such models are typically not unique; more than one combination of parameters can obtain the same solution. Moreover, errors in parameterisation can cancel each other out. Therefore, it is often impossible to find what has actually occurred during a simulation. Validation of individual parameters is only possible for parameters that are established independently, which is impossible for most if not all parameters in land-use models.
- Models are used for experimentation beyond the calibration period and over a time span that is far longer than the validation period. Errors that are not visible within the short time span of the validation period might grow incrementally and cause significant deviations in the experimentation results. In addition, there is no guarantee that future conditions will be similar to those in the validation period.

In this thesis however, both conceptual validation and operational validation are not used as philosophical terms as but as technical terms, such as proposed by Power (1993) and Rykiel (1996). Kleindorfer et al. (1998) provide a metaphor from justice to illustrate this interpretation: in court you do not need to prove that someone is guilty in a foundational sense. You just need proof beyond reasonable doubt. The validation of a land-use model is therefore not an activity that tests if a model is perfect, but an assessment how well it performs for the intended purpose. Moreover, model validation will not yield a binary outcome (Constanza, 1989; Kleindorfer et al., 1998). Increased testing will improve the insight into a model's quality, while the validity of a model increases as the model performance improves (Balci, 1997). As a result, validation is an open ended challenge without a clear ending (Aumann, 2007) and knowledge obtained from such models is always provisional. Passing the validation is at best an indication that the model is good enough for its intended use (Rykiel, 1996; Balci, 1997; Jakeman et al., 2006), which means that a minimum acceptable level or a benchmark needs to be established (Clarke 2004; Manson, 2007; Hagen-Zanker and Lajoie, 2008).

1.3 Operational calibration and validation of land use models

Many land-use models aim to explore future land-use changes, as part of a scenario study or a policy impact assessment (Veldkamp and Lambin, 2001; Verburg et al., 2004c; Sieber et al., 2010). For these purposes, land-use models are mostly calibrated to reproduce known historic land-use patterns, or land-use changes (Rykiel 1996, Silva and Clarke, 2002). Ideally, when sufficient data is available, the model is validated independently (Kok et al., 2001; Batty and Torrens, 2005). Independent refers to the fact that the data used for validation was not used for calibration. An independent validation is more likely to reveal biased models or overfitting (Balci, 1997; Sargent, 1998). For example, the model is calibrated from T1 to T2, and then validated from T2 to T3, as shown in Figure 1.2.

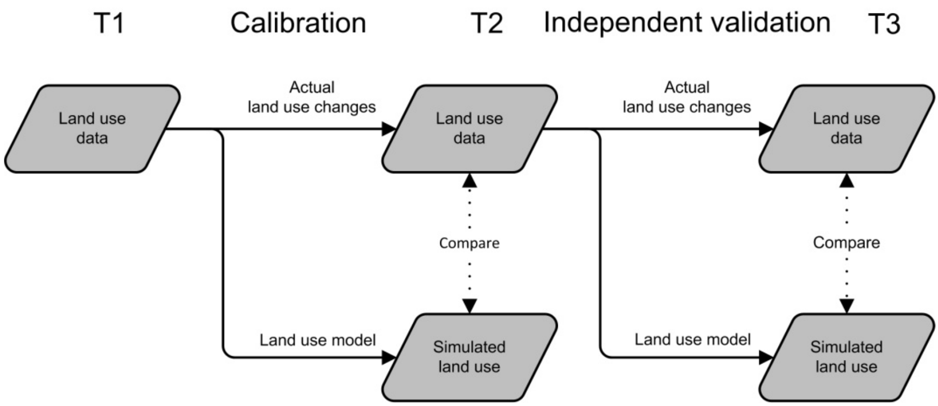


Figure 1.2: *Calibration and independent validation of a land-use model.*

The main question remains how to assess the results of a land-use model in this calibration and validation procedure. Several authors have indicated that land-use change processes are inherently uncertain, particularly since human decisions that drive these changes are rarely deterministic (Clarke, 2004; Brown et al., 2005; Manson, 2007). This property is reflected in many types of land-use models, such as random utility models, Markov models, cellular automata models, and agent based models as they typically include stochastic processes to simulate land-use changes. Consequently, each simulation will yield a different result and these results can never be expected to be perfect when compared to empirical data (Kleindorfer et al., 1998).

Additionally, many land use changes are not the results of one simple process. Instead, land-use changes are typically caused by a combination of biophysical and socioeconomic drivers that are mutually influential (Lambin et al., 2001). Moreover, land-use changes are often the combined result of many local actors, that are also mutually influential and that together shape global land-use patterns. The combination of the inherent uncertainty in land-use change processes and the feedback among these processes and actors make land-use change a path dependent process (Arthur, 1990; Krugman, 1991; 1998; Brown et al., 2005; Batty and Torrens, 2005; Torrens, 2011). Hence, land-use change processes can be characterised as complex, which yield emerging patterns and possibly bifurcations, making the assessment of land-use models not straightforward. The exact state at the local level of such a complex system cannot be known, but the patterns at the global level do show regularities (Manson, 2001; Batty and Torrens, 2005; Manson, 2007). Consequently, land-use models must be able to simulate complex processes, in order to represent the richness in behaviour that land-use change processes exhibit (Clarke, 2004).

Due to the inherent uncertainty and complexity of land-use change processes, it is not appropriate to validate land-use models only on thier capacity to reproduce historic land-use changes accurately, because this will cause an over calibration at the cost of realism (Kok et al., 2001). Instead a more comprehensive validation approach is required that assesses whether a model is accurate as well as realistic (Hagen-Zanker and Martens, 2008; Torrens, 2011). Brown et al. (2005) suggest the terms predictive accuracy and process accuracy for these two ways to assess land-use models.

Predictive accuracy answers the question whether land-use changes are allocated in the correct location. For instance, it can be measured from a pixel wise comparison of the simulated land-use map and the actual land-use map at T2. An overview of these methods is provided by Turner et al. (1989) and Couto (2003). More recently a number of methods have been presented that account for near-hits, using a fuzzy set approach (Hagen, 2003; Hagen-Zanker, 2009) or a multi resolution approach (Constanza, 1988; Pontius et al., 2004a).

Process accuracy on the other hand, assesses whether land-use changes are simulated realistically. This question can be answered by evaluating the model structure and equations, which are part of the conceptual validation. It can also be assessed indirectly from the aggregate statistics of the complete map. However, it has been mentioned that a realistic pattern in itself is not necessarily an indication of the quality of the underlying process (Manson, 2007). Visual inspection provides a first impression of the quality of the model results; however, measurements are preferable as they are objective and repeatable. Two types of pattern measures are currently being used. First, landscape metrics are a group of metrics that have their origin in landscape ecology (McGarigal and Marks, 1995, Riiters et al., 1995). This approach has been applied for example in Clarke et al. (1997) Another group of metrics find their origin in complexity science and describe complex patterns in the land use. Examples are the fractal dimension of patches (Batty and Longley, 1994), Zipf's law (Gabaix, 1999), the radial dimension and the urban cluster-size distribution (White, 2006).

Reviews actually show that few land-use models are validated at all (Agarwal et al., 2000; EPA, 2000). Other authors have noticed that developments in model calibration and validation have not kept pace with model developments (Grimm, 1999; Verburg et al., 2004c; Gardner and Urban, 2005; Auman, 2007). Consequently, there is no agreement on a set of methods to assess the results of land-use models (Silva and Clarke, 2002) and the available methods do not satisfy all demands. For instance, methods to assess the predictive accuracy typically asses the accuracy of simulated land use instead of simulated land-use changes. As a result such methods favour models that simulate little or no changes, while from a modelling point of view, no change is typically not a relevant model or not a realistic alternative since the topic of the model are precisely the changes (Pontius et al., 2004b). Methods that asses the process accuracy face a similar limitation, as they are often applied without an appropriate reference, which makes it impossible to gauge the true quality of a land-use model. Moreover, the fact that land-use patterns can be characterised using a specific metric does not necessarily mean that this characterisation is meaningful. Subsequently, there is a challenge to develop and apply more appropriate methods to assess land-use models.

1.4 Objectives of this study

The main objective of this study is to investigate the calibration and validation of land-use models. The Metronamica land-use model (Van Delden and Hurkens, 2011) will be used as the case study model for this research. Metronamica is a constrained cellular automata model (White and Engelen, 1993; 1997), and has been used in many integrated policy support systems (i.e. White and Engelen, 2000; Engelen et al., 2003; Van Delden et al., 2007; 2008; 2010). The core of constrained cellular automata models consist of the neighbourhood rules, which describe the influence of the existing land-use pattern on the allocation of new land uses. Although the existence of the neighbourhood effect has been acknowledged (Verburg et al., 2004b), their quantification remains an open question (Hagoort et al., 2008; De Nijs and Pebesma, 2010). Therefore, the calibration of neighbourhood rules will receive specific attention in this study.

The objective of this study is further specified in four research questions that each deal with one aspect of the calibration and validation of land-use models:

1. What characteristics of land-use models are important for assessing these models?
2. How can the predictive accuracy of a land-use model be assessed?
3. How can the process accuracy of a land-use model be assessed?
4. How can the neighbourhood rules in cellular automata land-use models be calibrated and validated?

Answers to these research questions will allow development more accurate models, and thereby improve our understanding of land-use changes as well as informed decision making based on scenario studies or policy assessments using land-use models (Sieber, 2010; Van Delden et al., 2010).

1.5 This thesis

The rest of this thesis is organized as follows:

Chapter 2 describes Kappa Simulation, a measure to assess the predictive accuracy of land-use models. Kappa Simulation is similar in form to the more frequently used Kappa statistic, but it applies a more appropriate stochastic model of random allocation of class transitions as a reference model. This effectively corrects for the amount of persistence in land-use changes and therefore provides a better indication of the predictive accuracy of a land-use model.

Chapter 3 presents Fuzzy Kappa Simulation, which is similar in form to Kappa Simulation but uses fuzziness in the interpretation of class transitions and locations. This allows to distinguish between near-hits and complete misses while assessing the predictive accuracy of land use models.

Chapter 4 describes the application of a variable grid cellular automata model for the city of Vancouver, Canada. The model applied in this chapter includes the entire study area in the neighbourhood effect and was calibrated and validated based on the predictive accuracy and process accuracy.

Chapter 5 elaborates on the neighbourhood effect, the core of cellular automata land-use models. This chapter presents an automated method to derive neighbourhood rules from observed land-use changes. Results from this method are used to simulate land-use changes, and the generated land-use maps are assessed in terms of their process accuracy and their predictive accuracy.

Chapter 6 introduces activity based land-use modelling. The activity based land-use model is an extension to constrained cellular automata land-use models, in that it can include dynamics in jobs or population separately from the land use it is associated with. Because actual data was missing, the model was calibrated and validated on synthetic data instead. This allowed for an assessment of the process accuracy only.

Chapter 7 provides a synthesis of the work that was performed within this study and places its results in the wider context of land-use modelling. It discusses how chapters 2 to 6 contribute to research questions as stated in this chapter and what topics still remain to be elucidated through further investigation.

2. Revisiting Kappa to account for change in the accuracy assessment of land-use models

Van Vliet, J., Bregt, A.K., and Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land-use models. Ecological Modelling 222 (8): 1367-1375. (Slightly adapted in order to improve consistency with other chapters).

Abstract

Land-use change models are typically calibrated to reproduce known historic changes. Calibration results can then be assessed by comparing two datasets: the simulated land-use map and the actual land-use map at the same time. A common method for this is the Kappa statistic, which expresses the agreement between two categorical datasets corrected for the expected agreement. This expected agreement is based on a stochastic model of random allocation given the distribution of class sizes. However, when a model starts from an initial land-use map and makes changes to it, that stochastic model does not pose a meaningful reference level. This paper introduces Kappa Simulation ($K_{Simulation}$), a statistic that is identical in form to the Kappa statistic but instead applies a more appropriate stochastic model of random allocation of class *transitions* relative to the initial map. The new method is illustrated on a simple example and then the results of the Kappa statistic and $K_{Simulation}$ are compared from the results of a land-use model. It was found that only $K_{Simulation}$ truly tests models in their capacity to explain land-use changes over time, and unlike Kappa it does not inflate results for simulations where little change takes place over time.

2.1. Introduction

Land-use models have been used to analyse of land-use dynamics, and increasingly to support spatial policy making (Verburg et al., 2004c; Sieber et al., 2010). This development is fuelled by an increase in spatial data and modelling tools. For scientific use as well as policy analysis it is important to assess the predictive accuracy of the results of land-use models and to be aware of their implications.

Visual inspection by experts is arguably the best way to assess simulation results (Hagen, 2003; Pontius and Malanson, 2005). Unfortunately this is highly subjective and not reproducible. Consequently, there is a need for methods to assess simulation results that are both objective and reproducible (Power et al., 2001; Hagen, 2003). Currently a multitude of methods are available for the assessments of simulation results (Turner et al., 1989; Couto, 2003), most of which compare simulated land-use maps with actual land-use maps.

However, many types of land-use models simulate land-use changes starting from an original land-use map, such as cellular automata models (van Vliet et al., 2009), Markov models (Riiters et al., 2009), and logistic regression models (Dendoncker et al., 2007), some of which are also used as parts of larger integrated systems (Van Delden et al., 2010). Since most locations do not change their land use over the length of a typical simulation period, the similarity between the simulated land-use map and the actual land-use map will be high for most calibrated models, regardless of the accuracy of simulated changes.

To rigorously assess the accuracy of the simulated land-use map, a meaningful reference level is required (Hagen-Zanker and Lajoie, 2008). The reference level for land-use models that start from an original map should therefore account for the information from this original map. In this paper we present Kappa Simulation ($K_{Simulation}$), a method that assesses the agreement between the simulated land-use map and the actual land-use map, adjusted for the information contained in the original land-use map. The method described is equally appropriate for gauging the accuracy of other types of simulation models, as long as the simulation starts from original values and simulates changes in categorical values.

The next section gives an overview of existing methods to assess the predictive accuracy of land-use models by means of map comparisons. Section 2.3 then derives $K_{Simulation}$ as a modification of the Kappa statistic, by adjusting the expected agreement for the information from the original land-use map. Section

2.4 demonstrates this method first with a simple example and then illustrates the interpretation of $K_{Simulation}$ with calibration results from a land-use model. Section 2.5 draws conclusions on the usage and implication of this new method.

2.2 Accuracy assessment of land-use models

2.2.1 Calibration and validation

The application of any land-use model to a specific region requires model calibration, where calibration is essentially the adjustment of parameters to improve the model's goodness of fit (Rykiel, 1996). Land-use models are typically calibrated to simulate known historical land-use changes. This procedure requires two land-use maps: one for the start of the simulation period (T1) and one for the end of the simulation period (T2). The simulated land-use map at the end of the calibration period can then be compared with actual land-use map at T2.

Independent validation requires at least a third map at time T3 that has not been used in the calibration procedure. The simulation result at T3 can then be compared with the actual land-use map at T3. It should be noted that the term validation is used here to refer to the process of assessing the quality of the parameter set with an independent data set. For other interpretations of model validation, both as a process and as a judgement can be found in Power (1993), Oreskes et al. (1994), Rykiel (1996), Sargent (1998), Pontius et al. (2004a) and Refsgaard and Henriksen (2004).

Both calibration and validation essentially comprise the same activity as the result of the simulation model is compared with an actual land-use map for the same moment. Hence both procedures require an assessment of the similarity between the two maps, only for different reasons. In calibration the purpose of the assessment is to evaluate the current parameters and look for improvements, while in the validation the assessment is to evaluate the quality of these parameters using an independent data set.

There are many aspects of maps that can be assessed, on the level of the data element, the data set and the spatial structure. Brown et al. (2005) indicate that results of land-use models can be assessed with respect to the allocation of land-use changes as well as the generated patterns, which they relate to predictive accuracy and process accuracy, respectively. A similar but more comprehensive view on this aspect has been offered by Hagen-Zanker and Martens (2008). They argue that land-use maps can be compared at three spatial scales, the local, the focal and the global scale, and with respect to presence and structure. The latter division roughly coincides with predictive accuracy and process accuracy, respectively.

2.2.2 Methods to assess results of land-use models

Predictive accuracy assesses the agreement between land uses in the simulated land-use map and the actual land-use map, typically based on a pixel by pixel comparison. One of the most commonly used methods for this is the Kappa coefficient of agreement (Cohen, 1960), alternatively known as the Heidke Skill Score (Heidke, 1926). It measures the agreement between two categorical datasets relative to the agreement that can be expected by chance, to avoid deflation or inflation of the perception of agreement (Doswell et al., 1990). The agreement that can be expected by chance is the agreement that is expected when the given sizes of classes are reallocated randomly.

Since land-use maps are categorical datasets, Kappa can be used to compute the agreement between a pair of land-use maps. Therefore it is frequently used for accuracy assessment of remote sensing image classifications (Foody, 2002; Wilkinson, 2005) and results of spatial simulation models (Monserud and Leemans, 1992; Hagen-Zanker and Martens, 2008). However, some authors argue that Kappa is not the appropriate measure for accuracy, see for example Allouche et al. (2006).

Others argue that the exact allocations of land-use changes cannot be predicted because they consider land-use change as a complex process (Batty and Torrens, 2005; Manson, 2007). Due to feedback loops, existing land-use patterns are found to be path dependent and therefore small variations in the original situation can develop into considerable variations in the final situation (Brown et al., 2005). However, the land-use patterns that are generated by this complex process have regularities that can be measured. Comparison of these regularities in the simulated land-use map with those from the actual land-use map can be used to assess the process accuracy of a model. Examples of such measures are landscape metrics which are frequently used in the assessment of simulation results (Turner et al., 1989; Wear et al., 1998; Power et al., 2001). In addition, urban modellers use fractal properties of patches (Batty and Longley, 1994) and the distribution of cluster sizes (White, 2006) to assess patterns of urban land use specifically.

2.2.3 Limitations of end-state assessment

Land-use models typically simulate changes for periods ranging from years to decades. Over these periods most locations do not change land use. This persistence is well illustrated by a study of the nature of land-use changes by

Pontius et al. (2004b). Accordingly, models will leave most cells in the same land-use category, as can be seen in land-use models that derive transition probabilities from historic land-use changes (Rutherford et al., 2008; Riiters et al., 2009). The probability of a location to maintain its original land-use class is typically very high in these models. As a consequence the agreement between the simulated land-use map and the actual land-use map will be high. However, this high agreement does not necessarily indicate an accurate model of change.

In the extreme, the agreement between a simulation result and reality for an area with only a few land-use changes will be high, even when all simulated changes are incorrect. Thus, the end-state agreement by itself is meaningless as a measure for model accuracy (Hagen-Zanker and Lajoie, 2008). Moreover, since the amount of change can vary considerably between applications of land-use models, their relative merits cannot be assessed on the basis of end-state agreement alone, because a reference level is missing.

In this respect the interpretation of results of land-use models is fundamentally different from the results of remote sensing image classification. In the latter the agreement between the classified land-use data and actual land-use data can be used as an indicator of the accuracy (Thomlinson et al., 1999), although it depends on the specific case what level of agreement is considered acceptable (Foody, 2008). In the evaluation of land-use models however, the agreement between the model result and the actual land-use data is meaningless when the amount of change is not considered.

There is some recent work that seeks to integrate the amount of land-use change in the interpretation of end-state agreement. Pontius et al. (2004a) and Pontius and Malanson (2005) use the original land-use map as a benchmark to compare with model results. The no-change model in their study has a higher accuracy in the original resolution than the result of their land-use model, even though it does not simulate any change at all. They coarsen the resolution of all three land-use maps by aggregation until the simulation result is more accurate than the no-change model and thus find the “null resolution”. Pontius et al. (2008) introduce the “figure of merit”, which assesses the agreement of land-use changes rather than just the land uses. Specifically, it measures the ratio of the intersection of the observed change and predicted change to the union of the observed change and predicted change. Still, because this statistic does not include a reference level it is not possible to interpret the absolute value of the figure of merit and results of different models cannot be compared.

Chen and Pontius (2010) elaborate further on the use of the original land-use map as they subdivide results into (a) land-use persistence that is simulated correctly, (b) land-use change that is simulated correctly, (c) land-use change that is simulated as persistence, (d) land-use persistence that is simulated as change, and (e) land-use change that is simulated as change into the wrong class. This subdivision is useful for the assessment of model results, because it informs on the type of errors in the simulation result as well as the number of hits. However, given two simulation results, it does not inform which of the two is more accurate as that requires some balancing between hits and misses.

Alternatively Hagen-Zanker and Lajoie (2008) propose the use of neutral models in the evaluation of simulation results. Particularly they propose a random constraint match (RCM) model to create land-use maps that comprise the correct distribution of land-use class sizes, allocated randomly over the original land-use map with minimal adjustments to this original land-use map. This method creates reference maps that can be used as a benchmark for the interpretation of model results. Still, this procedure is rather indirect, as it incorporates patterns of change in the interpretation of agreement but not in the measurement of agreement itself. Instead of interpreting the meaning of end-state comparisons for simulation results at length, we would prefer to have a comparison method that directly measures the agreement of changes.

2.3 Kappa Simulation

2.3.1 The Kappa coefficient of agreement

The Kappa coefficient of agreement is originally a statistic for discrete multivariate analysis. It expresses the agreement between two categorical datasets corrected for the agreement as expected by chance, which depends on the distribution of class sizes in both datasets only. In terms of map comparisons this can be interpreted as the expected agreement when the given class are allocated randomly over the map.

Kappa can be computed from the contingency table between two datasets. Table 2.1 gives the generic form of a contingency table from the comparison of actual land-use data, represented in map A with simulated land-use data, represented in map S. Land-use classes are indicated as $i = 1, 2, \ldots, c$. Elements $P(a = i \wedge s = j)$ in the table indicate the fraction of cells that have land use i in map A and land use j in map S and hence elements $P(a = i \wedge s = i)$ on the diagonal indicate cells that have the same land use in both maps. Row and column totals, $P(a = i)$ and $P(s = i)$, indicate the fraction of cells that have land use i in map A and map S, respectively.

Table 2.1: *Generic form of a contingency table.*

	Map S categories				
	1	2	...	c	Total map A
Map A categories					
1	$P(a = 1 \wedge s = 1)$	$P(a = 1 \wedge s = 2)$	...	$P(a = 1 \wedge s = c)$	$P(a = 1)$
2	$P(a = 2 \wedge s = 1)$	$P(a = 2 \wedge s = 2)$	...	$P(a = 2 \wedge s = c)$	$P(a = 2)$
...	...	...		...	...
c	$P(a = c \wedge s = 1)$	$P(a = c \wedge s = 2)$	...	$P(a = c \wedge s = c)$	$P(a = c)$
Total map S	$P(s = 1)$	$P(s = 2)$	...	$P(s = c)$	1

From the fractions as indicated in the contingency table one can compute the observed fraction of agreement, PO, the expected fraction of agreement given the distribution of class sizes, PE, and the maximum fraction of agreement given the distribution of class sizes, $PMAX$:

$$PO = \sum_{i=1}^{c} P(a = i \wedge s = i) \qquad \text{Equation 2.1}$$

$$PE = \sum_{i=1}^{c} P(a = i) \cdot p(s = i) \qquad \text{Equation 2.2}$$

$$PMAX = \sum_{i=1}^{c} min(P(a = i), p(s = i)) \qquad \text{Equation 2.3}$$

The observed fraction of agreement and the expected fraction of agreement are required to compute the Kappa coefficient of agreement. Cohen (1960) also indicates that the maximum obtainable agreement depends on the distribution of the class sizes and that therefore dissimilarities can be caused by disagreement in class sizes as well as disagreements in allocation, given these class sizes. In terms of map comparisons, this decomposition can be interpreted as a disagreement due to the amount of cells per land-use class, and dissimilarities due to the allocation of land uses given these class sizes. Hagen (2002) identifies these two components of Kappa in the context of map comparisons as $K_{Histogram}$ and $K_{Location}$:

$$Kappa = \frac{PO-PE}{1-PE} \qquad \text{Equation 2.4}$$

$$K_{Histogram} = \frac{PMAX - PE}{1 - PE}$$ Equation 2.5

$$K_{Location} = \frac{PO - PE}{PMAX - PE}$$ Equation 2.6

Hence Kappa is equal to the product of $K_{Histogram}$ and $K_{Location}$. Values for Kappa range from 1, indicating a perfect agreement, to −1 indicating no agreement at all. The value 0 represents the special case where the agreement is equal to the agreement that can be expected by chance. $K_{Histogram}$ can get values from 0 to 1, where 1 indicates a perfect agreement and 0 indicates that there is no agreement in the class sizes at all. $K_{Histogram}$ cannot be negative because $PMAX \geq PE$. $K_{Location}$ ranges from -1 to 1, where 0 indicates the agreement as can be expected by chance and 1 indicates an allocation which is as high as possible given the distribution of class sizes. It should be noted that $K_{Location}$ is undefined when $PMAX$ equals PE. However, this only happens in the case that either both maps have one and only one land use, or when each land use appears only in one of the two maps. In both cases the agreement between the allocation of land uses is not a meaningful measure, because this accuracy is already implicit in the distribution of class sizes.

2.3.2 Accounting for land-use persistence

Because the distribution of class sizes is not a meaningful reference level for models that start from an original land-use map, we would like to modify the Kappa statistic by integrating the amount of land-use changes in the expected agreement. This can be achieved by considering the distribution of class transitions, which can be interpreted as *conditional probabilities*; the chance of finding a certain class at a location will depend on the class that was originally there.

To compute the expected agreement as a function of class transitions, we need to express the size of class transitions as a function of the original land-use map and the simulated or actual land-use map. We express the fraction of cells that changed from land use j in the original map to land-use i in the simulated land-use map A as $P(a = i \mid o = j)$ and the fraction of cells that changed from land use j in the original map to land-use i in the actual land-use map S as $P(s = i \mid o = j)$. Because the original land-use map (O) is the same for both the simulated and actual land-use changes, the expected agreement between the simulated land-use map and the actual land-use map can be expressed as follows:

$$PE_{Transition} = \sum_{j=1}^{c} P(o = j) \cdot \sum_{i=1}^{c} p(a = i|o = j) \cdot p(s = i|o = j)$$

Equation 2.7

where $PE_{Transition}$ is the expected fraction of agreement, given the sizes of the class transitions. Similarly, the maximum obtainable agreement can be expressed as a function of land-use transitions:

$$PMAX_{Transition} = \sum_{j=1}^{c} P(o = j) \cdot \sum_{i=1}^{c} min(P(a = i|o = j), P(s = i|o = j))$$

Equation 2.8

where $PMAX_{Transition}$ is the maximum accuracy that can be achieved given the sizes of the class transitions.

Kappa Simulation ($K_{Simulation}$) and its components can be computed similarly to equations 2.4 to 2.6:

$$K_{Simulation} = \frac{PO - PE_{Transition}}{1 - PE_{Transition}}$$ Equation 2.9

$$K_{Transition} = \frac{PMAX_{Transition} - PE_{Transition}}{1 - PE_{Transition}}$$ Equation 2.10

$$K_{Transloc} = \frac{PO - PE_{Transition}}{PMAX_{Transition} - PE_{Transition}}$$ Equation 2.11

Where $K_{Simulation}$ expresses the agreement between the simulated land-use transitions and the actual land-use transitions. Similarly, $K_{Transition}$ expresses the agreement in the quantity of land-use transitions and $K_{Transloc}$ expresses the degree to which the transitions agree in their allocations.

Note that if the distribution of classes is independent from the original situation $P(a = i|o = j)$ equals $P(a = i)$, such as for remote sensing image classification which starts with an empty map, $PMAX_{Transition}$ becomes equal to $PMAX$ and therefore $K_{Transition}$ turns into $K_{Histogram}$ and $K_{Transloc}$ turns into $K_{Location}$. In that sense $K_{Simulation}$ is an extension of Kappa to be used when the initial land-use map is part of the model.

Values for $K_{Simulation}$ range from -1 to 1, with 1 indicating a perfect agreement, and 0 indicating the special case where the agreement is as good as can be expected from a random distribution of the given class transitions. Scores below 0 indicate that class transitions are less accurate than can be expected from a random allocation of the given class transitions. As many land-use changes are not random, values for simulation results will typically be above 0, and any score below 0 can be understood as a model that does not explain any land-use changes. $K_{Simulation}$ values above 0 can therefore be interpreted as how much more accurate than random a simulation explains land-use changes.

$K_{Transition}$ has values between 0 and 1, where 1 indicates the case where the sizes of class transitions in the simulation are exactly in agreement with the sizes of class transitions in reality, and the value of 0 indicates that there are no class transitions that appear in the simulation as well as in reality. $K_{Transloc}$ values range from -1 to 1. Here 0 indicates the agreement as can be expected by chance and 1 indicates an allocation which is as high as possible given the distribution of class transitions. Values below 0 indicate the case where the allocation of class transitions is worse than can be expected by random allocation of the given class transitions.

Similarly to $K_{Histogram}$, $K_{Transloc}$ is not defined for the case where $PE_{Transition}$ equals $PMAX_{Transition}$. This is the case under three circumstances: when all locations in reality and in the simulation change into one and the same class, when there is no transition that appears in reality and in the simulation, or when there are no land-use changes in either reality, or the simulation or both. The first is a highly unlikely case, but as the results are in perfect agreement there is no use for decomposition anyway. The second case can not have any agreement in the allocation of land-use changes, since there are no coinciding land-use changes. The third finally cannot have agreement in the allocation of land-use changes, because there are no changes.

***Table 2.2:** Overview of Kappa, $K_{Simulation}$, and their decomposition.*

Statistic	Agreement between	Corrected for
Kappa	two categorical datasets	the expected agreement given the distribution of class sizes
Kappa Histogram ($K_{Histogram}$)	the distribution of class sizes	the expected agreement given the distribution of class sizes *
Kappa Location ($K_{Location}$)	two categorical datasets	the maximum agreement given the distribution of class sizes
Kappa Simulation ($K_{Simulation}$)	two categorical datasets	the expected agreement given the distribution of class transitions from the same initial dataset
Kappa Transition ($K_{Transition}$)	the distribution of class transitions from the same initial dataset	the expected agreement given the distribution of class transitions from the same initial dataset*
Kappa Transition Location ($K_{Transloc}$)	two categorical datasets	the maximum agreement given the distribution of class transitions from the same initial dataset

* $K_{Histogram}$ and $K_{Transition}$ are corrected for the expected accuracy with the specific purpose of making them consistent with Kappa, respectively $K_{Simulation}$.

Kappa and all variations that are discussed in this paper express the agreement between two categorical datasets, and correct for a reference agreement. However, they differ in what agreement is expressed and what reference agreement it corrects for. Kappa, $K_{Histogram}$ and $K_{Location}$ are measures that express the agreement between two maps. $K_{Simulation}$, $K_{Transition}$ and $K_{Transloc}$ are measures that assess the agreement between land-use transitions and are therefore a replacement of Kappa that can be applied to assess results of simulation models. Table 2.2 presents an overview of Kappa, $K_{Simulation}$ and their components that are discussed in this paper. All comparisons methods are implemented in the Map Comparison Kit (Visser and de Nijs, 2006), version 3.2.1 and higher. The Map Comparison Kit is a freely available software tool designed to compare categorical maps which can be downloaded from www.riks.nl/mck.

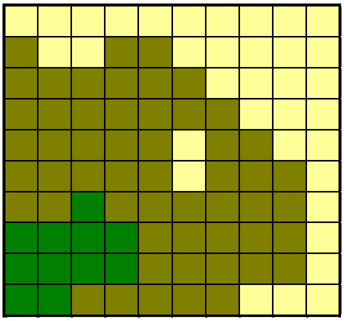
2.4 Results and discussion

2.4.1 A simple example

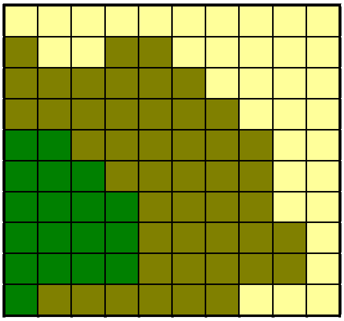
A simple example illustrates the use of Kappa Simulation ($K_{Simulation}$). Consider land-use maps O and A as shown in Figure 2.1. They represent the original and actual land-use maps at T1 and T2. The land-use changes that appear are an increase in the forested area, a small patch of arable land that changes into extensive grasslands and another small patch of extensive grassland that changes into arable land. Maps S1–S4 represent four different simulation results, also at T2, and maps C1–C4 represent the results of the comparison of these simulation results with the actual land use map A. The simulation results are constructed to show the effect of different types of errors.

Map S1 shows the result of the no-change model, hence the map is identical to the original land-use map. The comparison yields cells that are correctly simulated as persistence, and cells that are incorrectly simulated as persistence. Map S2 shows a simulation result where the area for forest decreases instead of increases. Hence the model simulated the wrong type of land-use transitions. Assessment of these land-use changes shows cells correctly simulated as persistence, cells incorrectly simulated as persistence, and also cells that are simulated as change into the wrong land-use class. Map S3 shows a simulation result that has exactly the right number of cells per land-use type, except that most of the increase or decrease per land-use is allocated in the wrong cells. Therefore the comparison includes cells that are correctly simulated as persistence, cells that are incorrectly simulated as persistence, cells that incorrectly simulated as change, and a cell that changed correctly. Map S4 finally shows a simulation result where the forest increases, but not as much as in the actual land-use changes. Comparison therefore yields cells that are correctly simulated as persistence, cells that are incorrectly simulated as persistence, and cells that are correctly simulated as change.

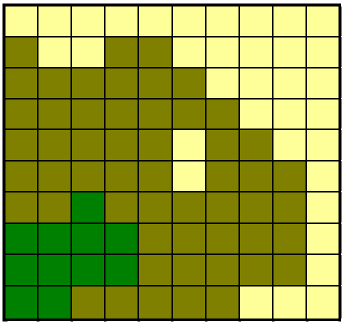
Table 2.3 presents the Kappa values for the comparison of the four simulation results with the actual land-use maps at T2. All scores are closer to 1 than to 0, which suggests that all simulations are fairly accurate. A decomposition of these Kappa scores into $K_{Location}$ and $K_{Histogram}$ gives only little indication of the nature of the errors that are made in the simulation. Because these values are relatively close to 1 they indicate that both the class sizes and the allocation of land uses is quite similar in the simulated and actual land-use maps. However, these scores do not inform on the accuracy of the simulated land-use changes.



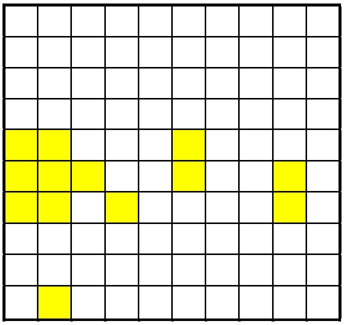
Map O: Actual land use at T1



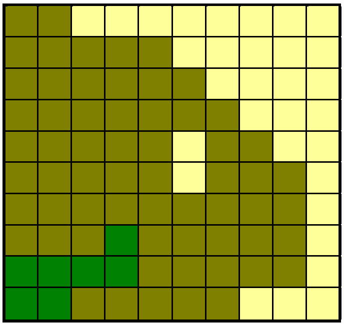
Map A: Actual land use at T2



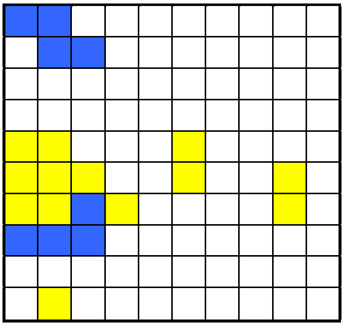
Map S1: Simulation result 1 at T2



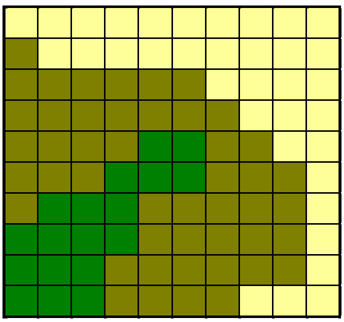
Map C1: Assessment of simulation result 1



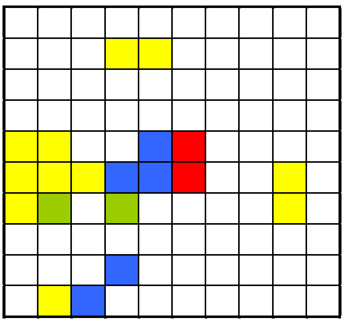
Map S2: Simulation result 2 at T2



Map C2: Assessment of simulation result 2



Map S3: Simulation result 3 at T2



Map C3: Assessment of simulation result 3

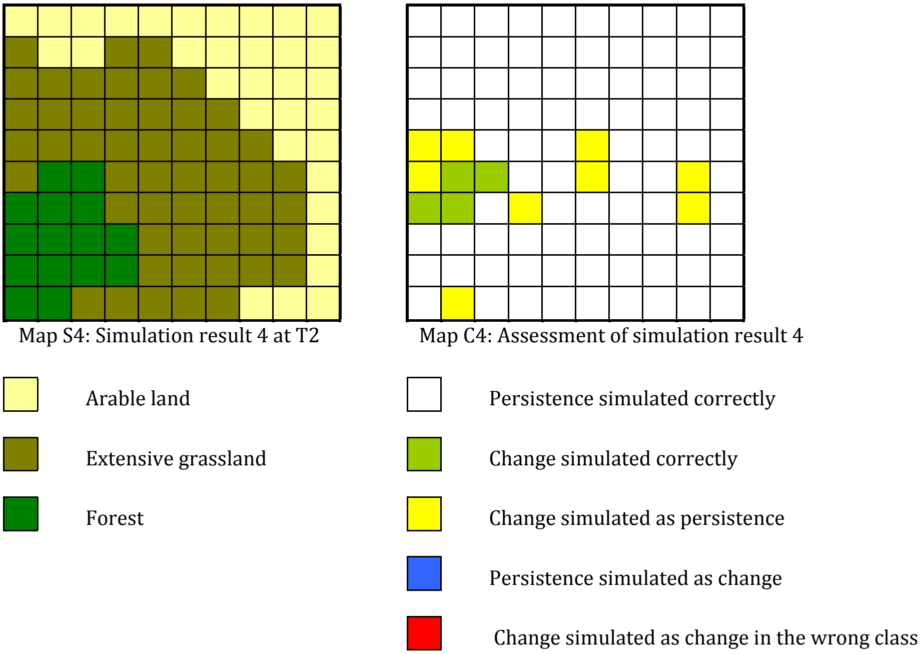


Figure 2.1: *Land-use maps and comparison results for the simple case study illustration of Kappa and $K_{Simulation}$. See section 2.4.1 for further explanation.*

Table 2.3 also presents the $K_{Simulation}$ scores for these four land-use simulations, which differ considerably from the original Kappa scores, because the expected agreement is corrected for the land-use that persists during the simulation period. Simulation model 1, the no-change model, yields a score of exactly 0. This indicates that the model doesn't explain any land-use changes. Simulation model 2 has a negative score for $K_{Simulation}$, and hence the accuracy of the simulated changes is worse than can be expected given the amount of class transitions. Simulation model 3 has a score that is little higher than 0, from which it can be concluded that it explains only some land-use changes. Simulation result 4 finally yields by far the highest score as it is closer to 1 than it is to 0, which indicates that this model does explain quite some land-use changes.

Decomposition of $K_{Simulation}$ scores into $K_{Transition}$ and $K_{Transloc}$ gives some useful information on the nature of the errors in the simulations. Simulation 1 does not simulate any changes, as expressed in a $K_{Transition}$ score of 0. Therefore these changes cannot be allocated either and hence the $K_{Transloc}$ is undetermined. $K_{Simulation}$ scores for simulation 2 are below 0, which indicates that the model explains no land-use changes. $K_{Transition}$ shows that this is due to the incorrect allocation of land-use changes, and not so much due to the types of land-use changes. Simulation 3 simulates the right type of change, but these changes are not allocated on the correct location. This is articulated in a high $K_{Transition}$ and a low $K_{Transloc}$. Results for simulation 4 indicate that the only errors are caused by incorrect amount of land-use transitions, while the allocation of all land-use transitions is completely correct.

Table 2.3: *Kappa scores and $K_{Simulation}$ scores obtained from the assessment of the results of the simple case study.*

	Comparison			
Statistic	Map A - Map S1	Map A - Map S2	Map A - Map S3	Map A - Map S4
Kappa	0.79	0.64	0.71	0.89
$K_{Histogram}$	0.88	0.75	1.00	0.92
$K_{Location}$	0.89	0.86	0.71	0.96
$K_{Simulation}$	0.00	-0.06	0.15	0.60
$K_{Transition}$	0.00	0.24	0.81	0.60
$K_{Transloc}$	n.a.	-0.25	0.19	1.00

A comparison of the Kappa scores with the $K_{Simulation}$ scores shows some remarkable differences. First, all Kappa scores are closer to 1 then to 0, which suggests that all model results are fairly accurate. However, $K_{Simulation}$ scores show that this is mostly due to the persistence of land uses, and not to the correct simulation of land-use changes, as three out of the four models yield scores close to 0. Simulation 4 is the exception as it does simulate several land-use changes correctly. Although it yields the highest score for both methods, only $K_{Simulation}$ indicates clearly that this model simulates land-use changes much more accurately than can be expected by chance, and that it outperforms the simulation results by a large margin.

Moreover, the ranking of the results differs between Kappa and $K_{Simulation}$. Kappa indicates that the result of simulation 1 is closer to the actual land-use at T2 than the result of simulation 3. However, $K_{Simulation}$ indicates that changes are simulated more accurately in simulation 3 as it ranks the two results in the reverse order. The reason for this is that Kappa doesn't account for the amount of land-use transitions, while $K_{Simulation}$ does. An incorrectly allocated pixel in a simulation result causes two disagreements. Land uses on both the location that incorrectly changes and the location that incorrectly persists do not correspond with the actual land-use map. On the other hand, the no-change model yields only one error per land-use change, which is on the location where the land use incorrectly persists. Therefore, Kappa generally favours models that generate less land-use changes over results that have the correct amount or more land-use change.

2.4.2 Assessment of a land-use model for Western Europe

To further demonstrate the utility of $K_{Simulation}$ it was used to assess the results for several land-use models applied to Western Europe. The modelled area comprises Ireland, Denmark, Germany, The Netherlands, Belgium, Austria and France. Land-use data was taken from the Corine Land Cover database (Haines-Young and Weber, 2006), for which maps were available for the years 1990 and 2000. The original 44 land-use classes were first reclassified into 17 aggregate land-use classes: natural vegetation, agriculture, residential, industry and commerce, tourism and recreation, forest, open spaces, infrastructure, port areas, airports, mineral extraction sites, dump sites, inland wetlands, marine wetlands, inland water, marine water, and beaches and dunes. The reclassified raster data was aggregated to cells of 1 km^2 each, using a majority aggregation. The original land-use map for the study area is shown in Figure 2.2.

We used three different models to simulate the known land-use changes from 1990 to 2000: the no-change model, the random constraint match (RCM) model and the Metronamica land-use model. The no-change model does not simulate any land-use change and therefore its result is identical to the initial land-use map from 1990. The RCM model is a neutral model; it allocates the correct amount of land use for the end year randomly over the original land-use map with minimal adjustments to the original land-use map (Hagen-Zanker and Lajoie, 2008). The size of the land-use classes is by definition similar in the result of the RCM model and the 2000 land-use map, while the allocation of land uses is not. It should be noted that the RCM model simulates the correct amount of net change, but not necessarily the correct amount of gross change, as it does not model land uses that have been interchanged. Metronamica is a constrained cellular automata land-use change model (Van Delden and Hurkens, 2011). In this model the total number of cells per land-use class is defined exogenously, while their allocation is computed using the cellular automata algorithm. Therefore the sizes of the land-use classes are exactly similar in the simulation result and the 2000 land-use map, but the allocation of the land-uses is not.

For all three models, the simulation results for the year 2000 were compared to the actual land-use map for the year 2000 to compute the Kappa statistics. In addition, the original land-use map for the year 1990 was used to compute the $K_{Simulation}$ scores for the results of the same three models. Results for both statistics are shown in table 2.4.

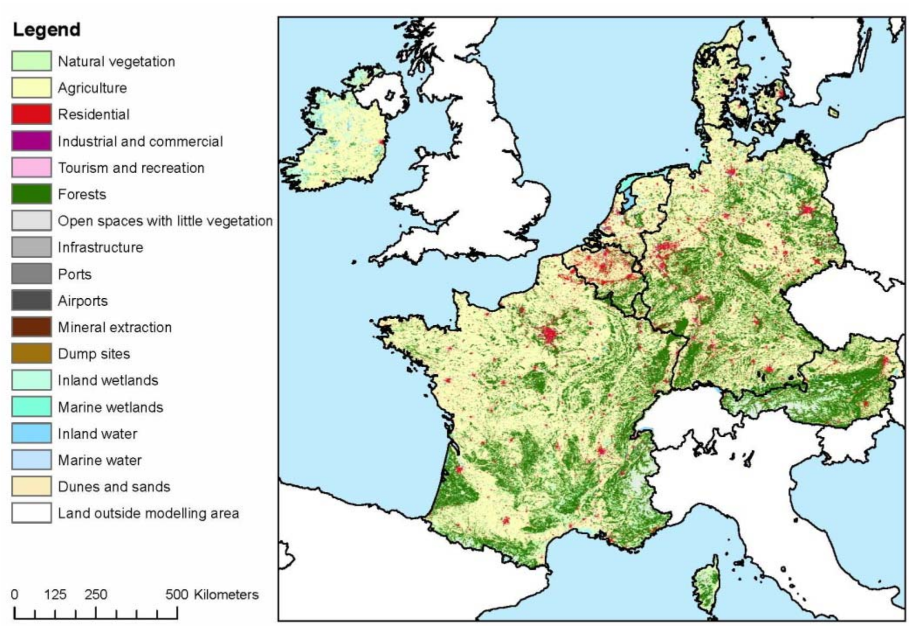


Figure 2.2: *Starting map for the simulation of land-use changes in Western Europe. National boundaries are included for spatial reference.*

2.4.3 Discussion of the results

Kappa values presented in table 2.4 are very close to 1 for all three comparisons. This indicates a very high agreement between all three simulated maps and the actual land-use map. It is not surprising that there were actually only a few land-use changes recorded over the simulation period. An analysis of land-use change in the Corine Land Cover database for 24 countries shows that more than 95% of the locations does not change land use between 1990 and 2000 (Haines-Young and Weber, 2006). Due to spatial aggregation and reclassification this persistence in land use is even higher when measured from the processed maps used in this study: only 1.8 % of the locations change land use over this period. Even if all land-use changes would be allocated incorrectly, the Kappa value would still be close to 1. Hence, the absolute value of Kappa does not give a good impression of the quality of the model, since it does not account for the amount of land-use change. $K_{Simulation}$ on the other hand accounts for the size of class transitions. Therefore $K_{Simulation}$ results will be similar for an application where a larger fraction of the cells change land use, if the model is equally accurate.

Table 2.4: *Kappa scores and $K_{Simulation}$ scores obtained from the assessment of the results of three different land-use models for Western Europe.*

	Comparison		
Statistic	No change - Actual land use	RCM model - Actual land use	Metronamica - Actual land use
Kappa	0.97	0.96	0.96
$K_{Histogram}$	0.99	1.00	1.00
$K_{Location}$	0.98	0.96	0.96
$K_{Simulation}$	0.00	0.00	0.26
$K_{Transition}$	0.00	0.45	0.71
$K_{Transloc}$	N.A.	0.00	0.37

Table 2.4 gives some insights in the nature of the disagreements between the maps. For both the RCM model and Metronamica, the number of cells per land use is exactly equal to that of the actual land-use map in 2000, due to the definition of the models and hence all errors are caused by an incorrect allocation of these land uses as expressed with a $K_{Location}$ lower than 1. Conversely, the no-change model result has both a $K_{Location}$ and $K_{Histogram}$ lower than 1 indicating that both the class sizes and the allocation of land uses have some disagreement when compared with the actual land-use map.

Table 2.4 also presents the results as assessed with $K_{Simulation}$. Because the scores are expressed relative to the value 0, they can be interpreted in absolute terms. The no-change model has a $K_{Simulation}$ score of exactly 0, which indicates that it does not explain any land-use change. The RCM scores are very close to 0 which shows that the RCM model can hardly explain any land-use changes either. This agrees with an intuitive interpretation of both models: the no-change model doesn't simulate any changes, while the RCM model has a complete random allocation of land-use changes. Metronamica on the other hand yields scores well above 0, but much lower than those obtained with Kappa. This indicates that a large part of the high Kappa scores is the result of land-use persistence, and not of accurately simulated land-use changes.

A more detailed look at $K_{Transloc}$ and $K_{Transition}$ indicates the character of the disagreements. The $K_{Transition}$ value of 0 for the no-change model indicates that the amount of land-use transitions does not explain anything. Since the no-change model does not simulate any changes, they cannot be located correctly either, which results in an undetermined $K_{Transloc}$. For the results of the RCM model, the scores can be explained a little differently. $K_{Transition}$ is above 0, indicating that the amount of simulated land-use changes have some similarity with the actual land-use changes. However, it is not equal to 1 as measured with $K_{Histogram}$. This is because the RCM model uses the net amount of increase and decrease per land-use class, but not the amount of land transitions. Hence the type of land-use transitions can differ, and land uses that have been interchanged are not considered. The value for $K_{Transloc}$ is almost equal to zero, which indicates that the allocation of land-use changes is about as good as can be expected by chance. This is exactly as expected given that allocation is random. Although the sizes of the land-use classes are defined endogenously to Metronamica, the scores show that $K_{Transition}$ is lower than 1. The reason for this is the same as for the RCM model. The type of land-use transitions and the amount of interchanged land uses do not necessarily correspond between the simulation and reality. Still both $K_{Transition}$ and $K_{Transloc}$ are well above 0, which indicates that Metronamica explains both the amount of land-use transitions and the allocation of these land-use transitions better than can be expected by chance, given the amount of land-use transitions.

It should be noted that a $K_{Simulation}$ score above 0 is not necessarily an indication that a simulation is good enough for the purpose of the modelling exercise. This method assesses the accuracy of simulated changes, and allows for a fair comparison of different model results. Yet it depends on the purpose of the modelling exercise what scores are considered "acceptable" or "good". However, a value below can be interpreted as an invalidation, as this indicates that a simulation is less accurate than random.

Similar to the simple case study, $K_{Simulation}$ yields a different ranking of the results compared to the Kappa results. Kappa results indicate that the no-change model result corresponds better with the actual land-use map than both the result of Metronamica and the RCM model even though it does not simulate any change at all. At the same time, $K_{Simulation}$ indicates that Metronamica yields the most accurate results. The reason for that is again that Kappa does not account for the amount of land-use persistence in the computation of the expected agreement, while $K_{Simulation}$ does. Hence Kappa generally prefers results with few land-use changes, because an incorrectly allocated land-use change yields two errors, while not allocating this change yields only one error.

As indicated in the introduction, we acknowledge that $K_{Simulation}$ does not capture all aspects of agreement between two land-use maps. This method assesses only the accuracy of simulated land-use changes. For a complete assessment of model results it should be complemented by other pattern oriented and scale-sensitive methods that account for the complex and path-dependent nature of land-use change.

2.5 Conclusions

The predictive accuracy of land-use models is often computed from the cell-by-cell comparison of actual land-use maps and simulated land-use maps, such as with the Kappa coefficient of agreement. However, the absolute value of this statistic depends heavily on the amount of land use that changes over the simulation period. Hence, applications that model a region with fewer land-use changes generally yield better results. However, this high agreement does not necessarily indicate an accurate simulation of land-use changes or a well calibrated model.

This paper introduces Kappa Simulation ($K_{Simulation}$), a statistic that is identical in form to the Kappa coefficient of agreement, but that uses a different underlying stochastic model. The Kappa coefficient of agreement corrects for the size of classes, whereas $K_{Simulation}$ corrects for the size of class transitions. The latter is a more meaningful reference level for the assessment of results of land-use models because it accounts for the amount of land-use change.

By correcting for class transitions, the absolute value of $K_{Simulation}$ can be interpreted in light of model calibration and validation, as it indicates how accurately a model can explain some land-use changes. This avoids a false impression of accuracy caused by high Kappa scores for models with only a small amount of change. Moreover, this method allows comparing results from models with different amounts of land-use change, as the agreement is corrected for the amount of land-use change. Although $K_{Simulation}$ gives an objective measure for the accuracy assessment of simulation models, it depends on the purpose of the modelling exercise to indicate what is accepted as sufficiently accurate.

$K_{Simulation}$ can be decomposed into $K_{Transition}$n and $K_{Transloc}$. This decomposition informs the modeller on the nature of the errors in the simulation result. $K_{Transition}$ values smaller than 1 indicate that the size of the class transitions for the entire application is not correct, while $K_{Transloc}$ values below 1 indicate that the allocation of land-use transitions is not entirely correct. This subdivision is especially helpful in models that explicitly simulate class transitions, such as cellular automata or Markov models, as it shows exactly what transition types are under- or over estimated. As such this method can improve the interpretation and communication of results of land-use models.

3. A fuzzy set approach to assess the predictive accuracy of land-use models

Van Vliet, J., Hagen-Zanker, A., Hurkens, J., and Van Delden, H. (Submitted) A fuzzy set approach to assess the predictive accuracy of land-use models. Ecological Modelling.

Abstract

The predictive accuracy of land-use models is frequently assessed by comparing two data sets: the simulated land-use map and the observed land-use map at the end of the simulation period. A common algorithm for this is the Kappa statistic, which expresses the agreement between two categorical maps, corrected for the agreement as can be expected by chance. This chance agreement is based on a stochastic model of random allocation given the distribution of class sizes. Two existing methods extend the Kappa statistic to make it more appropriate for the assessment of land-use models: Fuzzy Kappa uses fuzzy set theory to include degrees of agreement; this adds geographical nuance by distinguishing between small and large disagreement in position and in land-use classes. Kappa Simulation on the other hand addresses the stochastic model that underlies the expected agreement. When a model starts from an initial land-use map and subsequently makes changes to it, a stochastic model of random allocation given the distribution of class sizes has little relevance. The expected accuracy in Kappa Simulation model is therefore based on transition probabilities relative to the initial map. This paper presents Fuzzy Kappa Simulation, a method that combines the stochastic model of Kappa Simulation with the geographical nuance of Fuzzy Kappa. This new method is demonstrated on a case study example and results are compared with other variations of Kappa. The comparison confirms that Fuzzy Kappa Simulation is the only method to evaluate models in terms of land-use *change*, while being sensitive to geographical nuance.

3.1 Introduction

In the last decade, many land-use models have evolved into tools that can be used to study land-use change processes, perform scenario studies or do policy analyses for real world cases (see for example Van Delden et al., 2010; Hellmann and Verburg, 2011; Stanilov and Batty, 2011). Applying these models to any real world case requires calibration and validation, where calibration is defined as the adjustment of model parameters to improve the model's accuracy, and validation as the assessment of this accuracy using an independent dataset (Kok et al., 2001). Typically, this is done by reproducing known historic changes (for example in Wickramasurya et al., 2009; Wang et al., 2011). Applying these models to study land-use changes, perform scenario studies or do a policy analysis, requires an understanding of their performance, their strengths and weaknesses.

The accuracy of land-use models is often assessed using a pixel-by-pixel comparison of the simulated land-use map and the observed land-use map at the end of the simulation period. Several methods are available for this comparison, such as the Kappa statistic (Monserud and Leemans, 1992), the Tau coefficient (Ma and Redmond, 1994), the Average Mutual Information (Foody, 2006) and Receiver Operator Characteristics (Luoto et al., 2005). However, many available map-comparison methods have two important drawbacks for assessing land-use model results. First, they do not consider the amount of change during the simulation period; therefore scores for these comparison algorithms cannot be interpreted directly in terms of the predictive ability of a model. Second, these methods are crisp in their treatment of location and class boundaries; therefore near-hits are equivalent to complete misses while for a land-use modeller this is often not appropriate.

Some map comparison methods have been proposed that either address the amount of land-use change in a simulation, or a fuzzy interpretation of land-use maps, but not both. Hagen-Zanker and Lajoie (2008) use a reference model that simulates the same amount of net change relative to the original land-use map, and allocates these changes on the map randomly. Alternatively, Pontius et al. (2004b) use the original land-use map at the start of the simulation to distinguish between persistence and changes in the assessment of land-use simulations. Van Vliet et al. (2011) use the initial land-use map to implicitly account for the amount of change, by applying a stochastic model of random allocation of class transitions as a reference level. Fuzziness has been used in map comparison techniques with respect to location (Constanza et al., 1989 and

Pontius et al., 008), class boundaries (Townsend, 2000; Hagen, 2003; Fritz and See, 2005) or both (Hagen-Zanker 2009).

This paper presents Fuzzy Kappa Simulation (FKS), a map comparison algorithm that combines properties from Fuzzy Kappa (Hagen-Zanker, 2009) and Kappa Simulation (Van Vliet et al., 2011). FKS is a statistic similar in form to the original Kappa statistic (Cohen, 1960) but that instead applies a more appropriate stochastic model of random allocation of class transitions relative to the initial map and that uses a fuzzy interpretation of these land-use transitions. This new method has several advantages over other available map comparison methods: It allows to differentiate between change and persistence because it is based on land-use transitions rather than land-use classes, it allows to account for near-hits because it uses a fuzzy interpretation of land-use transitions, and the value of FKS directly indicates whether the model under assessment has any predictive capacity.

The rest of this chapter is organized as follows. Section 3.2 elaborates on two important aspects of land-use modelling: the difference between near-hits and complete misses and the difference between change and persistence, including its consequences for model assessment. Section 3.3 presents the algorithm for FKS. Section 3.4 discusses this map-comparison method by assessing the results of a case study model, and compares its results with other Kappa statistics. Section 3.5 then draws conclusions about FKS.

3.2 Assessing the predictive accuracy of land-use models

3.2.1 Near-hits and complete misses

The predictive accuracy of land-use models it typically assessed by comparing two maps at the pixel level: the simulated land-use map and the observed land-use map. For each pixel, such comparison indicates whether the land use is similar in both maps or not. Consequently, when a particular land-use change is simulated in the wrong location, it does not matter how far off this location is. However, from a modeller's point of view, simulating this change in the directly adjacent location can be considered as almost correct, while simulating this change at the other side of the study area would be a complete miss. Similarly, when a model simulates a land-use change in a particular location incorrectly, it is considered as completely different. However, for a modeller a change from cropland to dense residential areas simulated as a change from cropland into sparse residential land can be considered as almost correct. A crisp assessment of land-use classes would therefore be unnecessarily harsh for the comparison of two maps (Foody, 2008).

Because near-hits differ from complete misses, a fuzzy interpretation of locations as well as land-use classes is appropriate in the accuracy assessment of land-use models. Fuzziness of locations means that a cell has also a partial membership in the neighbouring locations. Therefore, the land use in a cell is also somewhat similar to the land uses found in its neighbourhood. A fuzzy interpretation of land-use classes means that a land-use class can have a partial membership in other land-use classes. For example, industrial land can be somewhat similar to residential land, but completely dissimilar to forested land.

Accounting for near-hits in land-use modelling is also justified by uncertainty that is inherent to land-use data. For example, the boundary between an urban area and its rural surroundings is often characterized by a transition zone, in which the fraction of build-up area decreases with an increasing distance from the city centre to the point where only a small part of the land is covered by buildings. Boundaries between many other land uses are similarly undetermined: instead of a crisp border there is a transition zone that shows a gradient from one land use to the other (Cheng and Molenaar, 1999; Fisher, 2000). This makes a fuzzy interpretation of the location of land uses appropriate: a location has also a partial membership in its neighbouring locations.

Land uses on a map are typically presented in mutually exclusive classes, while very few cells have only one land-use type: instead most include a combination of different land uses (Foody, 1996; Fisher, 2000; Foody, 2008). For example, many agricultural areas include some residential buildings, and many residential areas include some commercial activity. In addition to mixed land uses, the definition of a specific land-use class is also ambiguous: one can question how many trees it takes to classify a piece of grassland as a forest, or how many houses you need to classify a location as being residential. At the same time, an urban centre is quite different from a pristine wetland area. Because some land uses are more similar while other land uses are more different it is justified to use a fuzzy interpretation of land-use classes as well as land-use changes (Fisher et al., 2006).

The acquisition of land-use data adds additional sources for fuzziness in land-use maps. Many data sets use a minimum mapping unit, thereby omitting smaller patches from the map. Some time series, such as the Corine Land Cover data, even use a minimum area for recording land-use changes relative to another data set (Feranec et al., 2007). In addition to this, the acquisition process adds some more sources for uncertainty, such as the misregistration of pixels in remote sensing images, small misallocation introduced by the manual processing of aerial photographs, or subjectivity caused by the person interpreting aerial photographs or remote sensing images (Fritz and See, 2005, Foody, 1996). As a consequence land uses on a map and land-use changes in a series of maps can be considered an approximation, which justifies a fuzzy interpretation thereof.

3.2.2 Land-use change is different from land-use persistence

Most locations are characterized by land-use persistence rather than land-use change during a typical land-use change simulation (Pontius et al., 2004b). The predominance of persistence has a large implication for the assessment of land-use models, as the amount of change influences the similarity between the actual land-use map and the simulated land-use map at least as much as the accuracy of the simulated land-use changes. In other words: a model will perform well only because it reproduces a static landscape and not because changes are simulated accurately (Walker, 2003). Therefore, an end-state comparison is meaningless as a measure of the predictive accuracy without a meaningful reference level (Hagen-Zanker and Lajoie, 2008). Moreover, the relative merits of applications cannot be compared, since the amount of change can vary considerably between applications of land-use models. A meaningful reference level can be included when the initial map from the land-use simulation is included in the comparison, as this allows differentiating between change and persistence and, thereby, accounting for the amount of land use change (Van Vliet et al., 2011).

The differentiation between land-use changes and land-use persistence has an additional implication when fuzziness in location is incorporated in the model assessment. In reality, many land uses are strongly auto-correlated (Verburg et al., 2004a; Tang, 2008), and therefore locations that change into a land use are likely to be next to locations where that land use already exists. A typical example of this is urban growth, as the locations for new urban areas are often adjacent to already existing urban areas. An end-state assessment of model results cannot distinguish between persisting urban land and urban growth. Hence, when this urban growth is simulated in the wrong location but next to existing urban area, the comparison will interpret this as a near-hit, while it is be a complete miss for an urban growth model.

Figure 3.1 illustrates the difference between land-use change and land-use persistence for the assessment of model results: the simulated land-use changes are located next to the existing urban area, but they are not close the location of the actual land use changes, and vice versa. Hence when only the end-state is considered, these appear as near-hits, since the cells that are directly adjacent are also urban. However, when a model aims to simulate urban growth, it should compare the simulated land-use changes with the actual land-use changes. Hence, it is the distance between equivalent changes that determines whether it can be considered a near-hit or not. In Figure 3.1, the simulated land use changes and the observed land-use changes are located on opposite sides of the existing urban area. Therefore these can hardly be considered near-hits.

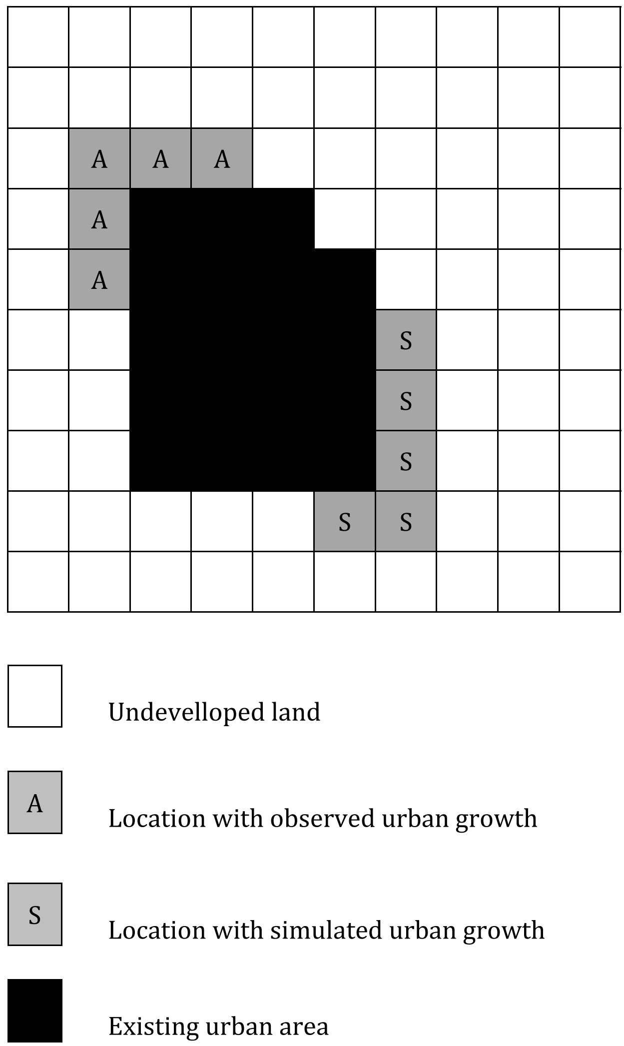


Figure 3.1: *A synthetic example to illustrate the difference between change and persistence. Both the simulated and the actual land-use changes are close to the existing urban land, while they are quite far from one another.*

3.3 Fuzzy Kappa Simulation

Fuzzy Kappa Simulation (FKS) combines properties from Fuzzy Kappa and Kappa Simulation as it expresses the agreement between observed land-use transitions and the simulated land-use transitions, corrected for the agreement that can be expected by chance and using a fuzzy interpretation of these land-use transitions. Three maps are required to compute FKS: the initial land-use map at the start of the simulation, the actual land-use map at the end of the simulation, and the simulated land-use map at the end of the simulation. Let O be the initial land-use map, let A be the actual land-use map, and let S be the simulated land-use map. All maps are raster maps with equal extent and resolution, and M is the set of all cells in these maps that are to be compared (possibly according to some mask). LU is the set of all land uses and T is the set of all possible land-use transitions ($T = LU \times LU$). In the following we will refer to a transition as $t \in T$ or with the form $\{o, s\}$ where $o \in LU$ and $s \in LU$.

The observed transition TA in a cell l is defined by the initial land use and the actual land use at the end of a simulation period. Likewise, the simulated transition TS in a cell is defined by the initial land use and the simulated land use at the end of the simulation period:

$$TA_l = \{O_l, A_l\} \quad \forall l \in M \quad (TA_l \in T) \qquad \text{Equation 3.1}$$

$$TS_l = \{O_l, S_l\} \quad \forall l \in M \quad (TS_l \in T) \qquad \text{Equation 3.2}$$

The similarity between two land-use transitions is expressed as a value between 0 (no similarity) and 1 (complete similarity). The similarity between pairs of land-use transitions is expressed in the similarity matrix X, hence $X(i, j)$ is the similarity between transitions i and j. Values in the similarity matrix are set by the modeller: they represent the modeller's interpretation of what land-use changes are similar and what not, or which land-use changes are important in the analysis and which are not. An example of the latter is given in Section 3.4.1. Because the fuzzy similarities are used to compute the observed agreement as well as the expected agreement fuzziness is not simply a way to inflate or deflate the overall comparison score, but instead it distinguishes almost similar pairs of land-use transitions from pairs of completely dissimilar land-use transitions.

3.3.1 Observed agreement

The observed agreement in a cell is computed based on both one-sided similarities for that cell. The one-sided similarity of a cell expresses the similarity of the simulated land-use transition in that cell, with any of the observed land-use transitions and vice versa. This similarity considers the fuzziness in land-use transitions as expressed in the similarity matrix X, and fuzziness in location as expressed in a distance decay function:

$$XA(l,t) = \max_{k\in M}\left(X(t, TA_k)\cdot f\big(d(l,k)\big)\right) \qquad \text{Equation 3.3}$$

$$XS(l,t) = \max_{k\in M}\left(X(t, TS_k)\cdot f\big(d(l,k)\big)\right) \qquad \text{Equation 3.4}$$

Where $f\big(d(l,k)\big)$ is a distance-decay function, $d(l,k)$ is the Euclidean distance between cell l and cell k, $XA(l,t)$ expresses how similar the transition map TA is to a transition t occurring in location l, and likewise for $XS(l,t)$. When the distance decay function f is cut off after a certain distance, the one-sided similarity is effectively limited to the neighbourhood of location l, but theoretically this is not required.

Then the observed agreement for cell l, PO_l, is the minimum of the two one-sided similarities:

$$PO_l = min\big(XA(l, TS_l), XS(l, TA_l)\big) \qquad \text{Equation 3.5}$$

The observed agreement PO for the entire map is than simply the average of the observed agreement over all cells on the maps

$$PO = \frac{\sum_{l\in M} PO_l}{|M|} \qquad \text{Equation 3.6}$$

where $|M|$ is the total number of cells in each map.

3.3.2 Expected agreement

The expected agreement between actual land-use transitions and simulated land-use transitions expresses the agreement that can be expected from a random allocation of class transitions relative to the initial land-use map. To derive the expected agreement, we follow the same structure as Hagen-Zanker (2009) for Fuzzy Kappa. The crucial distinction is that for Fuzzy Kappa the two compared maps are considered to be independent, whereas for FKS, we cannot consider the two compared transition maps to be independent, because they share the same original map. Instead, given an original land use o, we can consider the actual transition $\{o, a\}$ to be independent of the simulated transition $\{o, s\}$.

The logic behind the derivation of probabilities in this section can be understood if we consider a randomly selected cell in the map, such that the probability of each cell being selected is equal. The two one-sided similarities in this cell are unknown in general, because these require information on the location of the cell relative to all other cells in the map. However, given that we know (or can compute) the frequency with which each possible agreement in the range [0,1] occurs for the maps O, A and S, we can define a probability distribution for the similarity value that we observe in our randomly selected cell. Formally,

$$P\left(\mu A_{\{o,s\}} \geq x \middle| o, a\right) = \frac{\sum_{l \in M}(O_l = o \wedge A_l = a \wedge XA(l, \{o, s\}) \geq x)}{\sum_{l \in M}(O_l = o \wedge A_l = a)}$$

Equation 3.7

where $\mu A_{\{o,s\}}$ is a variable that is defined as the one-sided similarity $XA(m, \{o, s\})$ for a randomly picked location m that has original land use o and actual land use a. Analogously,

$$P\left(\mu S_{\{o,a\}} \geq x \middle| o, s\right) = \frac{\sum_{l \in M}(O_l = o \wedge S_l = s \wedge XS(l, \{o, a\}) \geq x)}{\sum_{l \in M}(O_l = o \wedge S_l = s)}$$

Equation 3.8

where $\mu S_{\{o,a\}}$ is a variable that is defined as the one-sided similarity $XS(m, \{o, a\})$ for a randomly picked location m that has original land use o and actual land use s. One-sided similarities $XA(m, \{o, s\})$ and $XS(m, \{o, a\})$ are computed similar to the one-sided similarity in the observed agreement:

$$XA(m, \{o, s\}) = \max_{k \in M}\left(X(\{o, s\}, TA_k) \cdot f\big(d(m, k)\big)\right)$$ Equation 3.9

$$XS(m, \{o, a\}) = \max_{k \in M}\left(X(\{o, a\}, TS_k) \cdot f\big(d(m, k)\big)\right)$$ Equation 3.10

The conditional probability that the two-sided similarity in a randomly selected location is at least x equals the conditional probability that both one-sided similarities are at least x. Formally,

$$P(\mu \geq x|o,a,s) = P\left(\mu A_{\{o,s\}} \geq x\middle|o,a\right) \cdot\ P\left(\mu S_{\{o,a\}} \geq x\middle|o,s\right)$$

Equation 3.11

The discrete nature of the categorical raster maps implies that the range of distance values that are passed to the distance-decay function is also discrete and hence that the range of values for local agreement is discrete (Hagen-Zanker, 2009). Therefore, there is a limited set of possible similarity values that will occur for any set of maps O, A and S. We define this set as $\{x_1, x_2, \dots x_n\}$, such that $x_i > x_j$ if $i < j$ (that is, sorted from high to low). Then we have:

$$P(\mu = x_k|o,a,s) = \begin{cases} P(\mu \geq x_k|o,a,s) & for\ k = 1 \\ P(\mu \geq x_k|o,a,s) - P(\mu \geq x_{k-1}|o,a,s) & for\ k > 1 \end{cases}$$

Equation 3.12

The conditional expected agreement in a cell is then taken as the sum of all possible two-sided similarity values x_k, weighted by the conditional probability of that value:

$$E(\mu|o,a,s) = \sum_{k=1}^{n} x_k \cdot P(\mu = x_k|o,a,s)$$ Equation 3.13

It is assumed that the conditional event of observing land use a in map A, given that the land use in map O for that cell is o, is independent of the conditional event of simulating land use s in the same cell, given o. Therefore, the probability $P(o,a,s)$ that we observe land uses o, a, and s after selecting a cell at random is:

$$P(o,a,s) = P(a|o) \cdot P(s|o) \cdot P(o)$$ Equation 3.14

$$\frac{\sum_{m\epsilon M}(A_m = a \wedge O_m = o)}{\sum_{m\epsilon M}(O_m = o)} \cdot \frac{\sum_{m\epsilon M}(S_m = s \wedge O_m = o)}{\sum_{m\epsilon M}(O_m = o)} \cdot \frac{\sum_{m\epsilon M}(O_m = o)}{|M|}$$

Equation 3.15

The expected agreement for the complete set of transitions, PE, is simply defined as the sum of the conditional expected agreements for all possible combinations of land-use classes in maps O, A, and S:

$$PE = \sum_{o\in LU} \sum_{a\in LU} \sum_{s\in LU} E(\mu|o,a,s) \cdot P(o,a,s)$$ Equation 3.16

3.3.3 Fuzzy Kappa Simulation

FKS is computed similar to the normal Kappa, as the observed agreement corrected for the expected agreement, but using the observed agreement and expected agreement as defined above:

$$FKS = \frac{PO-PE}{1-PE}$$ Equation 3.17

The definition of FKS is such that the normal Kappa statistic can be derived directly from this algorithm. When the fuzzy interpretation of land use maps is omitted from the FKS algorithm, it becomes exactly similar to Kappa Simulation, and likewise for Fuzzy Kappa and Kappa. This can be done by assigning a (fuzzy) similarity of 1 to those transitions that exactly match and a (fuzzy) similarity of 0 to all other combinations of transitions in the similarity matrix, and by changing the distance decay function so that it returns a no similarity disagreement for all distances > 0. Similarly, FKS becomes equivalent to Fuzzy Kappa when the original land-use map is replaced by any uniform map that has only one land use, and likewise for Kappa Simulation and Kappa. After all, the sizes of land transition classes then become exactly equal to the sizes of land-use classes in the simulated land use map and the observed land-use map.

The FKS algorithm and all other maps comparison algorithms applied in this paper are implemented in the Map Comparison Kit (Visser and the Nijs, 2006) version 3.3 and higher. The Map Comparison Kit is a software package that includes a multitude of map comparison algorithms that is a freely available from www.riks.nl/mck.

3.4 Results and discussion

3.4.1 A case study application

To demonstrate the use of Fuzzy Kappa Simulation (FKS) for model assessment and to compare it to other Kappa statistics it was used to assess the result of a case study application. Four different land-use models were applied to generate the case study maps: a null model, a random constraint match (RCM) model and two calibrations of a constrained cellular automata model. The null model does not simulate any change, and therefore the result is identical to the initial land-use map in 1990. The RCM model simulates the observed amount of net change, but these changes are allocated randomly on the map, while making only the minimal amount of adjustments relative to the initial map (Hagen-Zanker and Lajoie, 2008). The constrained cellular automata model is the Metronamica model (Van Delden and Hurkens, 2011); the two calibrations for this model differ in the value for the parameter that controls the strength of stochastic perturbations in the land-use model.

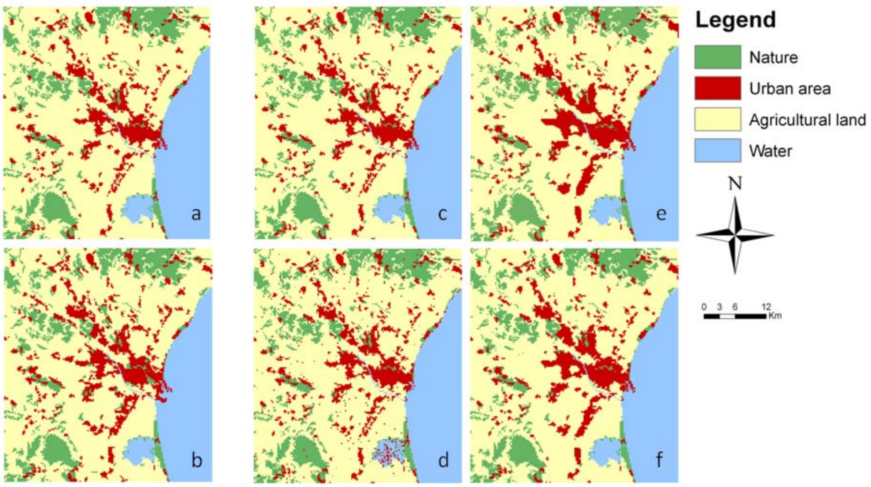


Figure 3.2: *The actual land-use maps for 1990 (a), and 2000 (b), and the simulated land use maps for 2000 using the null model (c), the RCM model (d), the cellular automata model with low stochasticity (e), and the cellular automata model with high stochasticity (f).*

Figure 3.2 shows the initial land-use map for 1990, the actual land-use map in 2000 and four different simulated land-use maps in 2000 for a small area around the city of Valencia, Spain. All maps have four land-use classes, which are nature, agriculture, urban area and water, and the specific aim of the modelling exercise is to simulate urban growth between 1990 and 2000.

The predictive accuracy of the results of the case study application are computed using the normal Kappa statistic as well as Fuzzy Kappa, Kappa Simulation and FKS. Both Fuzzy Kappa and FKS use exponential distance decay, with a halving distance of 2 and a cut-off distance of 10, and in both cases all four land uses are completely dissimilar from each other. As a consequence, this example only considers fyzziness with respect to location. Moreover, the similarity matrix in FKS is set so that land-use changes are dissimilar from land-use persistence, a distinction that can only be made in the proposed method. This similarity matrix and all other similarity matrices used in this case study example are shown in the appendix.

Table 3.1: *Predictive accuracy of the case study application, assessed using Kappa, Fuzzy Kappa, Kappa Simulation and Fuzzy Kappa Simulation.*

	Null model	**RCM model**	**Metronamica with low stochasticity**	**Metronamica with high stochasticity**
All land uses				
Kappa	0.970	0.952	0.971	0.971
Kappa Simulation	0.000	0.001	0.304	0.348
Fuzzy Kappa	0.979	0.958	0.979	0.979
Fuzzy Kappa Simulation	0.000	-0.005	0.389	0.446
Urban land use only				
Kappa	0.861	0.768	0.843	0.844
Kappa Simulation	0.000	-0.001	0.150	0.163
Fuzzy Kappa	0.915	0.810	0.907	0.910
Fuzzy Kappa Simulation	0.000	0.002	0.292	0.316

Table 3.1 presents the results of these map comparisons. All four model results yield very high scores for Kappa as well as Fuzzy Kappa, indicating that the

generated land-use maps are very similar to the observed land-use maps at the end of the simulation period. However, since only a small part of the landscape changed between 1990 and 2000, this is mostly due to correctly simulated persistence while it is not possible to interpret these values in terms of the predictive accuracy of the applied land-use model. Kappa Simulation and FKS account for this persistence as their expected agreement is based on a stochastic model of random allocation of class transitions relative to the initial map. Therefore Kappa Simulation and FKS scores are lower, but more importantly, they show a clearer distinction between the different simulation results. Specifically, they show that the null model and the RCM model do not explain any land-use changes, while both simulation models do, as their scores are well above zero.

The difference between Fuzzy Kappa and FKS is best illustrated by the comparison of the results of the Metronamica land use model. Both models yield a similar score for Fuzzy Kappa, suggesting that they are equally accurate. However, FKS scores indicate that the calibration that uses a higher stochasticity has a higher predictive accuracy. Apparently, this simulation yielded more near-hits than the calibration with a lower stochasticity. Hence, this method indicates which calibration is preferable, while the other methods were not able to identify clear differences between both calibrations.

Because the aim of the case study application was to simulate urban growth, FKS was also used to assess these changes specifically. The result of this comparison is also presented in Table 3.1. The results for the assessment of urban land use only are qualitively similar to the assessment of al land use changes, and are shown here only for purposes of illustration. The difference with the comparison of all land use classes is only in the similarity matrix. When urban land use is analyzed separately conversion between other land uses do not matter anymore and consequently they are considered similar. As indicated above, this does not inflate the overall similarity, as it affects both the observed agreement and the expected agreement. Instead, it specifically assesses those changes that are of interest to the modeller: changes from non-urban to urban land. More generally, by adjusting the similarity matrix, this method allows to focus on one or several particular types of land-use change instead all land use classes in a simulation. These particular changes can contain the appearance of land uses, such as urban growth, or the disappearance of land uses, such as tropical deforestation.

Figure 3.3 provides an illustration of the difference between the respective Kappa variations graphically. It shows the observed agreement of the map comparisons for the Metronamica model with high stochasticity. The top row (Figure 3.3a-c) shows the comparison of all land-use classes, while the bottom row (Figure 3.3d-f) provides the results for the urban land use only. From left to right results are shown for Kappa / Kappa Simulation, Fuzzy Kappa and FKS. Kappa and Kappa Simulation only differ in their expected agreement, hence the observed agreement is exactly similar. Clearly the results for Kappa differ from the others in their crisp interpretations: all cells show either black or white pixels indicating hits or misses. The difference between Fuzzy Kappa and FKS is more nuanced. FKS yields darker pixels for some near-hits, as illustrated by the patch in the dashed circle in Figure 3.3. These pixels indicate land-use changes that are simulated close to persisting land uses, but not close to the observed land-use changes. The results for all land uses compared to urban land use only differ in the total amount of errors: non-urban land use transitions are not ignored in the latter assessment and therefore the do not appear on the comparison maps. The dotted circle in Figure 3.3 indicates a location where these differences are visible.

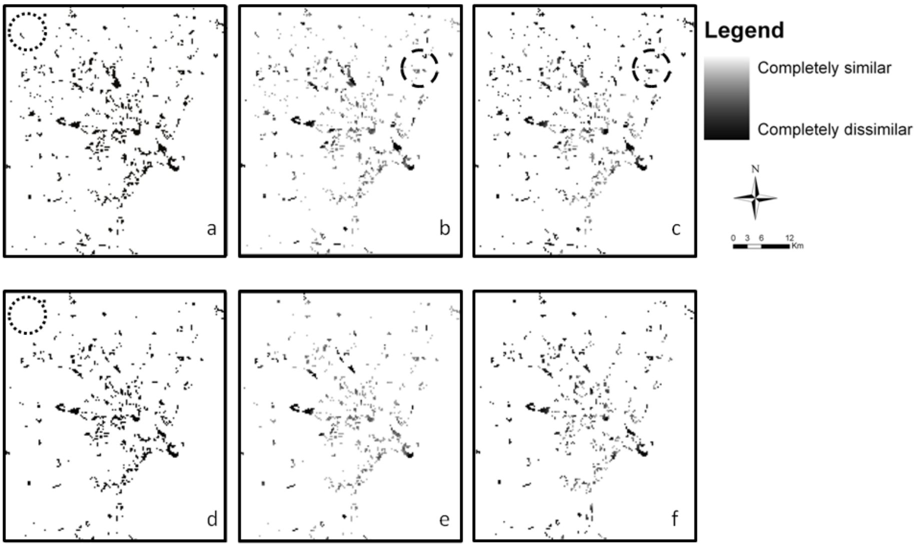


Figure 3.3: *Graphical results for the different map comparison methods. Results are shown for the Metronamica model with high stochasticity: a) Kappa – all land uses, b) Fuzzy Kappa all land uses, c) FKS – all land uses, d) Kappa – urban land use only, e) Fuzzy Kappa urban land use only, and f) FKS – urban land use only. See the text for further explanation.*

3.4.2 Fuzzy Kappa Simulation compared to other Kappa metrics

A number of variations on Kappa have been developed since Monserud and Leemans (1992) first applied this method for the accuracy assessment of spatial simulation models. Despite the existence of alternative map-comparison measures (Turner et al., 1989; Couto, 2003), Kappa and its varieties have since become the predominant measure to compare categorical maps. A full discussion of all methods for accuracy assessment goes beyond the scope of this paper; instead Kappa, Fuzzy Kappa, Kappa Simulation and Fuzzy Kappa simulation, all used in the case study application, are discussed in the context of assessing the results of land-use models. All four have the same basic structure – the observed agreement corrected by the expected agreement – but the definition of the expected agreement differs. Similarly all four have the same range of possible values – between -1 and 1, with 0 indicating an agreement as can be expected by chance and 1 meaning a perfect fit – but the interpretation of the absolute values differs as a function of the expected agreement. Table 3.2 provides an overview of characteristics of the four Kappa measures.

Kappa and its variations differ in their consideration of the amount of land-use change and the consideration of near-hits. Both Kappa and Fuzzy Kappa do not consider the amount of change in the simulation period. For that reason the absolute values of these measures have no intrinsic meaning in the context of land-use modelling. Land-use models applied to areas and/or periods with little change will generally yield higher scores than when applied to areas and/or periods with many land-use changes, regardless of the accuracy of these changes. Therefore a benchmark is required to interpret Kappa values. In this respect the assessment of land-use models is fundamentally different from the classification of remote sensing images, since the latter does not start from an initial land-use map. Kappa Simulation and Fuzzy Kappa Simulation on the other hand compute the expected accuracy as a function of the sizes of class transitions. Hence, the amount of change is included implicitly and the absolute value of both measures can be interpreted for the accuracy assessment of land-use models. A value > 0 indicates that a model has some predictive accuracy. Kappa and Kappa Simulation only consider crisp land-use classes and crisp locations. This means they cannot account for near-hits in terms of classes that are somewhat similar but not entirely, or in terms of nearby locations with the same land-use type or transition. As near hits can be valuable result for a land-use modeller, it is justified to also account for them in model assessment methods. Due to fuzziness in location, only Fuzzy Kappa and FKS are explicitly spatial, while Kappa and Kappa Simulation can also be applied to other, non-spatial, categorical datasets. The combination of accounting for the amount of change and near-hits makes FKS arguably the most appropriate method out of the four Kappa variations discussed for the assessment of the predictive accuracy of land-use models. However, this advantage comes at the cost of subjectivity, introduced by the modeller in the values in the similarity matrix.

However, the use of Kappa to assess the predictive accuracy of land-use models is not undisputed: Pontius and Millones (2011) criticize Kappa statistics for several reasons; however, they do not provide an appropriate alternative, as their suggested approach does not assess the predictive accuracy of land-use models, but instead provides insights in the types of error made. We see two possible directions to improve accuracy assessment methods: this paper presents an improved model for the expected agreement, which solves part of their critique; an alternative approach is the application of reference models such as presented in Hagen-Zanker and Lajoie (2008). Some modellers estimate model parameters empirically; however, although this removes the requirement to calibrate land-use models, it does not eliminate the necessity to gauge model performance using these estimated parameters.

Table 3.2: *Characteristics of Kappa and its variations.*

Algorithm	Location and class boundaries	Expected agreement based on
Kappa	Crisp	A stochastic model of random allocation of land use classes
Fuzzy Kappa	Fuzzy	A stochastic model of random allocation of land use classes
Kappa Simulation	Crisp	A stochastic model of random allocation of class transitions relative to the initial map
Fuzzy Kappa Simulation	Fuzzy	A stochastic model of random allocation of class transitions relative to the initial map

3.5 Conclusion

This paper presents Fuzzy Kappa Simulation (FKS), a map-comparison algorithm that expresses the agreement between simulated land-use changes and observed land-use changes, corrected for the expected agreement and using a fuzzy interpretation of land-use transitions. This algorithm combines properties of Fuzzy Kappa (Hagen-Zanker, 2009) and Kappa Simulation (Van Vliet et al., 2011) in a single map comparison method. FKS has several important advantages over other map comparison methods available to assess the predictive accuracy of land use models: It allows to differentiate between land-use changes and land-use persistence because it is based on land-use transitions rather than land-use classes; it differentiates between near-hits and complete misses because it uses a fuzzy interpretation of land-use transitions; and no benchmark is needed in the assessment of land-use models because there is an appropriate reference model implicit to this method. Moreover, by adjusting the similarity matrix FKS can be tailored to assess specific types of land-use changes, such as the simulation of urban growth or deforestation, by changing the similarity matrix. The assessment of a specific type of land-use change can be very useful in relation to the aim of a particular modelling study, such as studying urban growth or deforestation.

Due to its properties, FKS is very suitable for the assessment of the results of land-use models. This was shown by applying this new method to assess land-use maps generated by different land-use models. Results show that FKS, like Kappa Simulation, can differentiate between similarity due to persistence and similarity due to correct changes, which is of crucial importance in land-use modelling. Moreover, a comparison between scores for FKS and Fuzzy Kappa show that this method is indeed capable to distinguish between near-hits and complete misses, which is very relevant for interpreting and communicating the results of land-use models. It should be noted that a comprehensive assessment of the results of land-use models includes an assessment of several map properties (Hagen-Zanker and Martens, 2008). Hence FKS is very suitable to assess the predictive accuracy but should be complemented with methods that assess the process accuracy (Brown et al., 2005).

Appendix 3.A: Similarity matrices used for the case study application.

This appendix presents the similarity matrices used to compute Fuzzy Kappa and Fuzzy Kappa Simulation Scores in section 3.4. Land uses are abbreviated as follows: A = Agriculture; F = Forest; U = Urban; and W =Water:

Table 3.A.1: *Similarity matrix to compute Fuzzy Kappa for all land uses.*

	Agriculture	Forest	Urban	Water
Agriculture	1	0	0	0
Forest	0	1	0	0
Urban	0	0	1	0
Water	0	0	0	1

Table 3.A.2: *Similarity matrix to compute Fuzzy Kappa Simulation for all land uses.*

	A-A	A-F	A-U	A-W	F-A	F-F	F-U	F-W	U-A	U-F	U-U	U-W	W-A	W-F	W-U	W-W
A-A	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A-F	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
A-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
A-W	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
F-A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
F-F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
F-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
F-W	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
U-A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
U-F	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
U-U	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
U-W	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
W-A	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
W-F	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
W-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
W-W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 3.A.3: *Similarity matrix to compute Fuzzy Kappa for urban land use only.*

	Agriculture	Forest	Urban	Water
Agriculture	1	1	0	1
Forest	1	1	0	1
Urban	0	0	1	0
Water	1	1	0	1

Table 3.A.4: *Similarity matrix to compute Fuzzy Kappa Simulation for urban land use only.*

	A-A	A-F	A-U	A-W	F-A	F-F	F-U	F-W	U-A	U-F	U-U	U-W	W-A	W-F	W-U	W-W
A-A	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
A-F	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
A-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
A-W	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
F-A	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
F-F	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
F-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
F-W	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
U-A	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0
U-F	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0
U-U	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
U-W	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0
W-A	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
W-F	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1
W-U	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
W-W	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1

4. Modelling urban growth using a variable grid cellular automaton

Van Vliet, J., White, R., and Dragicevic, S. (2009). Modelling urban growth using a variable grid cellular automaton. Computers Environment and Urban Systems 33(1): 35-43. (Slightly adapted in order to improve consistency with other chapters).

Abstract

Constrained cellular automata (CA) are frequently used for modelling land-use change and urban growth. In these models land use dynamics are generated by a set of cell state transition rules that incorporate a neighbourhood effect. Generally, neighbourhoods are relatively small and therefore only a limited amount of spatial information is included. In this study a variable grid CA is implemented to allow incorporation of more spatial information in a computationally efficient way. This approach aggregates land uses at greater distances, in accordance with a hierarchical concept of space. More remote areas are aggregated into consecutively larger areas. Therefore the variable grid CA is capable of simulating regional as well as local dynamics at the same time. The variable grid CA is used here to model urban growth in the Greater Vancouver Regional District (GVRD) between 1996 and 2001. Calibration results are tested for predictive accuracy by means of the Kappa statistic and for its process accuracy by means of cluster size analysis and radial analysis. Kappa results show that the model performs considerably better than a neutral refernce model. Cluster and radial analysis indicate that the model is capable of producing realistic urban growth patterns.

4.1 Introduction

Tobler's first law of geography states that "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Translated to land use this implies that the surroundings of a location are related to the land use in that location, but close surroundings have a stronger influence than more remote surroundings. The notion that land uses are spatially related and that nearby land uses have a stronger relation than land use at a greater distance was confirmed by empirical analysis of neighbourhood characteristics (Verburg et al., 2004b; 2004c). This influence of neighboring land uses is strongly embedded in cellular automata (CA) based land-use models by their neighbourhood effect.

CA models are used in several ways to model land-use changes (White et al., 1997; Clarke et al., 1996, Wu, 1998a), where they are found to be particularly applicable to simulate urban dynamics (White and Engelen 1993; Barredo et al., 2004). The latter is predominantly so for the ability of CA to create complex patterns (Wolfram, 1984) that are not unlike urban patterns (Batty and Xie, 1994; Batty, 2005). More recently, CA land-use models have been applied as tools to support land-use planning and policy analysis (Geertman and Stillwell, 2003) as well as to explore scenarios for future development (Engelen et al., 2003; Barredo et al., 2003a; De Nijs et al., 2004).

A CA essentially comprises the following elements: (1) a cell space or lattice, (2) a finite set of cell states, (3) a definition of a cell's neighbourhood, (4) a set of transition rules to compute a cell's state change and (5) time steps in which all cell states are simultaneously updated (White and Engelen, 2000). To make CA applicable for geographical modelling, the strictly defined CA rules are frequently loosened. These models are therefore referred to as relaxed cellular automata models (Couclelis, 1997). In constrained CA models, the total amount of area per land use is not a function of the transition rules, but determined exogenously instead, while the allocation of these land uses is computed by the CA (White et al., 1997). For example, in an urban growth model the total area for residential land use can be derived from historic data or extrapolations thereof. This area demand is then imposed on the CA model that allocates a corresponding number of cells on the map, based on the transition rules.

4.1.1 On a cell's neighbourhood

A cell's neighbourhood is the region that serves as an input to calculate the neighbourhood effect in the transition rules. This effect is a function of a cell's own state and the state of the cells within its neighbourhood. In land use terms, this represents attraction or repulsion of neighboring land uses. Hence the size of the neighbourhood determines the amount of land-use information that is considered in the neighbourhood effect. Originally, in CA only directly and diagonally adjacent cells were included. In human induced land-use change, however, information at greater distances also influences land-use changes, although the effect typically decreases with increasing distance. Hence larger neighbourhood configurations are used to model land use change and urban growth (White and Engelen, 1993). In current applications this size ranges up to an 8-cell radius, enclosing 196 cells (Engelen et al., 2003; Barredo et al., 2004). Since larger neighbourhoods include more land-use information, they allow for better models. The number of cells in a neighbourhood is directly related to the radius of the neighbourhood. Therefore, increasing this radius would include more land-use information. However, the required computation time would increase dramatically, as the number of cell-to-cell relations grows with the square of that radius. At the same time, this approach would use spatial information at larger distances at a higher level of detail then required.

Still, intuitively, more distant areas also influence land-use change (Andersson et al., 2002a). This notion that information can travel over greater distances and thus have influence further away than just adjacent areas is well established in Hägerstrand's innovation diffusion (1967). To incorporate effects operating over larger distances, it has been necessary to combine two or more models that operate on different spatial levels. In these integrated models, a gravity based regional model calculates regional demands for land uses and a constrained CA model then allocates these demands on the map (White and Engelen, 2000). To overcome this problem, a more complete hierarchical conceptualization of space was introduced in Andersson et al. (2002b). The assumption is that humans intuitively use a similar indexation to interpret and divide space: A city has several parts, each part consists of several blocks and every block again has a number of houses. The closer a feature is, the more in detail we think of it. Close surroundings, like neighboring houses, are of prime importance in spatial decisions. The more remote environment is considered with respect to its place in a spatial hierarchy: the next block is less important than immediate adjacent houses, but more important than the other side of town (Andersson et al., 2002a). In analogy to this hierarchical notion of space, cells at a greater distance can be aggregated to larger areas, while detailed information is kept for areas close by. This aggregation to area averages of land uses considerably reduces the number of spatial relations and thus the required computation time (White, 2005). Consequently, spatial information over much larger distances can be incorporated in the neighbourhood effect and interregional effects need no longer be calculated in a separate model.

The variable grid CA is an implementation of this concept in a CA environment that allows incorporating all available land-use information when calculating an individual cell's propensity to change. This is done by enlarging the neighbourhood to include cells at all distances by using a hierarchical representation of space in the neighbourhood definition. Specifically, this method uses a variable grid to aggregate more remote areas to mean field approximations (White, 2005). More distant cells are aggregated into increasingly bigger fields. This limits the number of spatial relations to be computed while nevertheless incorporating the maximum amount of land-use information. Thus, the model incorporates long-distant relations as well as local effects. In this study the variable grid CA is applied to simulate urban growth in the Metro Vancouver area (former Greater Vancouver Regional District - GVRD). Both its applicability to simulate actual urban growth and its ability to simulate regional dynamics were tested with this application.

Moreover, the variable grid as presented in White (2005) introduces levels of activity for land uses. In the present application these are not incorporated and therefore activities are not considered in this text.

4.2 The variable grid cellular automata model

For this study, the variable grid neighbourhood was implemented in a constrained CA model. Hence, the demand per land-use class is defined exogenously: for each year the demand for constrained land use classes is defined in terms of a number of cells for the whole area (White et al., 1997). The allocation of these cells is determined by the potential of each cell for all land-use classes as computed by the CA transition rules and using the variable grid neighbourhood configuration. Land uses are assigned to cells with the highest potential, until the demand for this land use is met. Potentials for each cell and for each constrained land-use class are calculated as follows (White and Engelen, 2000):

$$P_{i,l} = \nu \cdot A_{i,l} \cdot S_{i,l} \cdot Z_{i,l} \cdot N_{i,l} \qquad \text{Equation 4.1}$$

Where $P_{i,l}$ is the potential for cell *i* and land use *l*; ν is a stochastic perturbation term equal to *1 + (-log(random))*$^{\alpha}$,where α is a scaleable parameter and *random* is a randomly drawn number from a uniform distribution between 0 and 1; $A_{i,l}$ is the accessibility of cell *i* for land use *l* to transport networks; $S_{i,l}$ is the suitability of cell *i* for land use *l*; $Z_{i,l}$ is the zoning status of cell *i* for land use *l*; and $N_{i,l}$ is the neighbourhood effect for cell *i* for land use *l* as computed using the variable grid neighbourhood, as explained below. Calculation of variables other than the neighbourhood effect is discussed more fully in White and Engelen (2000).

The variable grid CA was implemented using the Geonamica spatial modelling framework. This modelling framework (without the variable grid) has been applied successfully in land-use models, for example the Environment Explorer (Engelen et al., 2003) and the MOLAND project (Barredo et al., 2003b), and in integrated spatial models, for example MedAction (van Delden et al., 2007).

4.2.1 Definition of the cell neighbourhood effect

The basic lattice with the highest resolution is referred to as the level 0 grid. At this level, every cell has only one state that represents its actual land use, formalized as:

$$C_k^0(x) = \epsilon\{0,1\} \qquad \text{Equation 4.2}$$

where $C_k^0(x)$ is 1 if land use k is present at location *x* and 0 otherwise. Now each successive level (L) then contains 3^{2L} level 0 cells. Thus level 1 cells are an aggregation of 3^2 = 9 level 0 cells and, level 2 cells of $3^{2 \cdot 2}$ = 81 lovel 0 cells. As a result, higher level cells are represented with cell counts of level 0 land uses instead of having one single state, and $C_k^l(x)$ is the cell count of land-use class k in a square of 3^{2l} cells centered at *x*. Each level 0 cell has eight adjacent cells: 4 rook adjacent and 4 bishop adjacent. Around this level 0 neighbourhood there are eight level 1 aggregated cells, which are again surrounded by eight level 2 cells, etc. More generally every level L contains four rook adjacent cells:

$$D_i^{rook}(L) = \{(i, i+3^l), (i+3^l, i), (i, i-3^l), (i-3^l, i)\} \qquad \text{Equation 4.3}$$

and four bishop adjacent cells:

$$D_i^{bishop}(L) = \{(i + 3^l, i + 3^l), (i + 3^l, i - 3^l), (i - 3^l, i + 3^l), (i - 3^l, i - 3^l)\}.$$

Equation 4.4

This neighbourhood template, as shown in Figure 4.1, is relative to each individual cell and therefore it moves cell by cell over the entire grid. Each aggregated cell holds cell counts for all land uses $l, k \in \{1,2, \ldots, m\} = K$, where K is the set of all possible land-uses states.

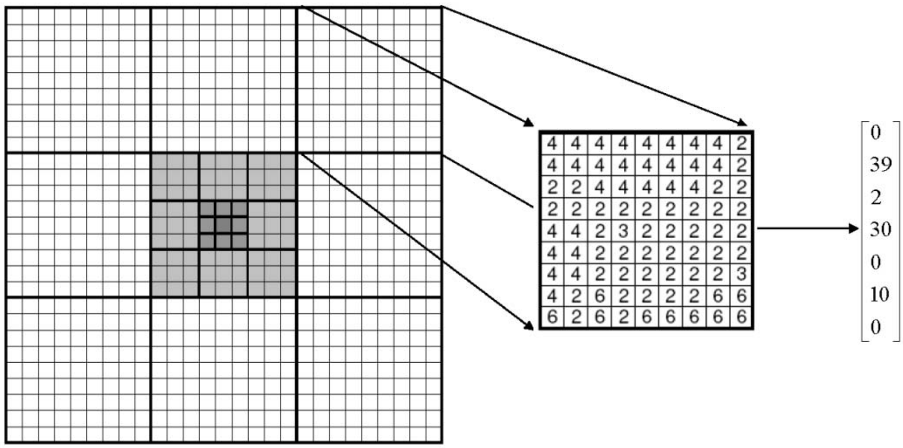


Figure 4.1: *Three aggregation levels relative to the central cell in the neighbourhood. Numbers characterize different land-use types. The vector represents the cell counts of level 0 cells per land-use type as assigned to the central point of the aggregated level 2 cells.*

Influence of land use is represented by a weight which represents the attraction or repulsion from one land use to another as a function of the distance. Since rook adjacent cells are closer than bishop adjacent cells, this requires two discrete weight values for each consecutive aggregation level. Since the variable grid incorporates the whole area in the neighbourhood, the neighbourhood effect is the cumulative effect of all weighted cell to cell land-use relations in all consecutive levels of aggregation:

$$N_{i,l} = \sum_{L} \left[\sum_{x \in D_i^{rook}(L)} \sum_{k \epsilon K} w_{l,k}(3^l) \cdot C_k^L(x) + \sum_{x \in D_i^{bishop}(L)} \sum_{k \epsilon K} w_{l,k}(\sqrt{2} \cdot 3^l) \cdot C_k^L(x) + \right]$$

Equation 4.5

Where N_{il} is the neighbourhood effect for cell i for land use l, $w_{l,k}$ *(d)* is the weight parameter representing the attraction or repulsion from land use k on land use l at distance d and $C_k^L(x)$ is the number of level 0 cells with land use k aggregated in the cell centered at x. Distance d is measured from the centre of cell i to the centre of each aggregated cell, x.

4.3 A case study on the Greater Vancouver Regional District

The Greater Vancouver Regional District is a highly urbanized and rapidly growing area located in the lower mainland of British Columbia, Canada. In the last century, population increased from just over 230 000 in 1921 to almost 2 000 000 in 2001 (Figure 4.2). Projections for the near future show no change in this trend, and population is expected to grow almost linearly to around 2 900 000 in 2031 (BC STATS, 2006). At the same time space for urban expansion is scarce. Greater Vancouver is surrounded by the sea to the west, the United States to the south and mountains to the north. The land that is suitable for urban land use is mainly protected and used for agriculture and natural areas. Hence, to prevent urban sprawl and protect both agricultural and natural areas, the government aims at concentration of population and restricted growth. Formally this is implemented in the Livable Region Strategic Plan (GVRD, 1999).

The Livable Region Strategic Plan defines four aims for a sustainable growth strategy. First, protect the green zone: the green zone protects Greater Vancouver's natural assets, including major parks, watersheds, ecologically important areas and resource lands such as farmland. It also establishes a long-term growth boundary. Second, build complete communities: the plan supports the public's desire for communities with a wider range of opportunities for day-to-day life. Focused on regional and municipal town centers, more complete communities would result in more jobs closer to where people live and accessible by transit, shops and services near home, and a wider choice of housing types. Third, achieve a compact metropolitan region: the plan avoids widely dispersed development and accommodates a significant proportion of population growth within the "growth concentration area" in the central part of the region. Fourth, increase transportation choice: the plan supports the increased use of transit, walking and cycling by minimizing the need to travel and by managing transportation supply and demand.

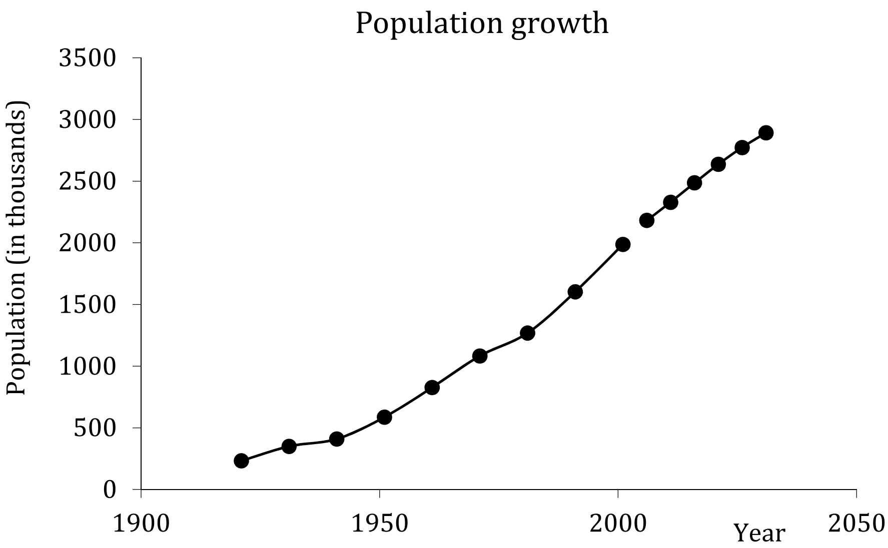


Figure 4.2: *Historic and projected population numbers in the Greater Vancouver Regional District (GVRD, 1999).*

4.3.2 Datasets

The GVRD area covers 2820 km^2. Raster data layers are represented on a grid of 760 by 635 cells and have a 100 meter spatial resolution. Land-use data was made available from the Greater Vancouver Regional District for the years 1996 and 2001. Hence, land-use change was simulated for this period, using time steps of one year. Overall land-use change, in terms of number of cells, was derived from the 1996 and 2001 land-use data and used as an exogenous constraint to the model. Land use-maps were classified in 14 classes in 1996 and 15 classes in 2001. The 14 classes were identical in both maps, which made comparison feasible. The one new class in 2001 is combined residential and commercial land use and is reclassified as commercial land. For use in the model, land use was reclassified to seven new classes: (1) agriculture, (2) forest and protected nature, (3) open and undeveloped, (4) commercial and industry, (5) residential, (6) extractive industry and (7) water. Of these only (4) and (5) are truly active classes: their total number is constrained exogenously, but their allocation is completely dependent on the potential as computed with the transition rules. Effectively, at each time step all cells in these classes are allocated again. However, the inertia effect results in only a few actual changes, mainly the increase in these classes. Classes (1), (2) and (3) represent passive land uses, they can only change as a result of change in active classes. Finally, classes (6) and (7) are fixed; they cannot change. However the presence of fixed land uses can influence the allocation of active land uses.

Additional data is used to derive accessibility information, a suitability map and a zoning map. Accessibility is computed in the model as a function of the Euclidean distance to the nearest cell that contains a transport network. Therefore, three transport networks are selected: skytrain, limited access highways and major roads. Information on transport networks was obtained from Greater Vancouver Transportation Authority (TransLink) in BC, Canada. To represent the physical suitability for urban land uses, a slope map is derived from a digital elevation model and aggregated to the appropriate cell size. The Digital elevation model was provided by the Greater Vancouver Regional District. Finally a zoning map is created to represent the restrictions on the development of residential and commercial and industrial land use in certain areas. This map reflects the GVRD Green Zone policy to preserve natural and agricultural areas (GVRD, 1999).

4.4 Model calibration and results

Calibration results are assessed from properties of the simulated land-use maps. Three different aspects of the output maps were measured: the goodness of fit on a pixel basis, the capability to produce realistic urban patterns, and the ability to model regional interactions. In the assessment, the simulation results were compared to results from a reference model. A random constrained match (RCM) model was used to create these reference maps (Hagen-Zanker & Lajoie, 2008). This model computes the amount of actual land-use change between two land-use maps and allocates this change randomly but with minimal change on the initial map, in this case the 1996 land-use map. As a result, the random map will have the same land-use frequency distribution as the actual 2001 land-use map.

Generation of both the model results and the reference results involves a random term. Therefore five model runs and five reference results were obtained to assess the quality. Maps a, b, c, d in Figure 4.3 represent the 1996 land-use map, the 2001 land-use map, a simulation result and a RCM result.

4.4.1 Goodness of fit

Accuracy of simulation results on a pixel by pixel basis was assessed using the Kappa statistic. This statistic measures the goodness of fit between two nominal datasets, corrected for accuracy by chance (Bishop et al., 1975). Since land-use maps are categorical maps, Kappa can be used to assess the goodness of fit between the simulation result and the real land-use map at the end of the simulation period (Foody, 2002; Pontius, 2002). Because the emphasis of this study is on simulating growth in urban land-use classes, i.e commercial and industrial land and residential land, Kappa statistics were also computed for these land uses separately.

Kappa values range from 1 to -1, where positive values indicate a better agreement than expected by chance, and negative values a worse agreement. However, the absolute value of Kappa is not an appropriate measure for model results since it is highly dependent on the number of cells that change. A simulation with very few changing pixels will result in high Kappa values, even if all newly allocated pixels are placed incorrectly. Therefore this statistic can only be used to compare different results from the same case study. Hence, Kappa values are considered here relative to the results of the random model.

A drawback of using Kappa statistics for model results is that slight displacements are classified as incorrect, whereas from a modeler's perspective this can be considered almost correct. For example, new residential land use that is allocated just one cell away from the actual location of this new residential area can be regarded as a good result. Therefore a Fuzzy Kappa was used (Hagen, 2003, Hagen-Zanker et al., 2005). This statistic uses a linear distance-decay function to account for slightly displaced pixels. Fuzzy Kappa was computed using a slope of 0.2 and a radius of 5 cells, i.e. a residential cell that is dislocated exactly 2 pixels, would count as 0.6 correct.

Model results are presented in Table 4.1. Although a random perturbation term is necessary to obtain realistic results, and although simulation results differ significantly from each other, Table 4.1 indicates that simulation results are similar in terms of goodness of fit. Both Kappa and Fuzzy Kappa scores indicate that the model performs considerably better than the RCM model. The relatively low Kappa scores for commercial and industry are caused by the appearance of a few large patches of this particular land use between 1996 and 2001. These are the results of one single planning decision and as such they cannot be simulated using a bottom-up approach like a CA land-use modelling.

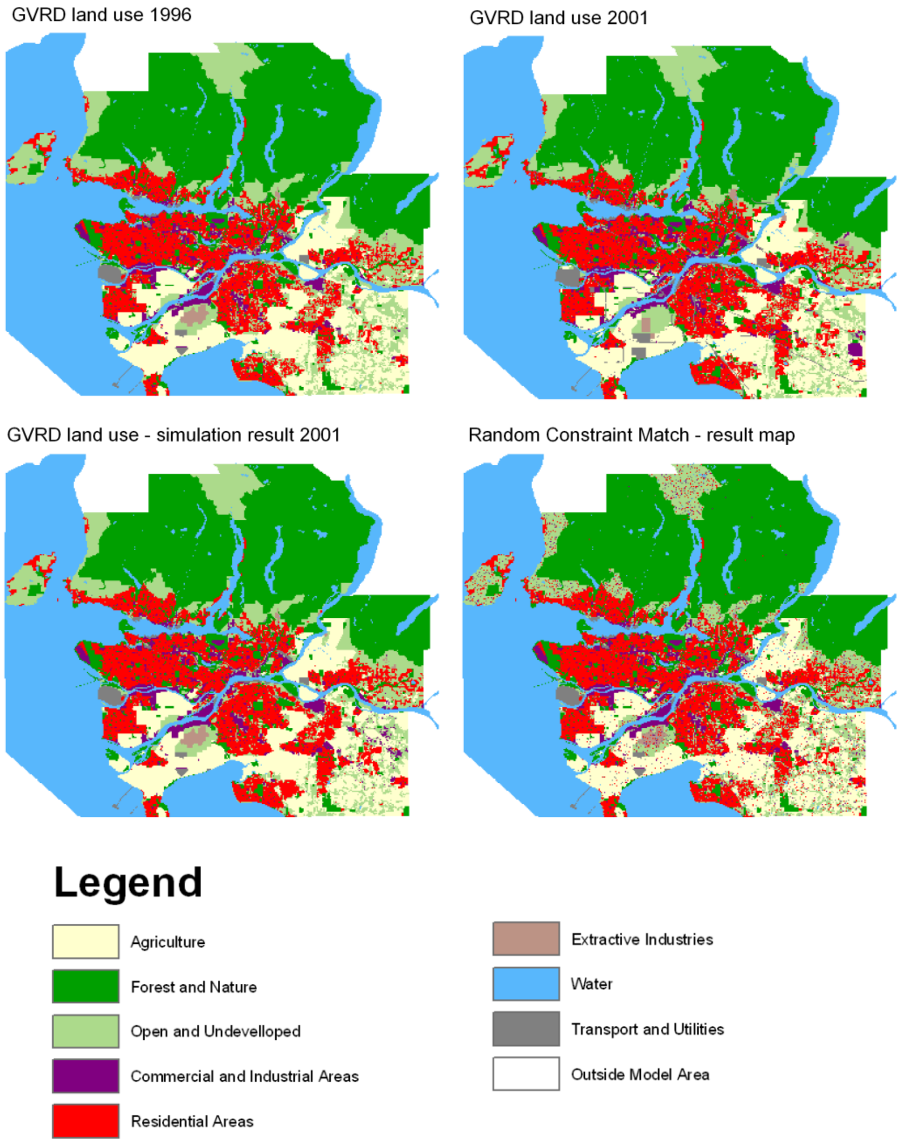


Figure 4.3: *Land-use maps representing (a) the 1996 actual land use, (b) the 2001 actual land use, (c) a simulation result and (d) a RCM result.*

Table 4.1: *Kappa and Fuzzy Kappa results for the calibration period (1996 – 2001). Scores are derived by comparing the simulated land-use maps from the land-use model and the random constraint match model with the actual 2001 land-use map.*

	Kappa			Fuzzy Kappa
	Overall	Residential	Commercial and Industrial	Overall
Simulation 1	0.866	0.871	0.750	
Simulation 2	0.866	0.871	0.752	
Simulation 3	0.866	0.871	0.750	
Simulation 4	0.866	0.872	0.751	
Simulation 5	0.866	0.871	0.752	
RCM 1	0.841	0.846	0.738	
RCM 1	0.841	0.846	0.738	
RCM 1	0.841	0.846	0.738	
RCM 1	0.841	0.846	0.737	
RCM 5	0.841	0.846	0.738	

4.4.2 Pattern analysis

Because land-use models often use randomness to simulate complex processes, some authors argue that accuracy assessment is not the appropriate way at all to measure simulation results (Power et al., 2000; Remmel and Csillag, 2003; Parker and Meretsky, 2004). As bifurcation and emergence occur in complex processes like land-use dynamics (Batty, 2005), results are generally path dependent and the same model can generate different outcomes (Brown et al., 2005). Although such outcomes do not match the actual land-use change, they may still represent realistic dynamics thus indicating a proper underlying model. Pattern-based measurements are a good alternative to assess a model's quality. In recent applications several metrics are used to measure maps, based on patch characteristics (Riiters et al., 1995), polygon matching (Power et al., 2001), or fractal analysis (Frankhauser, 1994; 2004). In this research, two pattern analyses were used to assess simulation results, both associated with fractal properties of urban systems (Batty and Longley, 1994): cluster analysis and radial analysis (White, 2006).

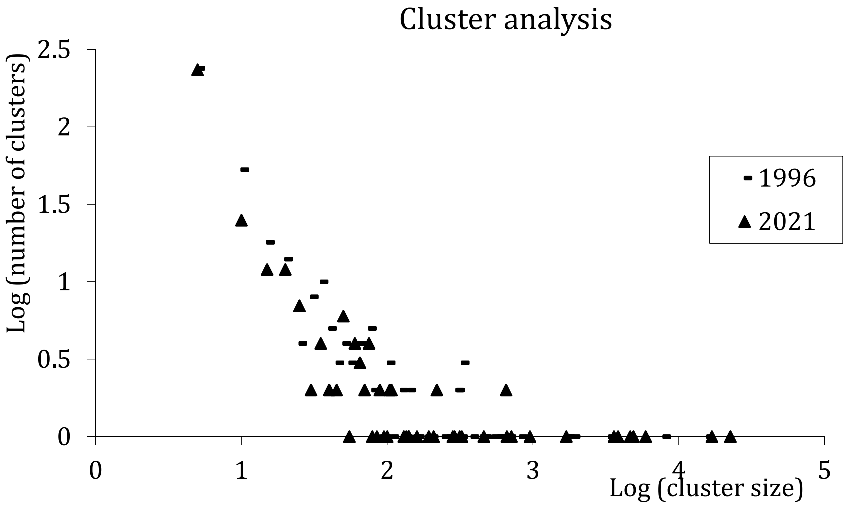


Figure 4.4: *Cluster size frequency analysis of the 1996 land-use map and the result of the simulation extended to 2021. The analysis was performed on clusters with residential land use.*

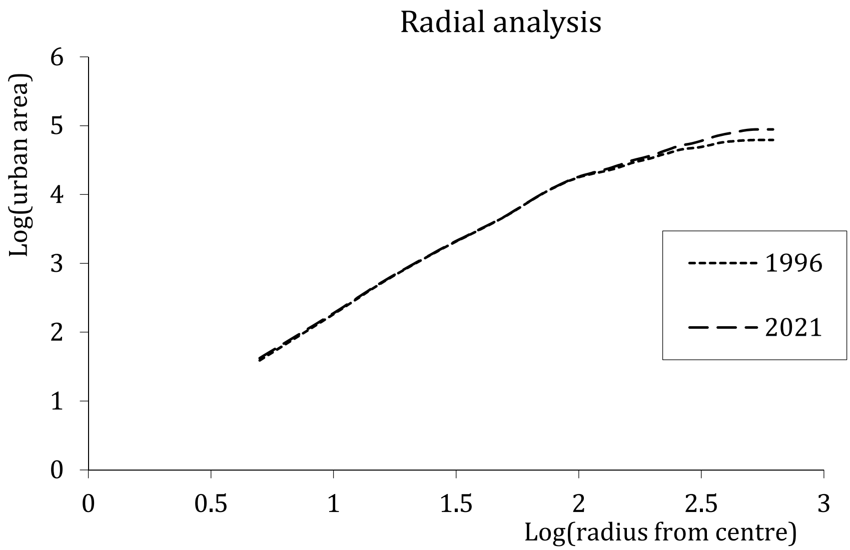


Figure 4.5: *Radial analysis of the 1996 land-use map and the result of the extended simulation. The analysis was performed on a combination of commercial and industrial, and residential cells. The boundary between the inner core and the outer zone of the urban area is visible as the bend at x=2.*

Cluster analysis measures the relation between the size and frequency of urban land-use clusters. On a logarithmic scale, this relationship is linear. Hence, it can be used to calibrate and validate urban growth models. Radial analysis investigates scaling properties by measuring cumulative area (pixels) against radius on a logarithmic scale. Processes like urban growth, which evolve outward from a nucleating centre, show such properties. On a logarithmic scale a linear relation can be observed, with a slope of 1.90 to 1.95 for dense urban centers, and approximately 1.0 in the outer urbanizing zone. A clear bend appears in the plot at the transition points between the urbanised and urbanizing points (White and Engelen, 1993). Because the amount of change over the calibration period is not very large, the simulation period was extended to 20 years, using the same rate of change. This generates enough spatial dynamics to investigate whether sufficient new clusters of urban land use appear and whether the urban area indeed maintains its characteristic radial dimensions. In this research, cluster analysis was performed on residential land use only. Radial analysis was computed for residential and commercial and industrial land use together.

Table 4.2: *RMSE scores for the simulation and two reference models for both residential, and commercial and industrial land uses. Results are obtained by comparing the number of cells per municipality for the simulation and reference results with the real 2001 land-use map.*

	Residential	Commercial and industry
Simulation 1	127	80
Simulation 2	127	80
Simulation 3	128	80
Simulation 4	128	79
Simulation 5	128	80
Constant share model	371	170
RCM 1	364	144
RCM 1	364	137
RCM 1	365	139
RCM 1	360	138
RCM 5	363	137

Results of the cluster analysis for one simulation are presented in Figure 4.4. The other four simulations show similar results. To define clusters, only rook adjacency was considered here. Clusters were aggregated in size classes, and frequencies are measured from all clusters within the boundaries of that class. The graph indicates that in general the model preserves the characteristic relationship between the cluster size and the frequency. However, from the graph it becomes clear that the simulation generates more small clusters than appear in reality. An explanation for this is the strict planning policy in the GVRD, which prevents these scattered settlements. Therefore, in reality most newly developed areas are larger patches from the beginning. This is hard to simulate in a CA environment. At the other end of the range of class sizes, an uneven distribution is visible. This uneven distribution is an effect of the local physical constraints. The shape of land between the rivers causes some urban patches not to grow any further.

Radial analysis results are presented in Figure 4.5. For reasons of visibility, only one simulation result is shown, but the other four results show similar figures. Since urban land use is a combination of commercial and industrial, and residential land uses, this analysis was performed on both land-use classes together. The centre for this radial analysis was chosen just southeast of the downtown area, where Vancouver was founded originally. In the graphs for the 1996 land-use map and the simulation result, the bend between the inner core and the outer zones is clearly visible. The difference between both graphs shows that new urban land use is mainly allocated at the fringes of the city.

Table 4.3: *Parameter values for the neighbourhood functions in three scenarios for land-use change. Values represent the attraction of a specific land use class on residential land use as a function of the distance, for the respective scenarios. Weight values are defined for discrete distance values. Weights for diagonally adjacent cells are linear interpolations.*

Level	0	1	2	3	4	5	6	7	8
Distance to central cell	0	1	3	0	27	81	243	729	2187
Standard									
Residential to residential	1000	10	0	0	0	0	0	0	0
Scenario 1									
Residential to residential	10	10	0	0	0	0	0	0	0
Forest to residential	0	0	0.005	0.01	0.005	0	0	0	0
Scenario 2									
Residential to residential	10	10	0	0	0	0	0	0	0
Water to residential	0	0	0.005	0.01	0.005	0	0	0	0

4.4.3 Regional distributions of land uses

To assess regional land-use distributions, the GVRD area was divided in municipalities. For all municipalities the modelled growth or decline in commercial and industrial, and residential land use was compared to the actual change per municipality. The root mean square error (RMSE) is used as a summary statistic for the whole map. Results for five simulation runs are presented in Table 4.2. In this table, model results are compared to the five results from the RCM model as well as a constant share model. For the constant share model, an increase in residential and commercial and industrial area was distributed over municipalities, proportionally to the existing amount of land use in these classes. This increase was equal to the overall increase in these land-use classes. These results indicate that the model performs considerably better than both the RCM model and the constant share model. Hence this indicates that it is capable of modelling regional interactions.

4.4.4 Model behavior for long range interactions

Model behavior, and specifically sensitivity analysis, is often neglected in land-use models (Kocabas and Dragicevic, 2006). In this study, only a qualitative investigation of model behavior was performed to assess the effect of land use interactions over a greater distance. To this effect, a very simple scenario was created where only the amount of residential land use increases. This residential land use is allocated using a self-attraction over a limited range, decreasing with the distance. This range of influence roughly coincides with the eight cell radius. All other possible land-use interactions were set to zero. No suitability maps, zoning restrictions or transportation network were used in this scenario, only the random perturbation term was included. This model basically creates urban growth at the edge of existing urban areas, which is what occurs in reality (Batty and Longley, 1994).

Then, to assess the effect of long distance land-use interactions, two alternative scenarios were created, indicating different preferences for new residential areas. The first includes a long range attraction from forest and natural areas to residential land use. The second includes a similar relation from water to residential land use. Weight functions used for all three scenarios are presented in Table 4.3. All three simulations were generated for a period of five years, similar to the simulations in the calibrated model, and result maps were compared with each other. Figure 4.6 presents the location of residential land use under the different scenarios. These result maps indicate that the long range interactions make a significant difference in the allocation of new residential cells.

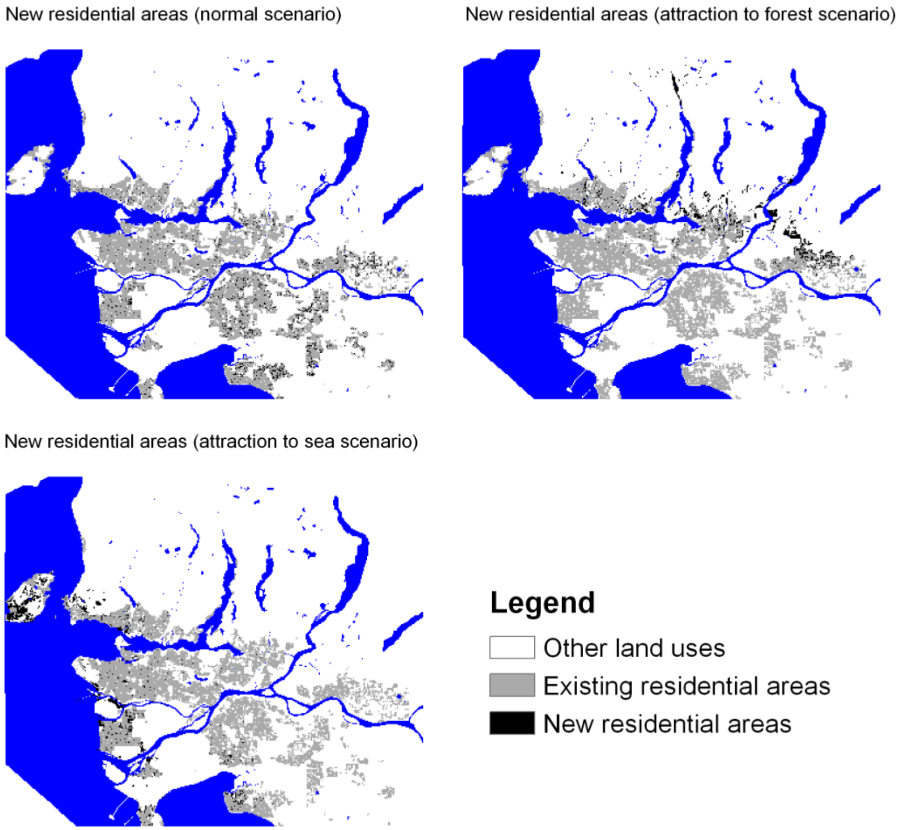


Figure 4.6: *Location of new urban land use for (a) the baseline scenario, (b) the attraction to forest and protected nature scenario and (c) the attraction to water scenario (c). Water surfaces are depicted in blue for spatial reference.*

4.5 Discussion and conclusion

In this study, an implementation of a variable grid CA was assessed for its ability to model urban dynamics and long distance land-use interactions in particular. Model results indicate that the variable grid CA approach is capable of reproducing historic urban growth and that it produces realistic urban patterns. Moreover, the effect of long distance interactions is significant in the allocation of land-use change, and simulation results improved considerably when they were used in the neighbourhood effect. This indicates a subdivision in the allocation procedure: long-distance interactions determine in which part of the area new developments take place, while the effects at short range determine the exact allocation of pixels on the land-use maps. However, errors are not distributed evenly over the municipalities. Because of strict zoning maps and a lack of transportation networks in specific areas, the model underestimates urban growth in those areas. Still, these long range interactions can be interpreted as an additional effect in land-use allocation. First, land uses for the GVRD are determined, exogenously. Then, the long-range effects determine in which part of the city people will live, while the short-range interactions determine the exact allocation within that part.

Land-use data limited the simulation period for this application to the 5 years between 1996 and 2001. This allows for a calibration, but not for independent validation. Since more recent land-use data for the GVRD was unavailable, extrapolation of the simulation could not be tested against real-world data. Moreover, simulations over longer periods, with more land use to change, might give a stronger confirmation of the variable grid concept and therefore a stronger argument for using more remote land-use information in dynamic spatial models.

5. Measuring the neighbourhood effect to calibrate land-use models

Van Vliet, J., Naus, N., van Lammeren, R.J.A., Bregt, A.K., Hurkens, J., and Van Delden, H. (Submitted). Measuring the neighbourhood effect to calibrate land-use models. Computers, Environment and Urban Systems.

Abstract

Many spatially explicit land-use models include the neighbourhood effect as a driver of land-use changes. The neighbourhood effect includes the inertia of land uses over time, the conversion from one land use to another, and the attraction/repulsion of surrounding land uses. The neighbourhood effect is expressed in the neighbourhood rules, but calibration of the neighbourhood rules is not straightforward. This research aimed to characterise the neighbourhood effect of observed land-use changes and use this information to improve the calibration of land-use models. To do this we measured the over- and underrepresentation of land uses in the neighbourhood of observed land-use changes using a modified version of the enrichment factor. Enrichment factors of observed land-use changes in Germany between 1990 and 2000 indicate that the neighbourhood effect exists. This suggests that it is appropriate to use neighbourhood rules to simulate urban land-use changes. Observed enrichment factors were used to calibrate a land-use model for Germany for 1990 to 2000 and the obtained neighbourhood rules were validated independently for 2000–2006. The results show that both the predictive accuracy and the process accuracy of the land-use model were improved in the calibration period, as well as in the independent validation period. This indicates that enrichment factors can be used to improve the calibration of the neighbourhood rules in land-use models.

5.1 Introduction

Land-use models typically include a combination of drivers to simulate land-use changes over time (Veldkamp and Lambin, 2001; Poelmans and Rompay, 2009; Wang et al., 2011). One very important driver that is included in many models is the interaction between land uses in space and in time (Irwin and Bockstael, 2002; Verburg et al., 2004b). This spatial and temporal interaction between land uses is known as the neighbourhood effect, which is represented in many land-use models by the neighbourhood rules (Hagoort et al., 2008). Examples of land-use models that include a neighbourhood effect are LUCIA (Hansen, 2007), Dyna-CLUE (Verburg and Overmars, 2009), and LUMOCAP (Van Delden et al., 2010).

Many land-use models exist as generic modelling frameworks, which can be calibrated for a specific case study application. This calibration includes the definition of the shape and parameter values of the neighbourhood rules. However, calibration of the neighbourhood rules is not a straightforward task. Several automated methods have been developed (Jeanerette and Wu, 2001; Li and Yeh, 2002; Li and Yeh, 2004; Straatman et al., 2004; Arai and Akiyama, 2004), but, despite these efforts, Hagoort et al. (2008) observe that the current practice of calibrating neighbourhood rules is predominantly manual. This is inherently subjective, not repeatable and highly dependent on the knowledge and skills of the modeller. One limitation of automated calibration methods is that most methods deal with the allocation of one land-use type only and cannot handle the interaction between multiple land uses. Another drawback of these calibration methods is that model parameters are assessed indirectly from the predictive accuracy of the simulation result: such assessment does not indicate directly which parameters should be changed and in what direction.

The research presented in this paper aimed to measure the neighbourhood effect of observed land-use changes and use this information to improve the calibration of land-use models. To do this we measured the over- and underrepresentation of land uses in the neighbourhood of observed land-use changes using a modified version of the enrichment factor (Verburg et al., 2004a). First, enrichment factors were measured for observed land-use changes to test the existence of the neighbourhood effect. The enrichment factors of the observed land-use changes were then used to calibrate the neighbourhood rules in a cellular automata land-use model. Two methods were employed to calibrate an application for land-use changes in Germany between 1990 and 2000: an automated procedure and a manual procedure. Both methods were validated independently by simulating land-use changes in Germany between 2000 and

2006. Calibration and validation results for both methods were compared with results from a reference model using a null calibration to assess their predictive accuracy and their process accuracy.

In the next section we discuss in more detail how land uses interact in space and in time and how this is reflected in the neighbourhood rules in land-use models. Section 5.3 presents the methodology for this study, including a description of the land-use model, the case study application, the details of the calibration procedures, and the assessment methods. Section 5.4 presents the simulation results and discusses these in relation to the applied calibration methods. In section 5.5 we draw conclusions and provide some directions for further research.

5.2 The neighbourhood effect

5.2.1 Inertia, conversion, and attraction/repulsion

Existing land-use patterns influence future land-use patterns in three ways: (1) through the inertia of land uses in a location, (2) through the ease of conversion from one land use to another, and (3) through the attraction or repulsion effects exerted by land uses situated in the neighbourhood of a location. The combined influence of inertia, conversion and the attraction/repulsion effects of existing land uses is known as the neighbourhood effect, which therefore includes the effects of land uses in a location as well as land uses in surrounding locations.

The existing land-use pattern is a good indication for future land-use patterns, first and foremost because the land use in most locations persists over time, at least when time is limited to periods from years to decades (Pontius et al., 2004b). The reasons for this inertia are socioeconomic as well as biophysical. Many land uses, such as residential areas or industrial activities, require a large initial investment and are therefore unlikely to change again afterwards. Similarly, some agricultural uses, like viticulture, are only profitable after some years or decades, and forests only exist because the trees were able to grow for several years. Other land uses, such as natural land uses, are influenced by biophysical circumstances, such as soil conditions and climate. These circumstances change relatively little over time, which means inertia is the rule rather than the exception for these land-use types.

When the land use in a location does change, for example because the demand for another activity increases over time, this change is highly dependent on the types of land use concerned. In areas where space is scarce or where land use is very dynamic, there may be competition between land uses for the best location. In this competition, the land uses likely to be converted are the less powerful ones, mostly in economic or political terms (Torrens, 2011). For example, urban development is often located on former agricultural land, even when the soils are very productive, because real estate developers typically have more economic influence than farmers. Similarly, many unprotected natural lands are developed for agricultural uses, even when these natural lands have a high ecological value. Hence, there is a hierarchy of land uses based on the economic or political power of the associated actors. Another factor that influences the likeliness of land conversion is the ease of conversion itself. For example, arable lands are usually reasonably flat and therefore easier to develop into urban land than dense forests on steep slopes.

The attractiveness of a location for a new use is influenced not only by the existing land use in the location itself, but also by the surrounding land uses. For instance, while it is mostly agricultural land that is converted into urban land, it is typically only those locations in the vicinity of existing urban land that are urbanised. More generally, the interaction between land uses and their associated actors can be expressed as a mutual attraction or repulsion (Hagoort et al., 2008). Neighbouring land uses can represent push and pull actors that shape land-use patterns (Anas et al., 1998; Krugman, 1998; Hansen, 2012). Examples of land-use relations are nuisances like noise and smoke from industrial sites that have a negative influence on the attractiveness of adjacent locations for residential land use, and nearby forests that have a positive effect as they provide clean air and opportunities for recreation.

5.2.2 Representation of the neighbourhood effect in land-use models

The notion that land uses are in competition for the most favourable locations was already acknowledged in some of the earliest land-use models: the Von Thünen model (Von Thünen, 1826) and Alonso type models (Alonso, 1964; Anas et al., 1998) allocate the economically most powerful land uses to the most favourable locations. The most favourable location in these cases is the location closest to the city centre, which is taken as the central market, because the distance to this central market determines transportation costs. These models implicitly include the competitive hierarchy of land uses and their associated actors. However, they describe a static situation and do not treat land-use change explicitly: inertia, conversion and attraction/repulsion are not included.

Another type of land-use model is based on Markov Random Fields (MRFs). MRF models are dynamic models and are able to simulate land-use inertia as well as conversion. They do not explicitly include hierarchy and competition between land uses, but their combined effect is expressed in the transition probabilities, which can be measured from data. For example, a study by Rutherford et al. (2008) indicates that most land uses persist year on year, while a study by Zhang et al. (2011) shows that new urban land is mostly allocated on agricultural land. The attraction or repulsion of neighbouring land uses, however, is not included in MRF models.

On the other hand, land-use models based on Cellular Automata (CA) include inertia, conversion and attraction/repulsion as drivers for land-use changes. The defining element of CA are the neighbourhood rules, which express the influence exerted on land-use dynamics by both the land use in a location and the land uses in neighbouring locations. Most contemporary CA models are able to simulate multiple types of land-use changes (Arai and Akiyama, 2004; Van Delden and Hurkens, 2011; Wang et al., 2011). It should be noted that CA models often include other drivers for land-use change as well, such as accessibility to transport networks, landscape elements, and spatial planning measures.

5.2.3 Neighbourhood rules

Spatially explicit land-use models, such as CA models, typically consist of a lattice of square cells, where the cell state represents the predominant land use in that location. The neighbourhood rules can therefore be defined as a function of the land uses in all cells in the neighbourhood and their distance to the location of interest. Inertia and conversion are the effects exerted by the land use in a cell itself, while attraction or repulsion are the effects exerted by cells at distance > 0. Because spatial actors typically consider a larger area in their allocation decisions (White et al., 1997; Verburg et al., 2004b; Van Vliet et al., 2009), neighbourhood rules often include more locations than only the directly adjacent locations. Consistent with Tobler's first law of geography (Tobler, 1970), the influence of neighbouring land uses typically decreases with increasing distance, and will eventually approach zero.

The influence of a land use on its own location and the influence the same land use exerts in the vicinity of its location can be different. For example, commercial land use can attract residential developments to nearby locations, whereas it is unlikely to convert into residential land itself. Additionally, the influence exerted by a land use can differ qualitatively at different distances. For example, industrial land use can have a repulsion effect on residential land use over a short distance because of noise and air pollution, but an attraction effect at greater distances because it generates employment. Therefore, the neighbourhood rule that describes the interaction between a pair of land uses can be simplified to the effect at a location itself, the effect at a short distance from the location, and the effect at greater distances. Combinations of these three effects yield 8 basic types of neighbourhood rules, which are shown in Figure 5.1.

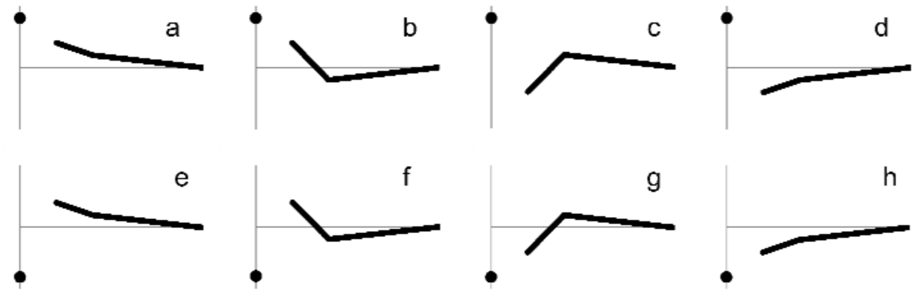


Figure 5.1: *Possible types of neighbourhood rules based on the combination of the effect at a location itself, the effect at a short distance from the location and the effect at a greater distance (adapted from White and Engelen (2003) and Hagoort et al. (2007)). Values on the x-axis indicate the distance to a land-use change location, while values on the y-axis represent attraction and repulsion.*

The challenge in the application of land-use models that include neighbourhood rules is to find the appropriate shape and parameter values for these neighbourhood rules (White and Engelen, 2003; Straatman et al., 2004). Empirical estimation of the parameters in neighbourhood rules is not possible for several reasons (Verburg et al., 2004a). First, many land uses are highly autocorrelated, and therefore the attraction or repulsion effect exerted by one location of a particular land use cannot be determined independently. Second, land-use dynamics are typically path dependent: changes at one point in time can influence changes at later points in time. Because data is typically only available for a few moments in time, it is very difficult to investigate the effect of these feedback mechanisms. Third, site characteristics, such as the presence of transportation networks, physical characteristics or policy regulations, also influence allocation decisions. These drivers are highly correlated with the existing land-use pattern and are therefore not independent of the neighbourhood effect itself (Irwin and Bockstael, 2002). It is for these reasons that parameters in the neighbourhood need to be calibrated by some manual or automated procedure.

Calibration is essentially an iterative process that includes an assessment of the current calibration result, followed by an adjustment of one or more parameters based on this assessment to improve the calibration. However, most assessment methods, such as map-comparison methods, do not directly generate directions in which parameters should be adjusted. The calibration procedures presented in this paper use the over- or underrepresentation of land uses in the neighbourhood of land-use changes as a measure to improve the calibration of the neighbourhood effect. This is an important improvement over other assessment methods as it allows consideration of each combination of land uses separately, and therefore provides insights into which parameter should be adjusted to improve calibration results.

5.3 Methodology

5.3.1 Characterization of the neighbourhood effect

Verburg et al. (2004a) introduced the enrichment factor to characterize the neighbourhood effect. Enrichment factors are defined as the over- or underrepresentation of a land use in the neighbourhood of a particular location, relative to the average land-use distribution:

$$F_{i,l,d} = \frac{n_{i,l,d}/n_{i,d}}{|N_l|/|N|} \qquad \text{Equation 5.1}$$

where $F_{i,l,d}$ is the enrichment of land use l on location i at distance d, $n_{i,l,d}$ is the number of cells of land-use type l at distance d of location i, $n_{i,d}$ is the total number of cells at distance d from location i, $|N_l|$ is the number of cells with land use l, and $|N|$ is the total number of cells in the study area. Distance is measured as the Euclidean distance between the centres of two cells. The enrichment factor at distance 0 relates to the inertia effect or conversion effect, while enrichment factors at distance > 0 relate to the attraction or repulsion exerted by neighbouring land uses. This method was applied earlier to the Netherlands and showed that the location of some combinations of land uses are strongly correlated (Verburg, 2004b).

Enrichment factors can be computed for any subset of cells, for example for all locations that changed into a selected land use in a given period, derived from the difference between two land-use maps. This is not necessarily similar to the enrichment factor measured from all locations with a particular land use on a single map, because land-use patterns are known to be path dependent (Brown et al., 2005). The above-mentioned selection of land-use changes implicitly assumes that land use is actively allocated, not actively removed, as it measures the attraction effect on newly allocated land uses and not the repulsion effect on the replaced land uses. This is typically what happens in land-use changes like urban growth or agricultural expansion, and is similar to what many land-use models do effectively. Therefore, it is the location of land-use changes that is of interest for the calibration of land-use models.

For a specific land use k, the enrichment factor for newly allocated land uses can be computed as follows:

$$F_{k,l,d} = \frac{\frac{1}{|N_k|}\sum_{j\in N_k} n_{j,l,d}/n_{j,d}}{\frac{1}{|N|}\sum_{i\in N}|N_l|/|N|} \qquad \text{Equation 5.2}$$

where $F_{k,l,d}$ is the enrichment of land use l on all locations that changed into land use k at a distance d and $|N_k|$ is the number of cells that changed into land use k between T1 and T2. Neighbouring land uses are measured from the land-use map at T1. Equation 5.2 differs from Equation 5.1 and from Verburg et al. (2004a) in that the reference neighbourhood is computed based on cells in the study area only, rather than all cells on the map. This corrects for edge effects, as land uses outside the study area can influence land-use changes, while they themselves are not analysed. This is the case, for example, with marine water in coastal areas.

Equation 5.2 was used to measure the enrichment factors of observed land-use changes to test the existence of the neighbourhood effect. These were measured for Germany in the period 1990–2000 and 2000–2006 using the data that was subsequently used for the calibration and the independent validation of the cellular automata (CA) based land-use model. In the rest of this paper, the logarithm of the enrichment factor will be used to express over- and underrepresentation on a comparable scale. Positive values indicate that a land use is overrepresented, while negative values indicate an underrepresentation.

5.3.2 The CA model and case study applications

For this study we used the Metronamica land-use model (Van Delden and Hurkens, 2011), which is a constrained CA model (White et al., 1997). Metronamica uses three types of land-use classes: function, feature and vacant land uses. Function land uses are actively allocated by the model. Generally, urban land uses are represented as function classes. Feature land uses are land uses that do not change location during a simulation. Typical examples are water bodies or infrastructure elements. Vacant land uses are assigned to all locations that are not occupied by a function or feature land use, based on their suitability only. These often include natural vegetation and sometimes agricultural land uses.

In each time step, representing one year, function land uses are allocated to those locations that have the highest potential for this land use. Potentials are computed for each cell and for each land use based on transition rules:

$$P_{k,i} = \nu \cdot A_{k,i} \cdot S_{k,i} \cdot Z_{k,i} \cdot N_{k,i} \qquad \text{Equation 5.3}$$

where $P_{k,i}$ is the potential for land use k in cell i, ν is a scalable random perturbation term for land use k in cell i, $A_{k,i}$ is the accessibility for land use k in cell i, $S_{k,i}$ is the physical suitability for land use k in cell i, $Z_{k,i}$ is the zoning status for land use k in cell i, and $N_{k,i}$ is the influence of the neighbourhood rules for land use k in cell i. The neighbourhood rules contain the combined effect of the land use of all cells j including i in the neighbourhood:

$$N_{k,i} = \sum_j w_{d(i,j),k,l(j)} \qquad \text{Equation 5.4}$$

where $w_{d,k,l}$ is the parameter value in the neighbourhood rules that describes the effect of land use l at distance $d(i,j)$ on the potential of location i for land use k. Parameters in the neighbourhood rules are constant over time. Note that the absolute values of parameters in the neighbourhood rules have no intrinsic meaning, but they have a meaning relative to each other.

Two Metronamica applications were set up to simulate land-use changes in Germany: one calibration application, which simulates changes from 1990 to 2000, and one independent validation application, which simulates changes from 2000 to 2006. Both applications are strictly comparable in their parameter settings and model properties, such as cell size, temporal resolution, and land-use classes.

Land-use data for both applications were taken from the Corine land-cover database (Haines-Young and Weber, 2006). This database provides a comparable dataset for the start year and the end year of both periods. Original Corine land-cover classes were reclassified into eight aggregate land-use classes. These classes are: natural areas and agricultural areas, which are modelled as a vacant land use; recreational areas, residential land, and industrial areas (including commercial areas), which are modelled as function land uses; and mineral extraction and landfill sites, infrastructure, and water, which are modelled as feature land uses. This characterization represents the assumption that the selected function land uses have an explicit area demand, depending on population dynamics and economic developments. After reclassification, the original 100 metre raster data were resampled into a 500 metre resolution using a majority aggregation. This resampling was required to reduce the total number of cells to fit the computational capacities of the software.

In addition, data and parameters for suitability, zoning and accessibility were included in both applications. These data and parameters were taken from comparable applications, such as the LUMOCAP system (Van Delden et al., 2010). Specifically, we included the influence of elevation and slopes in the suitability, the influence of the Natura 2000 network in zoning, and the main road network for the accessibility effect.

5.3.3 Calibration of the neighbourhood rules

Since the aim of this study was to improve the calibration of the neighbourhood rules in land-use models, the calibration procedure was clearly separated from the assessment of model results after calibration. Starting from a null calibration, two parameter sets were obtained using different procedures: a manual calibration procedure and an automated calibration procedure. The parameters found using these calibration procedures were subsequently used to simulate land-use changes for the calibration period (1990–2000) and for the independent validation period (2000–2006). Land-use maps generated using these parameters sets were then assessed and compared with those generated by the null calibration. Figure 5.2 provides an overview of the complete calibration and assessment procedure.

The starting point for the calibration of the neighbourhood rules is the null calibration. The null calibration contains a very simple set of neighbourhood rules in which all function land uses have an inertia value of 100, while all possible conversions have a value of 1. All neighbourhood rules at distance > 0 are set to zero; hence there is no attraction or repulsion. This means that land uses typically persist in their current location, while all other locations have an equal but small chance to convert. The null calibration was used as the starting point for two different calibration procedures: a manual calibration procedure and an automated procedure. The null calibration was also used as a reference for comparison with other calibration results.

The manual calibration procedure uses a visual interpretation of the comparison between the enrichment factors of observed and simulated land-use changes. Based on this comparison, neighbourhood rules are adjusted manually, while keeping in mind the possible shapes as shown in Figure 5.1. This means that the shapes of neighbourhood rules were assessed visually. Each step in this calibration procedure can include the adjustment of parameter values at several different distances, or even from several different combinations of land uses.

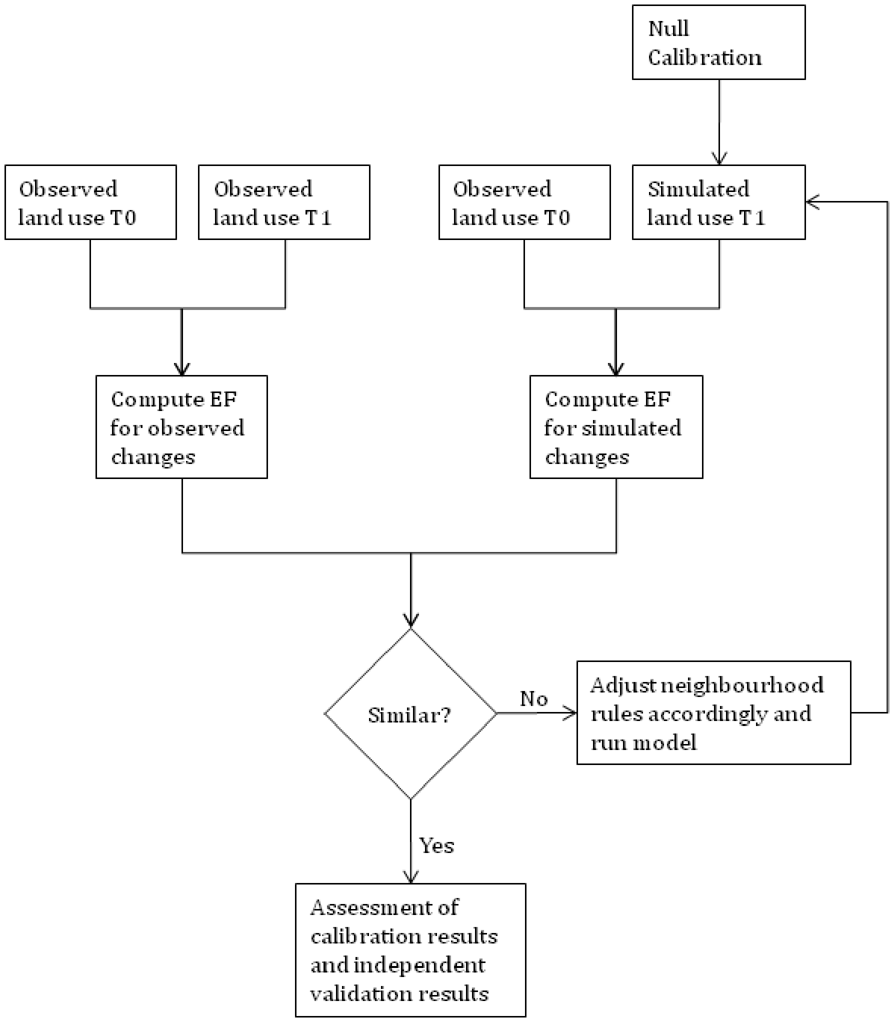


Figure 5.2: *Flowchart of the procedure used to calibrate the neighbourhood rules.*

The automated calibration procedure systematically compares the enrichment factor of observed land-use changes with those of simulated land-use changes for all distances and all combinations of land uses. The difference between observed and simulated land-use changes is computed as follows:

$$Deviation_{k,l,d} = log\left(\frac{\left(\frac{1}{|N_k^{Sim}|}\sum_{i\in N_k^{Sim}} n_{i,l,d}/n_{i,d}\right)/\left(\frac{1}{|N|}\sum_{i\in N}|N_l|/|N|\right)+c}{\left(\frac{1}{|N_k^{Data}|}\sum_{j\in N_k^{Data}} n_{j,l,d}/n_{j,d}\right)/\left(\frac{1}{|N|}\sum_{i\in N}|N_l|/|N|\right)+c}\right)$$

Equation 5.5

where $Deviation_{k,l,d}$ is a measure for the difference in the presence of land use l at distance d in the neighbourhood of locations that change into land use k in the simulated land-use changes and the observed land-use changes, $|N_k^{Sim}|$ is the number of cells that changed into land use k in the simulations, $|N_k^{Data}|$ is the number of cells that changed into land use k in the data, and c is a constant ($c > 0$). Constant c was added to avoid too much emphasis on land uses that make up a relatively small proportion of the map, because small absolute differences in the neighbourhood translate into large relative changes. Similar to the enrichment factor, Equation 5.5 uses the logarithm to express over- and underrepresentation on a comparable scale. To limit the number of parameters in this method, distances were divided into discrete rings of (cell) unit distance. Hence $Deviation_{k,l,2}$ refers to the influence of land use l at a distance between 1.5 and 2.5 cells on all locations that changed into land use k. Correspondingly, parameters in the neighbourhood rules are only defined at each unit distance, and interpolated linearly in between. Consequently, theoretical shapes of neighbourhood rules as shown in Figure 5.1 were not included in this procedure.

The automated calibration procedure adjusts parameters one by one, starting from the parameter set used in the null calibration. Using Equation 5.5 to compare the neighbourhood of simulated and observed land-use changes, the land-use combination and distance that showed the largest deviation were selected and the associated parameter adjusted accordingly. For example, when the largest difference is an underestimation of the number of cells with residential land at a distance of 1 cell from the locations that changed into industrial land, the parameter that expresses the attraction from residential to industrial land at distance 1 was increased. Since the number of cells allocated in the simulation is constrained exogenously, the total areas of each land-use class in the data and in the simulation result are similar. This means that deviations other than 0 are entirely due to the allocation of land uses. To improve the calibration result we applied the following procedure iteratively: run the simulation using the current parameter set, find the largest deviation, and adjust the associated parameter in the neighbourhood rules. Theory (Hagoort, 2008), empirical measurements (Verburg et al., 2004b) and experience from other land-use models (Van Vliet et al., 2009) all indicate that the influence of neighbouring land uses decreases with increasing distance. To anticipate this distance-decay effect, the size of adjustments was decreased with increasing distance, from a step size of 1 at distance 0 to $1 \cdot 10^{-4}$ at distance 4. At distance 5 and higher, the neighbourhood rules were set to zero. This systematic procedure was implemented directly in the Metronamica land-use model.

Both calibration methods differ in two ways from other studies that employ enrichment factors to simulate land-use changes (Verburg et al., 2004a; Geertman et al., 2008; De Nijs and Pebesma, 2010): first we computed enrichment factors only for locations that change land use, not for all locations; and second, we used the enrichment factor as a calibration measure, not as a model parameter.

5.3.4 Assessment of model results

To assess the two calibration procedures, their performance was tested for predictive accuracy and process accuracy (Brown et al., 2005). The first indicates how accurately land uses are allocated; the latter indicates whether land-use change processes are simulated realistically. Since the latter is hard to assess directly, process accuracy is typically assessed indirectly based on generated land-use patterns. In our assessment, the simulation results were compared with the results of the null calibration for the calibration period and also for the independent validation period. This was done using a visual comparison of the simulation results as well as using objective measures.

The predictive accuracy of land-use change simulations was assessed using Kappa Simulation (Van Vliet et al., 2011). This method expresses the similarity between simulated land-use changes and observed land-use changes, corrected for the amount of change. Kappa Simulation was measured for all land-use classes combined and for each function land-use class separately.

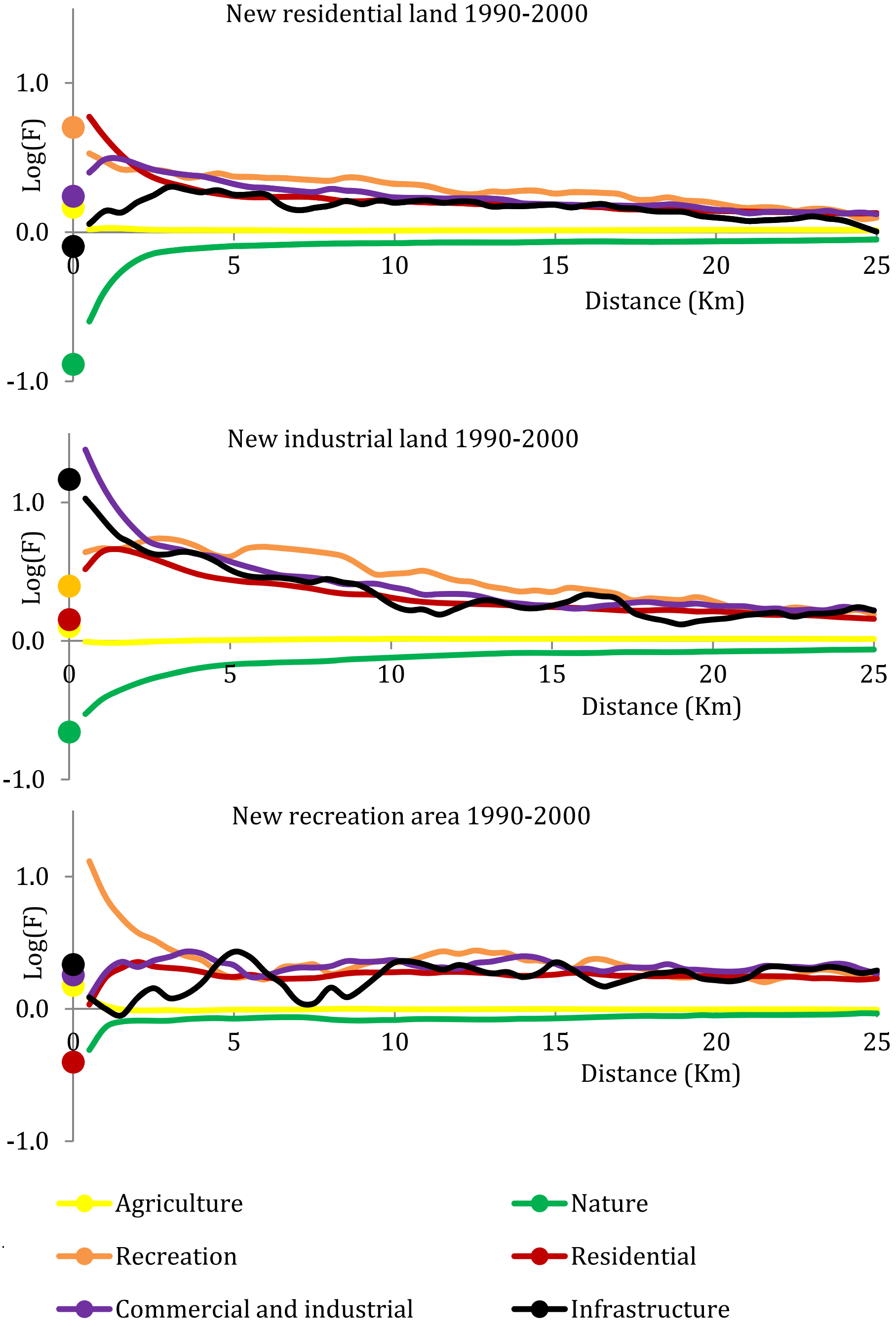
Process accuracy was assessed using the fractal dimension of patches of residential and industrial land combined. The fractal dimension is a metric that can be used to express the complexity of land-use patches. Several researchers have indicated that complexity is characteristic for patches of urban land (Huang et al., 2007; Schwarz 2010; Chen, 2011) and a realistic model will simulate land-use patterns that are similar to observed land-use patterns. We computed the fractal dimension using the FRAGSTATS definition (McGarigal et al., 2002).

5.4 Results and discussion

5.4.1 Enrichment factors of observed land-use changes

Figure 5.3 shows the over- or underrepresentation of land uses in the neighbourhood of residential land and industrial land that appeared in Germany between 1990 and 2000 and between 2000 and 2006 as a function of the distance to the locations of land-use changes. The graphs show two separate effects: the value at distance zero in the graphs indicates the conversion effect, while values at distance greater than zero indicate the over- or underrepresentation of land uses in the neighbourhood of locations of land-use changes. Positive values indicate an overrepresentation and negative values indicate an underrepresentation of the respective land uses. Results were computed for all combinations of land uses, however Figure 5.3 only shows results for land uses that were modelled as a function land use. Similar graphs for other land uses suggest that urban land uses have the strongest neighbourhood effect. Since we measured enrichment factors for changing land uses only, inertia could not be computed.

The graphs in Figure 5.3 show that new urban land uses tend to be located next to other urban land uses, but not next to agricultural or natural land uses. For example, commercial land uses are overrepresented in the neighbourhood of new residential land uses, and residential land uses are overrepresented in the neighbourhood of new commercial and industrial areas. It is tempting to translate under- and overrepresentations of land uses in terms of attraction and repulsion, but this is not always correct. For example, new residential land use is found less than average in locations near natural vegetation. This does not necessarily indicate that natural vegetation actively repels residential developments; other land uses could assert a greater attractive effect, or other drivers such as accessibility or protection of natural landscapes could explain this. Alternatively, over- or underrepresentation may be an indirect effect. For example, agricultural land is overrepresented in the vicinity of new residential areas, which could be explained by the fact that existing residential land attracts new residential uses, while existing residential areas are frequently surrounded by agricultural areas.


New residential land 1990-2000
Log(F)
1.0
0.0
-1.0
0
5
10
15
20
25
Distance (Km)
New industrial land 1990-2000
Log(F)
1.0
0.0
-1.0
0
5
10
15
20
25
Distance (Km)
New recreation area 1990-2000
Log(F)
1.0
0.0
-1.0
0
5
10
15
20
25
Distance (Km)
Agriculture
Nature
Recreation
Residential
Commercial and industrial
Infrastructure

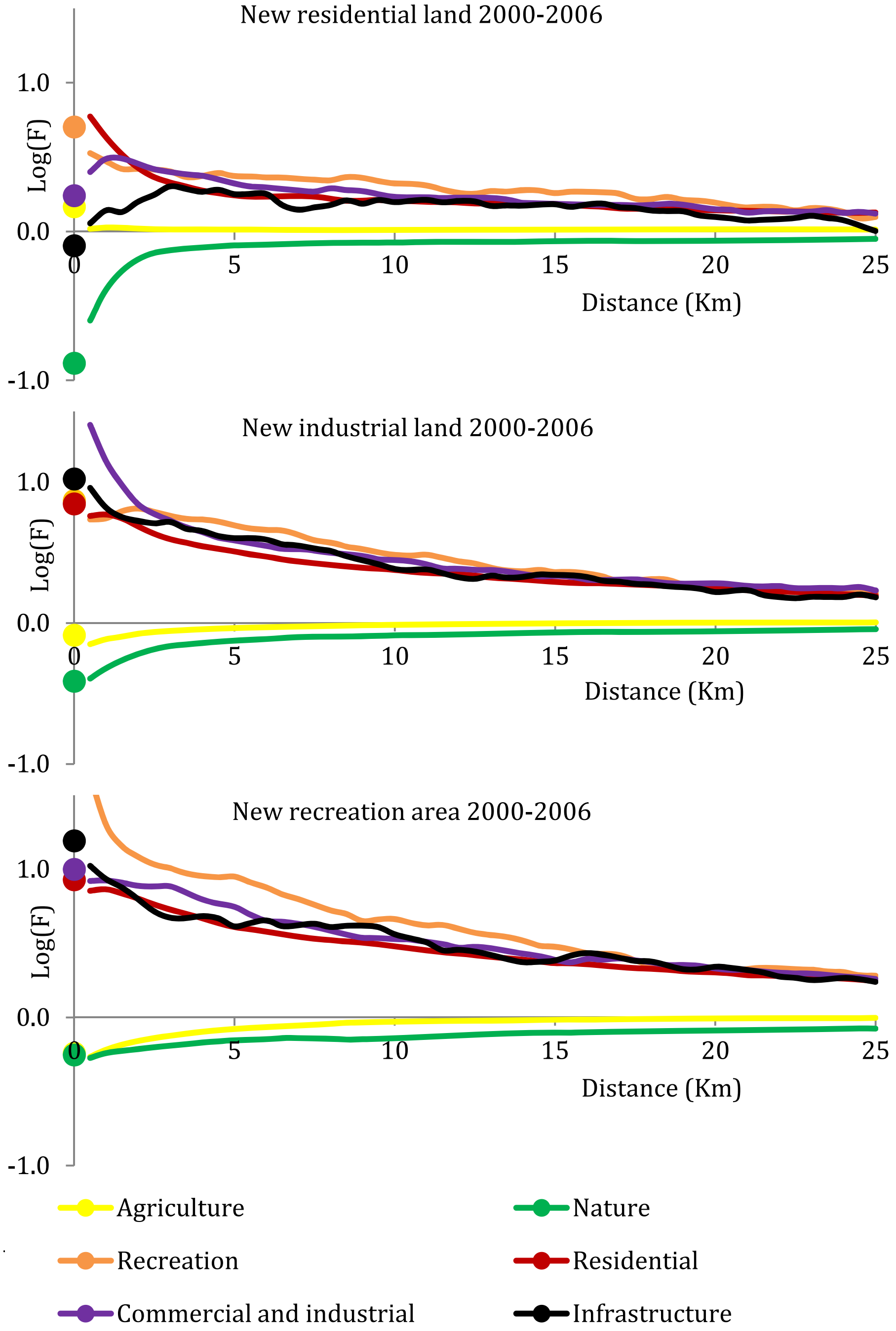


Figure 5.3: *Over- and underrepresentation of land uses in the neighbourhood of observed land-use changes between 1990 and 2000 and between 2000 and 2006.*

Figure 5.3 also shows that most of the possible land-use relations are either entirely positive or entirely negative. Hence, changes from over- to underrepresentation and vice versa (sign changes), as shown in Figure 5.1, are an exception, and two sign changes (as depicted in figures 5.1c and 5.1f) do not occur at all. Moreover, enrichment factors generally show a distance decay effect, as the values decrease with increasing distance. This decrease is not always smooth. Such fluctuating values are most visible in those changes for which there are fewer observations (e.g. the location of new recreational areas between 1990 and 2000). These local optima are therefore probably an artefact caused by this limited number of observations rather than a local optimum in the relation between both land uses.

Measured enrichment factors confirm the existence of a neighbourhood effect, which suggests that it is appropriate to include neighbourhood rules to simulate urban land-use changes. Moreover, measured enrichment factors for changes between 1990 and 2000 are quite similar to the changes between 2000 and 2006, which indicates that land-use relations are rather static over time periods typical for land-use models. Therefore, it seems appropriate to use a historic calibration to set neighbourhood rules for ex-ante studies or scenario studies.

5.4.2 Calibration of the neighbourhood rules

Figure 5.4 shows the over- or underrepresentation of residential land use in the neighbourhood of new residential land uses for observed land-use changes and simulated land-use changes in all calibration procedures. Results are shown for the calibration period and for the independent validation period. From these graphs it is clear that both the manual and the automated calibration procedure improve the simulation results considerably compared with the null calibration: the enrichment factors measured from simulated land-use changes became much more similar to the enrichment factors of observed land-use changes. It is also noteworthy that the neighbourhood rules obtained by both calibration methods only include attraction or repulsion within a 4-cell, 2-kilometre radius, while simulated land-use relations approach observed land-use relations over a much greater distance. This can be explained by the autocorrelation between land uses: when a location is neighbouring, say, residential land uses, there is a large probability that there is also more residential land use than average at some greater distances. These results indicate that similar neighbourhood rules yield large improvements for the calibration application (1990–2000) and the independent validation (2000–2006). In the independent validation period, the neighbourhoods of observed changes match those obtained from the manual calibration procedure a little closer than those obtained through the automated calibration period. However, the differences between both procedures are small.

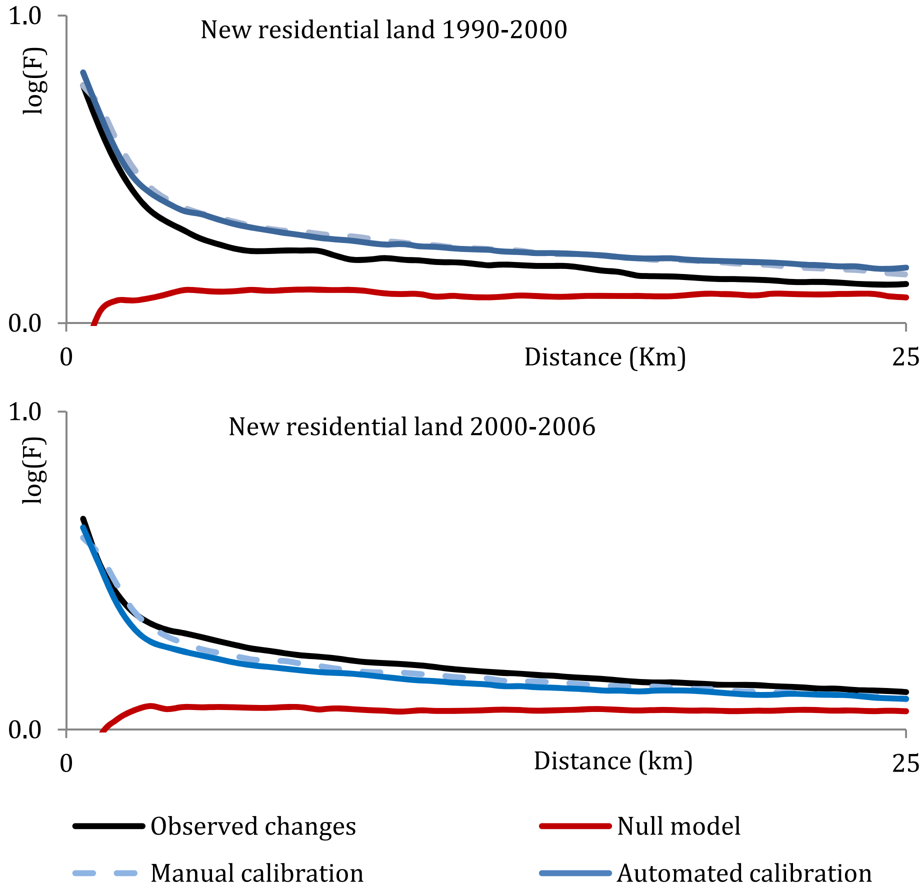


Figure 5.4: *Over- and underrepresentation of residential land use in the neighbourhood of new residential land use for observed land-use changes and simulated land-use changes in the calibration period (1990-2000) and the validation period (2000-2006).*

The neighbourhood rules obtained by the manual and automated calibration procedures are quite similar: both procedures yielded a mutual attraction between the urban land-use classes and both calibration procedures generally yielded rules that decrease with increasing distance. However, the rules generated by the automated calibration are less smooth and occasionally show sign changes between the attraction at a short distance and at a greater distance. Although these sign changes could happen in reality, their abruptness suggests that they are the results of path dependency in the automated calibration: a relatively high value at, say, distance 1 is then compensated by a relatively low value at distance 2. Generally, the higher the number of observed land-use changes, the smoother the obtained neighbourhood rules. Therefore, the application of such an automated procedure is more appropriate for land uses that show many changes, while it is less suitable for small case study applications. The manual calibration did not yield such sign changes or abrupt changes as there was no indication from theory or data that these were appropriate. In addition, both the data and the automated calibration procedure show a greater number of conversions in urban land uses, while the manual procedure yields a more persistent (urban) land-use pattern. In this respect, the manual calibration is probably more realistic because it is unlikely that these locations change from urban land to arable or natural land in reality. Conversions in the data might be due to errors in the data, or artefacts introduced by the aggregation of the land-use data.

Finally, the automated procedure was very sensitive to the value for c. We found an optimal result at values around 0.25, because larger values put too much emphasis on the conversion effect and predominant land uses, while smaller values put too much emphasis on the values at greater distances and scarce land uses. A similar trade-off is made implicitly by the modeller in the manual calibration procedure.

5.4.3 Simulation results

The simulation results of the null model and the results of both calibration procedures were assessed for their predictive accuracy and their process accuracy. Scores for the predictive accuracy measured using Kappa Simulation are presented in Table 5.1. Because the land-use model includes a stochastic term, the results of several runs were assessed. However, the differences in predictive accuracy were small (<0.001) and so results are only shown for one run. The table shows that both the automated and the manual calibration procedure improved overall predictive accuracy as well as predictive accuracy for each function land use separately. A similar observation can be made for the independent validation, which again confirms that both applications produce rules that are quite general in time, and not specific to one period. However, even though the automated calibration uses many more iterations and compares enrichment factors systematically, it does not outperform the manual calibration: the manual calibration yielded better results for the entire land-use map as well as for industrial land use. Nevertheless, in both cases the Kappa Simulation scores are not very high, which indicates that the location of land-use changes is still quite uncertain. Another explanation of these low Kappa Simulation values is that the calibration procedure used in this study only focuses on the neighbourhood rules and only uses enrichment factors to improve the calibration. Most other land-use model studies will consider a combination of all parameters, and the predictive and process accuracy would already have been assessed during the calibration period to further improve model performance.

Table 5.1: *Accuracy assessment and pattern analysis of the calibration results per land use. The results that are most accurate are underlined.*

	Kappa Simulation			
	All land uses	Residential land	Industrial land	Recreation
1990-2000 (Calibration)				
Null calibration	0.003	0.003	0.002	0.001
Manual calibration	<u>0.034</u>	0.038	<u>0.052</u>	0.003
Automated calibration	0.026	<u>0.044</u>	0.026	<u>0.004</u>
2000-2006 (Independent validation)				
Null calibration	0.002	0.003	0.003	0.001
Manual calibration	<u>0.016</u>	0.037	<u>0.038</u>	0.009
Automated calibration	0.014	<u>0.041</u>	0.017	<u>0.010</u>

The process accuracy was assessed using the fractal dimension of clusters of urban land use, which is comprised of residential and industrial land combined. The closer their fractal dimensions are to those measured from the actual land use map, the more realistic the simulation result. These measurements again show that both the manual and the automated calibration procedure are more realistic than the null calibration, and that the generated urban patterns are quite similar to those observed in reality. This indicates that the neighbourhood effect is better at generating realistic land-use patterns than precise allocations of land-use changes. It is not unexpected that neighbourhood rules simulate the process accuracy better than the predictive accuracy, because neighbourhood rules represent the behaviour of spatial actors and therefore represent the process underlying land-use changes. This is also in line with earlier observations that there is a limit to the predictive accuracy with which land-use changes can be simulated, due to the inherent uncertainty of the behaviour of associated actors (Batty and Torrens, 2005; Manson, 2007).

Table 5.2: *Fractal dimensions of observed land-use map and simulation results. Underlined values indicate those simulation results that best match the 2006 land use patterns.*

Land-use map	Fractal dimension
1990 data	1.344
2000 data	1.346
Null calibration 2000	1.387
Manual calibration 2000	1.358
Automated calibration 2000	<u>1.354</u>
2006 data	1.380
Null calibration 2006	1.427
Manual calibration 2006	<u>1.381</u>
Automated calibration 2006	1.377

Finally, Figure 5.5 shows the land-use changes between 1990 and 2000 as observed in reality as well as those simulated by the null calibration and by both calibrated models. This figure illustrates the differences between the procedures: the null calibration (Figure 5.5b) shows scattered urban developments, the manual calibration (Figure 5.5c) has a high persistence and consequently little urban land use disappears, while the automated calibration (Figure 5.5d) shows quite some disappearance of urban land and a somewhat scattered land-use pattern.

Although any visual inspection remains subjective, this figure clearly shows that the null calibration generated less realistic land-use changes than both calibrated models, as it only allocates single-pixel patches, while in reality larger patches of land-use change also appeared. The difference between the two calibrated models is more difficult to discern, but in our opinion the automated calibration procedure generated a land-use pattern that is most similar to the actual land-use pattern, as the patches of urban land are less blobby.

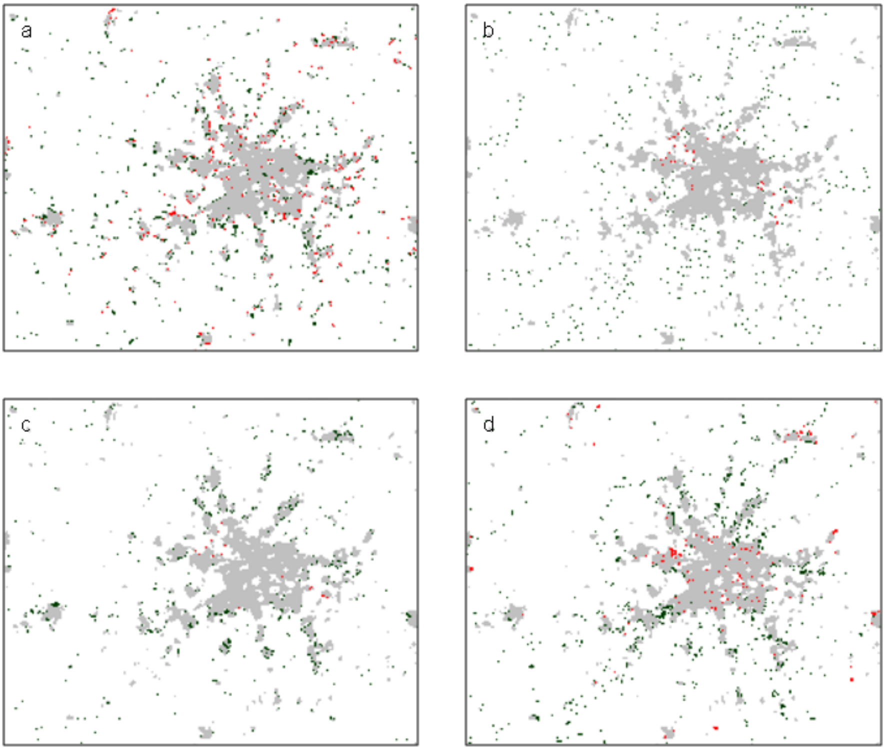


Figure 5.5: *Visual comparison of observed land-use changes (a) and land use changes using the null calibration (b), the manual calibration (c) and the automated calibration (d). Grey pixels indicate persiting urban land use, red pixels indicate disappering urban land, and green indicate appearing urban land.*

5.5 Conclusions

An analysis of observed land-use changes indicates that there is a relation between locations of land-use change and the land use in their neighbourhood. This was found for all combinations of land uses, but predominantly for the allocation of urban land uses. Measured relations between 1990 and 2000 were similar to those observed between 2000 and 2006, which indicates that the behaviour of spatial actors is quite constant over the investigated time period.

The neighbourhood effect is represented in cellular automata (CA) land-use models in the neighbourhood rules. In this study we applied a manual and an automated procedure to calibrate a CA model, both of which use enrichment factors as a calibration measure. Simulation results were assessed for their predictive accuracy and their process accuracy. In addition, the shape of the obtained neighbourhood rules and the generated land use maps were assessed visually. Both procedures improved the results considerably in the calibration application as well as the independent validation, which suggests that enrichment factors are useful as a calibration measure. The biggest improvement was in the process accuracy, which shows that neighbourhood rules represent the behaviour of spatial actors, while the exact allocation of these changes remains uncertain.

This research provides some directions for further research. The manual calibration has the drawback that it is highly dependent on the modeller performing the calibration, while the automated calibration procedure is systematic and repeatable. However, the results obtained from the automated procedure do not yet outperform the manual calibration. The automated procedure can probably be improved by introducing theoretical knowledge about the possible shapes of neighbourhood rules, because such knowledge is used implicitly in the manual calibration. Moreover, this study treated the calibration of neighbourhood rules only, while in reality the location of land-use changes is also driven by other factors, such as accessibility or physical landscape characteristics. Consequently, a more comprehensive calibration technique should address multiple drivers, whereas their effect was taken as a given in this research.

6. An activity-based cellular automaton model to simulate land-use dynamics

Van Vliet, J., Hurkens, J., White, R., and Van Delden, H. (2012). An activity based cellular automaton model to simulate land-use changes. Environment and Planning B 39: 198-212. (Slightly adapted in order to improve consistency with other chapters)

Abstract

In recent decades, several methods have been proposed to simulate land-use changes in a spatially explicit way. In these models, land is generally represented on a lattice with cell states representing the predominant land use. Since a cell can have only one state, mixed land uses and different densities of one land use can only be introduced superficially, as separate cell states. This paper describes a cellular automata based model that simulates dynamics in both land uses and activities, where activities represent quantitative information, such as the number of inhabitants on a location. Therefore each cell has associated with it (1) a value representing one of a finite set of land-use classes, and (2) a vector of numerical values representing the quantity of each modelled activity that is present on that location. This allows simulating incremental changes as well as mixed land uses. The proposed model is tested with a synthetic application that uses two activities: population and jobs. It simulates the emergence of human settlements over time from local interactions between activities and land uses. Assessment of results indicates that the model generates realistic urbanization patterns.

6.1 The ever-changing world around us

Land use is constantly changing. In urban areas new houses are being built at some locations while older ones are being demolished at elsewhere. Brownfield sites are being improved and industrial areas are reallocated outside city centres. Similar observations can be made outside urban areas. Some farmers cultivate pieces of natural land and build new houses, while others abandon their plots, leaving them to renaturalize. A crucial aspect these land-use changes have in common is that they are the result of human decisions (Parker et al., 2008). These decisions are not made in isolation, but instead influence each other. For example, people consider the availability of public services or transportation networks when choosing where to live. These facilities, however, are there because of earlier decisions on developments in the vicinity.

These examples illustrate the general observation that land use, both urban and rural, is a dynamic system. The land-use patterns one finds today are therefore essentially the result of a series of previous and incremental land-use changes that affect each other over time. These feedback mechanisms can cause developments that are initially small to grow over time and reinforce each other (Krugman, 1991; Arthur, 1999), which makes the present land-use pattern highly path dependent (Brown et al., 2005). In order to gain insights in and explain land-use patterns it is therefore essential to consider processes underlying these incremental changes.

For reasons of physical, economical, and ethical consideration we have only very limited possibilities to study land-use change processes with experiments (Janssen and Ostrom, 2006). Therefore models seem to be the appropriate tools to gain insights in these land-use changes and we argue that these models should be able to generate land-use patterns through decentralized local interactions in line with Epstein's (1999) generativist's question.

This paper presents an activity-based cellular automaton model, which can act as such a laboratory for land-use change. Section 6.2 first gives an overview of land-use modelling approaches in this direction. Section 6.3 then presents the activity-based model. Section 6.4 describes a case study application that was used for assessment of this model. Section 6.5 finally discusses the results of the case study to draw conclusions and give directions for future research.

6.2 Modelling land-use change

6.2.1 Existing types of land-use models

Over the last decades several approaches for modelling land use have been proposed. These approaches can roughly be divided in those originating from economy and those originating from geography. Economists have been using approaches, which are mostly founded on bid-rent curves as presented by Alonso (1964). These models compute an equilibrium situation, in which resulting land use or population density is depending on the distance to the urban centre, sometimes in combination with other factors (see for an overview Anas et al., 1998).

However useful these models are in the context of land pricing and urbanization, they have a few drawbacks that make them less suitable for the study of land-use change processes. Time is not treated explicitly, and therefore developments over time, which are elementary for some land-use dynamics, cannot be studied. Moreover, space is considered only as the distance to the urban centre, ignoring features that are not homogeneous over space such as elevation, transport networks or rivers. Following from the introduction, we aim to understand land-use changes as an explicitly spatial and dynamic process. Therefore Alonso-type models are not considered further.

Geographers have focused on models that simulate land-use changes in an explicitly spatial way, and hence developed model that can include spatially non-homogeneous factors, such as elevation or transport networks. Overviews of different approaches are among others available in Veldkamp and Lambin (2001) and Koomen et al. (2007). From this group of models, we would like to consider two concepts for land-use modelling in more detail: spatially explicit multi-agent systems (MAS) and cellular uutomata (CA) based models. Because both approaches include time explicitly, they allow for feedback mechanisms over time. Moreover, both methods approach human decision-making and generate land-use patterns from local dynamics (White and Engelen, 2000; Brown et al., 2008).

In this discussion, agents in MAS are actors that can act and move independently over space. The advantage of MAS is that they can represent the behaviour of agents in a very straightforward way, since agents can interact directly with each other and with the environment. Precisely these local interactions between agents and differences among them generate the patterns observed on a global scale.

However, since the agents are the basic unit of computation, MAS are computationally demanding. This is illustrated by an overview of case study applications of agent-based models for land-use modelling, presented in Parker et al. (2003). In addition, MAS require data on the level of actors, represented by agents, which makes them data demanding and poses difficulties for model calibration and validation (Robinson et al., 2007). The problem of data collection is further thwarted by privacy regulations that are related to personal data.

Cellular automata, although sometimes considered agent-based as well, differ from MAS in that sense that the basic unit for computation is a cell, not an agent. Together cells make up the lattice on which the CA exists, which makes them inherently spatial and therefore very suitable for the simulation of land-use dynamics. Since the cell is the basic unit of computation and cell sizes can be adjusted according to the scope of the application, models can keep a computational efficiency. Therefore, CA have been applied to simulate land-use changes on larger scales, from urban areas (for example Van Vliet et al., 2009) to groups of countries (for example Van Delden et al., 2010). Simulating land-use changes at a rather high level of abstraction brings the advantage that CA are less data demanding compared to MAS. Although calibration and validation of these models has been considered a major issue numerous calibrated examples are available (Hagen-Zanker and Lajoie, 2008; Van Vliet et al., 2009; Wickramasuriya et al., 2009).

The advantages of CA come at the cost of detail: individual actors are not considered. Instead, cells have a state, which generally represents the predominant land use. However, a cell can only have one land use, and land uses are thus by definition mutually exclusive and combinations of land uses are not possible on one location unless explicitly defined as a separate mixed class. Still, in reality mixed land use is the rule rather than the exception. Hence CA cannot represent the richness and diversity in land uses one observes in reality.

6.2.2 Bridging the gap between MAS and CA

Several efforts have been made to fill the gap between CA and MAS by adding quantitative information to CA land-use models. In regional applications of their constrained cellular automaton model, White and Engelen (1997) first compute for each region a density to translate the number of inhabitants or jobs in a cell demand for the associated residential or commercial land uses. Consequently, cells in their model have a density, but this density is similar in all locations with residential or commercial land use within one region. Hence, their model simulates regional differences but not local ones.

Loibl and Tötzer (2003) follow a similar approach to model migration on a regional level. However, their model adds more detail on the local level, as households and enterprises are allocated on the lattice based on local characteristics. Therefore densities can vary cell by cell and urban growth is simulated as an incremental process.

Wu (1998b) simulates urbanization by explicitly allocating residents on a lattice, which allows for incremental changes in population density. As such he is able to simulate both monocentric and polycentric urban land-use patterns, depending on the regimes used for allocation. Wu and Webster elaborate on the allocation of land development as they combine a CA approach with multi criteria evaluation and neoclassical urban economic theories (Wu and Webster, 1998; 2000). In their model land development increases incrementally, based on the profitability of the development of a location. This is again a function of the development in the neighbourhood of that location.

Yeh and Li (2002) also acknowledge that densities of urban areas differ considerably from one city to another and for different locations within one city. Their approach differs from Wu (1998b) and Wu and Webster (1998; 2000) in that land uses are considered separately from population densities. They model development density proportional to the development probability as computed from the CA transition rules. The CA model then assigns a density to those locations where the land use changes from undeveloped into urban. Depending on the parameters in the transition rules, the model can simulate different urbanization patterns, both monocentric and polycentric. Moreover the cells that are not urban can have a "grey value" between 0 and 1 that indicates how far a particular location is from urbanization.

From these examples it becomes clear that the addition of activity or information on density has been studied before in a dynamic and spatial environment, using several different approaches. Similar to Wu (1998a), Wu and Webster (1998; 2002) and Yeh an Li (2002), we see activity density as a cell property, which changes in small but incremental steps. In addition, like Yeh and Li (2002) we account for activities in a separate data layer. However, these models have in common that they focus on urban land-use change and particularly on growth. Typically, densities are changing incrementally, but changes are one-directional towards urbanization or an increase in development. Consequently, non urban land is background; it can influence the allocation of population or urban densification, but they do not change themselves as a result thereof, except when they change into urban land. In addition mixed land uses or multifunctional land uses only exist superficially as cells that are not fully urbanized. This does not reflect the richness that exists in reality, such as locations with a combination of commercial and residential uses. Hence these approaches pose problems for the simulation inner city dynamics or rural depopulation.

The model presented in this paper adds to the CA framework the notion of activities, where activities represent the general idea of a density. For example population can be the activity related to residential land use, and jobs can be the activity associated with industrial land use. Because a location can have more than one activity, mixed land uses can be represented explicitly as more than one activity in the same location, not necessarily relating to the predominant land use. As the amount of activity on a location can increase as well as decrease, the proposed model allows for the study of a range of land-use change processes, from urban growth to rural depopulation.

As the cell remains the basic unit of computation this model is an extension to the existing computational framework of CA land-use models. Because activity is considered as a cell property rather than a set of agents it is not considered to be a MAS.

6.3 The activity-based model

6.3.1 The Metronamica land-use modelling framework

Cellular automata (CA) land-use models generally exist on a lattice of regular squares. Each cell on the lattice has one of a limited number of cell states which represents the predominant land use on that location. In each discrete time step, cell states are updated simultaneously according to a set of transition rules. The characteristic of CA models is that the state of adjacent cells, and hence the land use in the neighbourhood of a location, is input to the transition rules. Additional factors are often added to represent heterogeneous geographic features.

The activity-based approach is founded on the CA model as presented by White et al. (1997) and further developed as the Metronamica land-use modelling framework (Van Delden and Hurkens, 2011). Metronamica uses three land-use types: constrained land uses, which are actively allocated by the CA, features, which are not supposed to change during a simulation, and unconstrained land uses, which only change as a results of other changes. "Constrained" refers to the notion that the total number of cells per land-use type is determined exogenously (White et al., 1997). In land-use terms, the total area demand for a constrained land use in a certain time step is defined externally, while this demand is then allocated by the CA. Land-use allocation is driven by the potential of cells, an endogenous variable that is calculated for every location and each constrained land-use class:

$$P_{g,i}^{t+1} = \nu \cdot A_{g,i}^{t+1} \cdot Z_{g,i}^{t+1} \cdot S_{g,i}^{t+1} \cdot N_{g,i}^{t+1} \qquad \text{Equation 6.1}$$

where $P_{g,i}$ is the transition potential for land use g on cell i, $N_{g,i}$ is the neighbourhood effect for land use g on cell i, $S_{g,i}$ is the physical suitability for land use g on cell i, $Z_{g,i}$ is the zoning status for land use g on cell i, $A_{g,i}$ is the accessibility for land use g on cell i, and ν is a stochastic perturbation term. The latter is added to represent the different preferences that individual actors have and to account for variation in factors that are not otherwise represented. Time dependent variables are indicated with superscripts, where t indicates that information is taken from the existing situation, and t+1 indicates that this information is used for activity allocation in the next time step. Factors that represent physical suitability, zoning and accessibility can have values between 0 and 1.

The neighbourhood effect is the dynamic component of the CA algorithm which accounts for the self-organizing behaviour of the model. It is calculated from the land uses in the neighbourhood of a location:

$$N_{g,i}^{t+1} = \sum_j w_{d(i,j),g,l^t(j)} \qquad \text{Equation 6.2}$$

where $w_{d,g,l}$ is the neighbourhood rule that describes the effect of land use l at distance $d(i,j)$ on the potential of location i for land use g . This land use taken from the previous time step. Locations that do not have a feature of constrained land use will get the unconstrained land use for which they have the highest suitability.

6.3.2 Adding activities to the Metronamica framework

In the original Metronamica modelling framework, land uses are constrained by an area demand. However, in reality many land uses are not constrained by an area, but by an amount of activity instead. In the case of urban expansion it is not an amount of land surface that needs to be covered by urban land, but the population that needs a place to live. Therefore, in the activity-based approached, the total demand is not constrained in terms of cells, but in terms of activity instead. Hence for each time step the amount of activity that needs to be allocated on the lattice is defined exogenously. After the activity allocation, land uses that are associated with activities are assigned based upon the new activity distribution. The number of cells per land use then depends on the density of the activity: a high density requires fewer cells than a lower density.

Additionally, the land uses in the surrounding of a location are input to the neighbourhood effect in the original constrained CA. For example, commercial areas in the vicinity make a location more attractive for residential development as they represent jobs and services. As quantitative information is available in the activity-based model, it is now possible to define the attraction as a function of the number of jobs rather than just the presence of commercial land use. Hence locations that have a high job density in their commercial areas are more attractive to residents than locations with only a low density. At the same time, the compatibility of existing land use will influence the amount of activity that can be allocated on a location.

Since not all land uses can be associated with an activity, we have three types of land uses in our model. These are activity-constrained land uses, area-constrained land uses and unconstrained land uses, where the activity-constrained land uses are added to the existing Metronamica framework. Activity-constrained land uses are assigned based upon their activity distribution, area-constrained land uses are allocated after that, and finally, all cells that are not assigned one of the constrained land uses get the unconstrained land uses. The order of allocation of land-use types represents the economic influence or power that is associated with these land uses.

An obvious example of an activity-constrained land use is residential land use, with population as the associated activity. It should be noted that each activity constrained land use has only one associated activity. An example of an area-constrained land use is agriculture, for which the demand is expressed as a number of cells. Unconstrained land uses are often natural land uses, which occupy all locations that are not occupied by land uses which are driven by an external demand. Similar to the original Metronamica framework, the model requires at least one unconstrained land-use class to make sure that all cells will have a land use at any time.

To keep track of activities, an additional data layer is required per type of activity. Hence, a cell has no longer only one discrete cell state. Instead it has a land-use state, and one numerical value for each activity. Computation of land-use dynamics therefore becomes a two-step process, as presented in figure 6.1. In each time step, first the demand for activities (a) is distributed according to the potential of each location for that activity (b). This potential for each activity is a function of the land use and activity distribution in the previous time step (c and d), as well as the suitability, zoning status and accessibility of that location (e). Activity-constrained land uses are then assigned based on the updated activity distribution (f). Similarly, the total potential for area constrained land uses is computed based on the land use and activity distribution from the previous time step (g and h) and the suitability, zoning status and accessibility of that location (i). These land uses are allocated accordingly (j) until the area demand (k) is fulfilled.

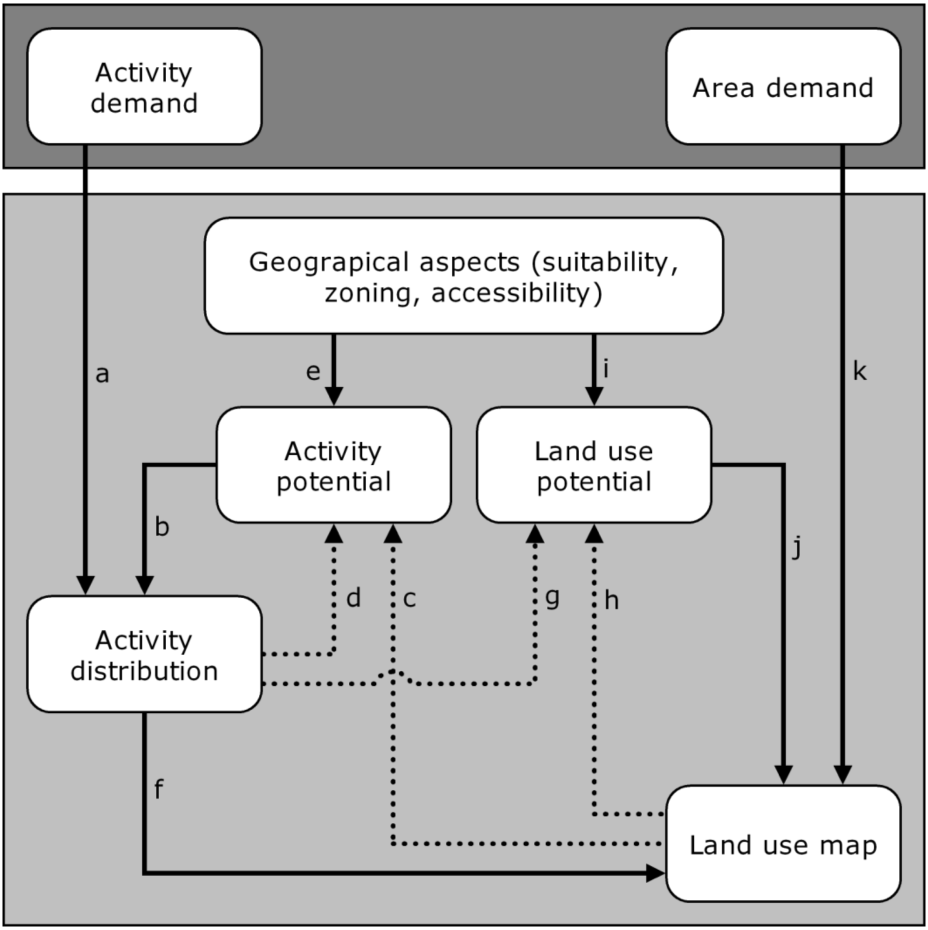


Figure 6.1: *System diagram for land use and activity distribution. Arrows show the flow of information, where solid lines represent current values and dashed lines represent values from the previous time step. Other symbols are explained in the text.*

An important characteristic of the proposed activity based model is that it is a generic model that aims to simulate urbanization from the bottom up. According to urban economic theory, urbanization is the result of the interplay between centripetal and centrifugal forces (Krugman 1996; Furtado et al. 2012). The activity-based model incorporates both forces. Agglomeration effects are simulated by the mutual attraction of activities creating economies of scale. However, this agglomeration advantage will generate competition for space which leads to higher land prices, congestion, and pollution that are associated with higher activity densities. These diseconomies of scale cause centrifugal forces. Both forces are included in the proposed activity based model for which details are provided in the next section.

6.3.3 Activity distribution and land-use allocation

Activity is distributed proportionally to a cell's potential for this activity, which is computed according to equation 6.3:

$$P_{k,i}^{t+1} = c_{k,i}(g^t) \cdot f\left(N_{k,i}^{t+1} + E_{k,i}^{t+1} + \varepsilon\right) \cdot A_{k,i}^{t+1} \cdot Z_{k,i}^{t+1} \cdot S_{k,i}^{t+1}$$

Equation 6.3

Where $P_{k,i}$ is the potential for activity k in cell i, $c_{k,i}(g)$ is the compatibility coefficient that indicates how well activity k can be accommodated by existing land use g in cell i, $f(x)$ is a transformation function to avoid negative potentials, $f(x) = log_2(1+2^x)$, $N_{k,i}$ is the neighbourhood effect for activity k in cell i, $E_{k,i}$ is the diseconomies of scale effect for activity k in cell i, ε is a scalable stochastic variable drawn from a normal (0, α) distribution, $S_{k,i}$ is a factor that represents the physical suitability for activity k on location i, $Z_{k,i}$ is a factor that represents the zoning status for activity k on location i, and $A_{k,i}$ is a factor that represents the accessibility to transport networks for activity k on location i. Time dependent variables are indicated with superscripts, where t indicates that information is taken from the existing situation, and t+1 indicates that this information is used for activity allocation in the next time step. The factors that represent physical suitability, zoning and accessibility can have values between 0 and 1. The stochastic variable is drawn independently for each cell and in each time step, and its effect is scalable with parameter α.

The neighbourhood effect $N_{k,i}$ is a function of the existing activity distribution in the neighbourhood of a cell. It is computed as the sum of the effects of all cells j on all distances d in the neighbourhood. This includes the activity and the land use of cell i itself.

$$N_{k,i}^{t+1} = \sum_j w_{d(i,j),k,h(j)} \cdot X_{h(j)}^t$$ Equation 6.4

where $w_{d,k,l}$ is the weight function representing the attraction or repulsion of activity h on activity k at distance d, and $X_{h(j)}^{t}$ is the amount of activity l in cell j. The activity of area constrained land uses and unconstrained land uses follow the Kronecker delta function: they are 1 wherever that land use is present, and 0 in all other cells.

Diseconomies of scale represent the negative effects of agglomeration. This effect is computed as a function of the neighbourhood effect for that activity in a specific location:

$$E_{k,i}^{t+1} = \gamma_1 \cdot (N_{k,i}^{t+1})^{\gamma_2} \qquad \text{Equation 6.5}$$

where γ_1 and γ_2 are parameters that need to be calibrated. Typically $-1 < \gamma_1 < 0$ and $\gamma_2 > 1$ to make sure that negative externalities are small initially and grow more than proportionally with the present activity.

Once the activity potentials are calculated for all cells j, the amount of activity k that is assigned to a particular cell i, $X_{k,i}$, is proportional to the activity potential of that cell:

$$X_{k,i}^{t+1} = \frac{P_{k,i}^{t+1}}{\sum_j P_{k,j}^{t+1}} \qquad \text{Equation 6.6}$$

Once the amount in each activity on each cell is known, the associated land uses are assigned to those cells that have an activity higher than a predefined threshold value. When more than one activity exceeds this threshold value, the cell is assigned the land use for which the activity is highest relative to the threshold value. The threshold values can differ for the different types of activity. Regardless of the assigned land use, all cells maintain their activities.

To allocate area-constrained land uses, the potential for these land uses is computed for each cell and for each area-constrained land use. Note that similarly to the potential for activities, this potential is computed based upon the activity distribution from the previous time step. Since area-constrained land uses are not represented with a quantity, there are also no diseconomies of scale.

$$P_{g,i}^{t+1} = f\left(N_{g,i}^{t+1} + \varepsilon\right) \cdot A_{k,i}^{t+1} \cdot Z_{k,i}^{t+1} \cdot S_{k,i}^{t+1} \qquad \text{Equation 6.7}$$

Where $P_{g,i}$ is the total potential for land use g in cell i, and $N_{g,i}$ is the neighbourhood effect for land use g in cell i. The neighbourhood effect for area-constrained land uses is computed exactly similarly to the neighbourhood effect for activities. This differs from the original Metronamica in that the neighbourhood effect is a function of the activity distribution rather than the land-use pattern.

Area-constrained land uses are assigned to cells with the highest potential until the demand for this land use is met and as long as these cells are not already occupied by an activity constrained land-use state.

Finally, all cells that are not occupied by either an activity-constrained land use or an area-constrained land use receive an unconstrained land use.

6.4 A synthetic case study

6.4.1 The case study aplication

A synthetic application was used to test the hypothesis that settlement patterns can grow from local interactions between activities and land uses. Hence, the case study does not include geographical information such as physical suitability, spatial planning or accessibility to transport networks. The case study is defined on a regular lattice of 200 by 200 cells, which for illustrative purposes can be taken to represent one hectare each. Thus the total area comprises 40 000 cells or 400 km^2. The application is used to simulate land-use change over a period of 1000 time steps that represent one year each.

The case-study application has two activity-constrained land-use types. These are residential land and industrial land, and associated to this are the activities population and jobs, respectively. An initial small amount of population and jobs was distributed randomly over the area. Over time population increases linearly from 592 at t = 0 to 60 000 at t = 1000, while the number of jobs increase linearly from 445 at t = 0 to 30 000 in t = 1000. Next to residential and industrial land use, there is one area-constrained land use, which is agricultural land. We assume that 1 hectare is sufficient to feed 10 persons, therefore the area demand for agriculture increases from 1 cell in t = 0 to 6 000 cells in t = 1000. Finally, there is one unconstrained land use, which is natural land. All cells that are not occupied by residential, industrial or agricultural land use become natural land.

To evaluate the activity-based CA model we considered the land-use patterns that were generated in T = 1000, which were assessed visually and by the statistical signature (Moss, 2002). Additionally, the meaning of the parameters in the transition rules as well as the path towards the eventual land-use pattern was evaluated.

6.4.2 Parameterization

The calibration of parameters in the activity-based model is a manual process. However, the parameter space is limited as we require neighbourhood rules to represent real-world interactions between actors. As argued before, these interactions are typically local. Because the case-study application includes three constrained land uses and one unconstrained land use, there are twelve possible interactions between activity types. The numerical values of the compatibility coefficients and the weights in the neighbourhood effect that were used for this case study are given in Table 6.1 and Table 6.2, respectively. It should be noted that weights that represent the effect of population and jobs are low relative to the effect of other activities. This is because the weights are multiplied with the respective activities on a location. The density for population and jobs is typically much higher than the activity density of 1 associated with agricultural and natural land uses.

Population, associated with residential land use, can be allocated on all land-use types, but has a natural preference for residential land uses, represented in its compatibility factor. Population is attracted by existing population in the same location as well as in the direct vicinity. This represents the social interactions or the availability of general facilities implied by the clustering of a larger number of people. Jobs and agriculture attract population at a small distance, representing employment and the availability of food in the neighbourhood.

Jobs show behaviour similar to population. They can be allocated on all land uses, although natural land use is by far the least attractive, as represented by the relatively low compatibility coefficient. Both agriculture and jobs attract jobs in their vicinity, since both represent employment. Clustering creates additional employment, for the processing of products, and benefits of scale such as described among others by Arthur (1990). Jobs are attracted by population, because people represent both customers and employees.

Table 6.2: *Compatibility coefficients as used in the case study application.*

Land use	Compatibility with population	Compatibility with jobs
Natural land use	0.6	0.5
Residential land use	1	0.7
Industrial land use	0.4	1
Agricultural land use	0.7	0.7

Table 6.1: *Weights of the neighbourhood effect as used in the case-study application. Weights are interpolated linearly between the indicated values.*

	Distance (cells)				
Activity interactions	0	1	2	3	4
From population to population	30	0.25	0.001	0	0
From population to jobs	0.1	0.4	0	0	0
From population to agriculture	0	3	0.5	0.25	0
From jobs to population	0	0.5	0	0	0
From jobs to jobs	20	0.45	0	0	0
From jobs to agriculture	0	2	0	0	0
From agriculture to population	4	1.5	0.2	0.1	0
From agriculture to jobs	0	2	0	0	0
From agriculture to agriculture	300	5	0	0	0
From nature to population	0	0	0	0	0
From nature to jobs	0	0	0	0	0
From nature to agriculture	0	0	0	0	0

Agricultural land use is attracted by areas where there is population, because population represents first of all farmers to work on the field and second, customers for their products and therefore a lower distance to markets. The latter also explains why an accumulation of population is more attractive for agricultural land use than a mere presence of it. A similar relation but much less strong exists between jobs and agriculture, because of the benefits of the scale effect as well as jobs indicating a processing industry. Natural land use is not of any influence on agriculture since it does not represent any apparent attraction or repulsion.

Natural land use itself is not allocated according to transition rules, since it only occupies locations that are not in use by residential, industrial or agricultural land. Therefore, there is no interaction effect from other land uses or activities on natural land use defined.

6.4.3 Simulation results

Figure 6.2 shows snapshots of the land use, the population distribution and the job distribution, at regular intervals in time. Maps are taken from one single run, but because all simulation runs show similar results, this is taken as a representative example. The maps show that initially activities are distributed more or less randomly over space, as are the plots of agricultural land. Activity is not yet clustered to the extent that any residential or industrial land appears. As time progresses and the amount of activity increases, activity clusters and so does the agricultural land. The first small patches of residential land appear in the centre of larger agricultural areas, and continue to grow because the population increases. Eventually, some urban clusters grow bigger over time, while most of them remain of smaller sizes. As jobs cluster on locations with a concentration of population, some locations with industrial land appear on the edge of larger urban clusters, while others are more isolated.

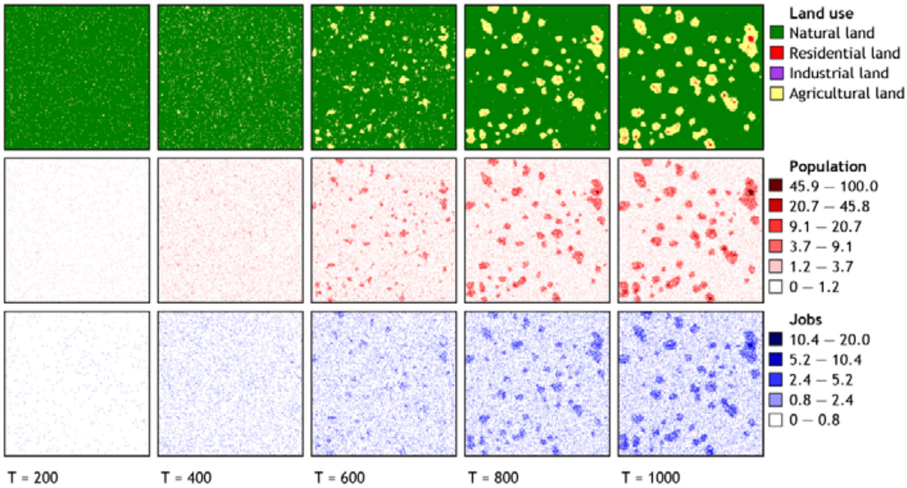


Figure 6.2: *Time series representing land use and activity distributions of a typical simulation run for regular intervals in time.*

The emergence of the settlement pattern over time occurs in several stages, which are associated with an increa--gly more developed economy. Initially people cluster together only a little bit in agricultural areas, representing the development of the primary sector. Then, as the population grows, the first settlements appear, which equates to the secondary sector as some form of organization is required. Most settlements stay small, while some grow over time to more central cities. It is mostly around these larger cities that also jobs clusters to the extent that industrial land use appears, indicating a developing tertiary sector.

6.4.4 Urban cluster distribution

The described settlement pattern emerges from strictly local interactions between activities. These incremental activity changes eventually exhibit themselves as land-use changes, in accordance with the initial hypothesis.

However, as a description is only subjective, the generated land-use patterns and activity distribution are also assessed objectively.

Models for land-use change are often evaluated by their capability to accurately simulate historical land-use changes (Pontius et al., 2008), where accuracy is typically assessed on a pixel level using map comparison techniques (Hagen, 2002; Pontius et al., 2004a). Since the aim of this study is to test the ability to generate realistic urbanization patterns rather than to simulate changes accurately, we use a synthetic application and therefore we have no historic land-use changes to compare with.

For reasonably large areas the distribution of urban cluster sizes is known to follow Zipf's law, also known as the rank – size rule (Krugman, 1996; Gabaix, 1999; Reed 2002; Cordoba, 2008). For this, clusters of direct adjacent residential and industrial land are ranked from one onwards, where 1 represents the largest cluster, 2 the next largest, and so on. Cluster sizes are measured from the population in a cluster in the simulation results at T = 1000. The rank-size rule indicates that this distribution approximates a power law as follows:

$$Size = a \cdot rank^b \qquad \text{Equation 6.8}$$

Figure 6.3 shows this rank-size distribution for one simulation result, including the power law that approximates this distribution best.

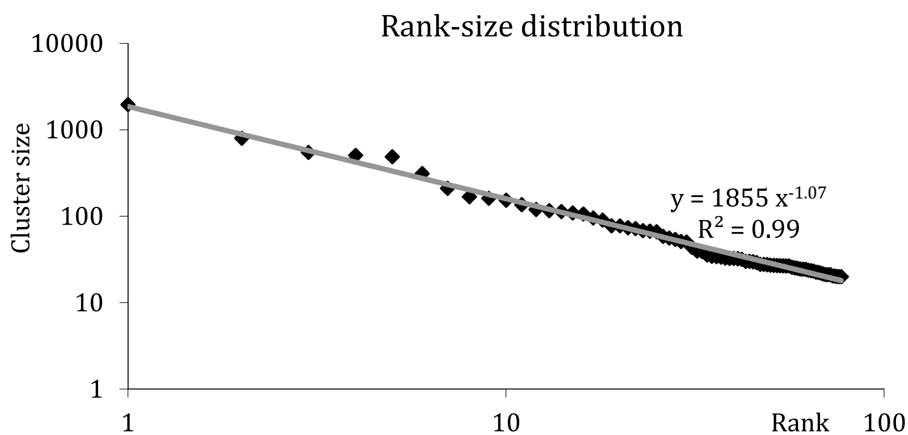


Figure 6.3: *Cluster size distribution for the result of one typical simulation run at time T = 1000.*

As described in section 6.3, to generate realistic patterns the model includes some noise, represented with a normally distributed random variable. Therefore it is not sufficient to assess only one simulation result as this could be an outlier. Hence the model was run several times starting with a different random seed and results were assessed. Results for all ten simulation results are presented in Table 6.3.

R^2 values that are close to 1 indicate that the land-use patterns that are generated closely follow the expected rank-size rule. Moreover, the slope of the power-law function that describes the cluster-size distribution is close to the value of -1 as was expected from literature (Gabaix 1999; Cordoba 2008). This shows that the activity based CA can generate realistic land-use settlement patterns.

The exact shape of this power law is a function of two forces, the neighbourhood effect works as a centripetal force while the diseconomies of scale and the stochastic perturbation work as centrifugal forces. Adjustments to the size of both forces adjust the rank-size distribution as well, with a tendency to several large settlements or more numerous small ones, depending on the exact parameterization. In the extreme case, without centrifugal forces, the model generates the von Thünen solution: one large residential area surrounded by agricultural lands in the middle of a natural area.

Table 6.3: *Generated land-use patterns, expressed as estimated parameters for the power law that best describes this distribution. See text for further explanation.*

	a	b	R^2
Simulation run 1	1937	-1.02	0.97
Simulation run 2	2422	-1.12	0.97
Simulation run 3	1614	-0.99	0.98
Simulation run 4	2204	-1.10	0.97
Simulation run 5	1855	-1.07	0.99
Simulation run 6	1642	-1.01	0.97
Simulation run 7	1916	-1.03	0.99
Simulation run 8	2071	-1.14	0.97
Simulation run 9	2048	-1.18	0.96
Simulation run 10	2286	-1.16	0.97

6.5 Conclusions and directions for further research

Using activities to simulate land-use dynamics combines aspects of cellular automata (CA) land-use models and multi-agent systems (MAS): The cell is still the basic unit of computation, but this is complemented with information and behaviour of actors, represented by activities. This approach adds to the constrained CA framework in three ways. First, the presence of activity gives more information than the, rather superficial, land-use states only: the model explicitly allows mixed land uses as well as a variation in densities for activities per cell, from sparse residential areas to densely populated locations. Second, land-use dynamics are the results of many smaller and incremental changes rather than sudden changes in cell states only. Third, the activity-based model explicitly allows including the feedback between activities and hence better represents the process of land-use change.

Assessment of the behaviour and results shows that the activity-based model generates realistic land-use patterns. As the inclusion of activities allows representing land-use change processes more accurately we expect that this approach will improve the results of real world applications for land-use modelling. This is particularly the case for processes such as rural depopulation or urban sprawl, which are the result of incremental changes in activity location rather than sudden changes in the land use.

Therefore, the obvious next step is to test the activity-based CA model for its ability to mimic historic changes in land use and activities accurately. For this the activity-based CA model should include real world aspects, represented by physical suitability, zoning maps and transport networks. A first attempt shows promising results in that direction (Van Vliet and Van Delden 2008), but requires more extensive testing on aspects such as robustness and scale sensitivity such as already done for other CA models (Menard and Marceau 2005; Kocabas and Dragicevic 2006).

An advantage of the activity-based approach over agent-based models is that the data demand is relatively low. To calibrate an application for historical changes, land-use data as well as data for activities like population and jobs are required for at least two points in time. For independent validation, another dataset for a third point in time is required. Due to developments in remote sensing, land-use data is currently easily available. For population data or data for other activities, this is not yet the case. However, several approaches have been developed to estimate population density on the level of cells from satellite imagery (Harvey 2002; Wu and Murray 2007) or from land-use maps (Gallego and Peedell 2001). We expect that these developments will make the required data more widely available in the nearby future.

Additionally, the activity-based model offers several opportunities for integrated land-use modelling. As an example Luck (2007) gives an overview of the relation between population density and its effect on biodiversity. He concludes that biodiversity and population density are significantly correlated and points at the need to focus on anthropogenic drivers of environmental changes. This possibility appears in models that explicitly link socio-economic and biophysical aspects of the land stem, such as the MedAction PSS (Van Delden et al., 2007).

7. Synthesis

7.1 Main findings

The land use that we observe today is constantly changing. These changes are driven by a multitude of driving forces, such as a growing population, shifts in lifestyle, an increase in food demand, climate changes, or changes in other biophysical conditions. Moreover, land-use changes have many and far reaching consequences: they influence traffic patterns and congestion, the production of food, the sequestration or removal of carbon, and biodiversity. Therefore, it is important to understand land-use changes to anticipate future changes under alternative conditions and to assess the impact of policy measures.

In the last two decades, land-use models have been increasingly applied as tools to study land-use changes (Claessens et al., 2009; Wyman and Stein, 2010), to explore future land-use scenarios (Yang and Lo, 2003 and Verburg et al., 2006) and assess the impact of policy measures (Jantz et al., 2004; Sieber et al., 2010; Van Delden et al., 2011b; Hellman and Verburg, 2010). These developments witnessed in land-use change models is at least partly fuelled by an increase in the availability of spatial data, and an increase in computation power, which together facilitate the application of high resolution spatial models. These land-use models, in order to apply them to a real world case, require calibration to adjust parameter values, and validation to assess the model results after calibration (Chen and Pontius, 2010). However, developments in land-use modelling have not been matched by developments in model assessment (Gardner and Urban, 2005). This is partly due to a lack of interest in model testing and partly to a lack of proper methods for model testing and the lack of a proper framework for this purpose (Agarwal et al., 2000; Refsgaard and Henriksen, 2004; Dietzel and Clarke, 2007).

This thesis contributes to the calibration and validation of land use models in general and of cellular automata (CA) based land-use models in particular. In order to do this the following four research questions were addressed:

1. What characteristics of land-use models are important for assessing these models?
2. How can the predictive accuracy of a land-use model be assessed?
3. How can the process accuracy of a land-use model be assessed?
4. How can the neighbourhood rules in cellular automata land-use models be calibrated and validated?

This chapter reflects on these four research questions. The rest of the paper is structured as follows: section 2 indicates the main findings of the research presented in the context of the stated research questions. Section 3 provides some critical reflection on these main findings, and section 4 subsequently points at some directions for future research.

7.1.1 What characteristics of land-use models are important for their assessment?

Many land-use changes are inherently uncertain, because the drivers causing these changes cannot be identified sufficiently accurate. This is particularly true for human decisions that are underlying many land-use changes. For example, when facing the same set of biophysical and economic conditions, one farmer might decide to quit farming for financial reasons while another might decide to continue because she or he feels it is his obligation to keep the family farm up and running. Similarly, the preference of one or a small number of politicians might determine the location of a new urban expansion area. This uncertainty is represented in many land-use models, including Metronamica, by some stochastic perturbation or random component in the simulation of land-use change processes. For these reasons, the exact location of future land-use changes cannot be known and therefore land-use change models cannot be expected to yield a perfect fit with reality (Clarke, 2004; Batty and Torrens, 2005; Manson, 2007), even though it might be able to explain some land-use changes. Therefore, modellers have indicated that a comprehensive assessment of land-use models should not only assess whether land-use changes are simulated exactly, but also whether the land-use change processes are simulated realistically (Brown et al., 2005; Hagen-Zanker and Martens, 2008).

Land-use change models typically start from an initial land-use map and make changes to this map, for example by allocating new urban areas, based on the existing urban areas, or by simulating land abandonment from existing agricultural land. During a typical land-use simulation - one to several decades - there is only a fraction of the land that actually changes while the majority of the land persists in its original state. This implies that assessing land-use models only on the basis of the generated land-use map is not appropriate because the amount of change and persistence is not included in the assessment: in a case where land-use changes only little over the simulation period, most simulation results will yield a high accuracy, even in the case where all changes are simulated incorrectly. A proper assessment therefore requires assessment of the simulated land-use *changes*, rather than the simulated land-use *pattern*. The main implication of this difference is that assessment methods, both for the predictive accuracy and the process accuracy, require an appropriate reference level in the sense that they needs to account for the information that is already available in the initial land-use map (Hagen-Zanker and Lajoie, 2008).

It is not always possible to test results of land-use change models against real world observations, because real world observations are not available for all experiments. Examples include simulation of land dynamics beyond the period of data availability or simulation of land-use changes in a synthetic case study. In both cases it is not possible to assess the predictive accuracy as there is no observed truth to compare with. However, it is possible to some extent to assess the process accuracy simulations by assessing the generated land-use patterns, as illustrated in chapter 4 for long term simulations and in chapter 6 for synthetic models. The latter qualifies as generative social sciences, which requires growing a phenomenon in order to explain it (Epstein, 1999). Such measures, which can include objective and measurable as well as more subjective or visual methods, allow assessing general model behaviour.

7.1.2 Assessment of the predictive accuracy of land use models

The predictive accuracy of a land-use model indicates how accurately land-use changes are allocated by this model. This is typically assessed by a pixel by pixel comparison of the simulated land-use map at the end of a simulation period with the actual land-use map at the end of the simulatin period. The commonly used method for this comparison is the Kappa statistic, which expresses the agreement between two categorical datasets corrected for the expected agreement. This expected agreement is based on a stochastic model of random allocation given the distribution of class sizes. However, as explained previously, this is not a meaningful reference level because the amount of change and persistence is not included in the assessment.

Chapter 2 of this thesis presents Kappa Simulation, a statistic that is similar in form to the Kappa statistic, but instead applies a more appropriate stochastic model of random allocation of class transitions relative to the initial map. This is a more relevant reference for land-use models, as it includes the amount of change or persistence in the simulation as well as in reality. As a consequence, Kappa Simulation truly tests models in their capacity to explain land-use changes over time, and unlike Kappa it does not inflate results for simulations where little change takes place. An application of both Kappa and Kappa Simulation on several case study results essential illustrates this difference as both methods rank results differently: while Kappa indicates that a no-change model is more accurate than the case study land use model, Kappa Simulation shows that a no-change model does not explain any land-use changes, while the land-use model does.

Most map comparison methods are crisp in their interpretation of land-use changes and their location. As a consequence, these methods do not differentiate between near-hits and complete misses in the assessment of land-use change models. However, from a modeller's point of view near-hits are a much better result than complete misses. Fuzzy Kappa, therefore, is a map comparison method that does account for near-hits in both location and class definition. Chapter 3 of this thesis presents Fuzzy Kappa Simulation, a method that combines properties from Fuzzy Kappa and Kappa Simulation. It assesses the similarity between two land-use maps corrected for the agreement expected from a stochastic model of random allocation of class transitions relative to the initial map, and it applies a fuzzy interpretation of land-use changes and the location of land-use changes. Application of this method to case study results shows that, similar to Kappa Simulation, Fuzzy Kappa Simulation truly tests models for their ability to simulate land use changes over time. In addition, case study results also show that Fuzzy Kappa Simulation is able to distinguish near-hits from complete misses, as it could differentiate between two simulation results that yielded similar scores when a crisp interpretation was applied.

7.1.3 Assessment of the process accuracy of land use models

The process accuracy of a land-use model expresses how realistic land-use change processes are represented in the model. As discussed in section 7.1.1, this is an essential addition to the predictive accuracy since the exact allocation of land-use changes cannot be known and therefore land-use change models cannot be expected to yield a perfect fit with reality. However, modelled land-use change processes are difficult to assess directly: land-use change is typically influenced by a combination of drivers, and therefore the effect of individual drivers cannot be measured in isolation, if it can be measured at all. Therefore, process accuracy is typically assessed indirectly from the land-use patterns that are generated by the model.

The assessment of the process accuracy of a land-use model requires characterisation the land-use pattern. One way to do this is by means of landscape metrics, which include indices that express the diversity of the land use pattern, or that characterize the shape of patches (Riiters et al., 1995; McGarigal et al., 2009). Landscape metrics can be used to compare the simulated land-use pattern with the actual land-use pattern at the same moment, or to compare the simulated changes in the land-use pattern with the observed changes in the land-use pattern. These comparisons assess if the end state and the direction of change in the simulation is similar to that of the observed land-use changes. This approach is used to assess the results of land-use model that was applied in chapter 2. Similar to the assessment of the predictive accuracy, such comparison requires a reference level for a meaningful interpretation. In chapter 4, a neutral land-use model, the random constraint match model (Hagen-Zanker and Lajoie, 2008), was used as a reference level.

Another way to characterise land-use patterns is by means of fractal measures. Fractal measures originate from complexity theory and can describe regularities in land-use patterns, such as the distribution of urban cluster sizes, which is known to describe a power-law distribution, or the relation between the area and perimeter of urban patches, the fractal dimension. Although both measures can be used to assess some urban land uses only, they have the advantage that values of these measures have been found to be quite constant. Therefore, they can be used to assess synthetic applications for which no observed reality exists, as is demonstrated in chapter 6. In addition, they can be used similarly to landscape metrics: by comparing the actual land-use pattern with the simulated land-use pattern, as well as with a land-use pattern generated by some reference model, as was shown in chapter 5.

7.1.4 Calibration and validation of neighbourhood rules in cellular automata based land-use models

The neighbourhood effect represents the influence of neighbouring land uses and their associated agents on the allocation of land-use changes. Therefore, the neighbourhood effect is especially apparent in those land uses that change as a consequence of human decisions, such as urban land uses. An example is the attractive effect that a commercial area invokes on new residential houses. Because residents generally prefer a certain level of services within their vicinity, and commercial enterprises require sufficient customers and employees, the two are mutually attracted. The neighbourhood effect is represented in land-use models in their neighbourhood rules, the defining element of cellular automata (CA) based land-use models. Parameter values for these neighbourhood rules need to be set in a calibration procedure. This thesis addresses three aspects of the calibration of neighbourhood rules.

The distance over which land uses still influence each other remains an issue of debate. Initially, the size of the neighbourhood effect was constrained by the computational power available. As this computation power increased, neighbourhood sizes increased as well. Chapter 4 tests the variable grid CA model, a CA model that aggregates land uses at greater distances from a location so that the entire map can be included in the neighbourhood rules. This allows all land to be included in the neighbourhood. Results for a small scale application for Vancouver indicate that the model results improve by introducing neighbourhood rules at two different scales: first, the directly adjacent cells in the model representing the scale of a city block influenced the allocation of land-use changes, and second, introducing neighbourhood rules at the scale of a city region also improved the simulation results. This reflects the idea that agents consider different hierarchical scales in their allocation decisions.

Chapter 5 dives deeper into the calibration of parameters in the neighbourhood rules as it employs enrichment factors to characterize the neighbourhood effect. Enrichment factors of observed land-use changes confirm the hypothesis that land-use changes are related to the existing land-use pattern, especially for urban land-use changes. However, as the location of land uses is correlated with other drivers for land-use changes, such as the accessibility to transport networks or the suitability of the location, these measurements cannot be used as model parameters directly. Therefore, chapter 5 further applies these enrichment factors to compare observed land-use changes with simulated land-use changes from two different calibration procedures: a manual procedure and an automated procedure. This comparison shows that both calibration procedures improve the predictive accuracy and the process accuracy of the land-use model. Simulation results also show that neighbourhood rules mainly improve the process accuracy of land-use models.

Neighbourhood rules in CA models are typically based on the predominant land use in a cell. However, since they actually represent the interaction between agents associated with this particular land use, this effectively ignores density effects as the number of agents associated with the predominant land use is not represented in the model and as a consequence their quantity will not affect land-use changes. Moreover, activities other than those associated with the predominant land use in a location are not represented in these models, essentially ignoring mixed land uses. Chapter 6 presents an activity based CA model. The neighbourhood rules in this model are a function of the amount of activity in a cell, rather than the predominant land use only, where activity denotes a quantity or density related to a land use. This study confirms that the local interaction between spatial actors represented as an activity level (in this case residents and jobs) can generate a realistic urban settlement structure.

7.2 Reflection on the main findings

For land-use modelling to take a prominent position in land-use science, accepted procedures for model testing are required. In addition, land-use models are increasingly being used as tools to support policy making, either by assessing specific policies or policy options or by scenario studies that inform about possible future land developments. However, in order to study land-use change processes or to inform policy makers, the results of a land-use change model should be of sufficient quality. Procedures for testing land-use models are currently not well established and as a consequence, many models are poorly tested (Silva and Clarke, 2002; Agarwal et al., 2000; Pontius et al., 2004a). This dissertation builds on existing frameworks for simulation models by adjusting these to fit the specifics for land-use change modelling in the introductory chapter and by developing and applying methods that can be used within this framework in the subsequent chapters. While neither the framework nor the presented methods are panaceas to model testing, they do provide a next step to develop land-use modelling as a science.

This thesis presents and applies several methods to assess the results of land-use models. However, the question that remains unanswered is what results would indicate that a model is good enough. Many authors have argued that a perfect fit is not achievable because there is a limit to the ability with which we can predict the exact location of land-use changes (Clarke 2004; Manson, 2007). This is due to the uncertainty that is inherent in land-use changes, particularly those that are driven by human decisions. This uncertainty is reflected in the fact that many land-use models include a stochastic term, causing the land-use system to show properties of a complex system, including associated non-linearity. Therefore, the uncertainty in the predictive accuracy will increase over time, which places a limit to the period that can be simulated reasonably. A cross section of currently available land-use outlooks indicates that one to several decades is the range for which researchers feel comfortable simulating land-use changes (Schulp et al., 2008; Hellman and Verburg 2010; Haasse et al., 2010; Van Delden et al., 2011b). Moreover, the representation of reality in a model is a simplification, which implies that some drivers that also influence land-use changes are excluded from the model. An additional complication is the feedback that exists in land-use changes: deviations in a simulation that are small initially can grow over time and have large effects on the generated land-use pattern.

This thesis uses benchmarks to validate results of land-use change models. These benchmarks can be applied in the assessment of the predictive accuracy and the process accuracy and serve as an indication of the minimum accuracy that a model requires in order to pass the validation. Some benchmarks are implicit to the assessment methods, such as Kappa Simulation and Fuzzy Kappa Simulation. In thaese cases a minimum score of 0 is required. In other cases the benchmark is explicit, taking form of a reference model, such as the random constraint match model or a null calibration, as applied in this thesis. There, the land-use model has to generate better results than these benchmark models in an accuracy assessment. Both the implicit and explicit benchmarks are low thresholds to pass: they serve as a minimum condition to pass the validation, but they are not necessarily a sufficient score that proof a good model. The challenge is to find an appropriate benchmark to avoid a false sense of accuracy. For land-use modelling specifically, this means that a benchmark should acknowledge that simulations start from a land-use map and that the end result is for a large part the result of persistence. Therefore, benchmarks based on random allocation of land uses are inappropriate as they will inflate the accuracy of the model, while benchmarks based on a random allocation of changes are more appropriate as they do not show this inflation.

Any calibration and validation of a land use model remains case specific. This poses some limitation on the application of such models to explore future land-use changes, or changes in other regions. Passing the validation means that under the given circumstances, the model simulates changes sufficiently accurate. This poses limitations on the period for which extrapolations can be made, as uncertainty increases over time. It also means that a model cannot, and cannot be expected, to predict "black swan events" (Makridakis and Taleb, 2009) or "unknown unknowns" (Pawson et al., 2011), such as changes in political regimes or the introduction of new technologies. These events are not only unknown but also unknowable and therefore fall outside the scope of land-use modelling. At best they can be explored in terms of what-if scenarios.

Land-use modelling typically involves other stakeholders in the process, either as users of the model, or as users of the model results. In this context, an accurate model is not necessarily sufficient to make a model acceptable for application. Instead, a model needs to be credible, where credibility denotes the belief a stakeholder has in the model, which is inherently subjective. Therefore, although increased model validity generally increases the credibility of a model, the two are not necessarily related (Balci, 1997; Aumann, 2007). Many aspects besides accuracy can influence this credibility. Van Delden et al. (2011a) identify several factors that determine the credibility, including transparency, saliency, usefulness and usability of a model, all of which benefit greatly from the interaction with stakeholders in the development process. Hence, while accuracy is not sufficient for a model to be credible, it should certainly be an essential component.

7.3 Directions for future research

This thesis contributes to the field of land-use science by developing and applying methods for the calibration and validation of land-use models. However, research in this direction is not finished as several issues remain open to further investigation. At least three main issues can be identified: 1) the selection of measures for process accuracy, 2) the influence of decisions upon model setup, and 3) new questions that arise from developments in land-use models.

Process accuracy in this thesis is assessed by applying metrics that characterize the generated land-use pattern, although the selection of these metrics remains subjective. Many metrics exist that can be used to characterize land-use patterns or changes in land-use patterns (Herold et al., 2002; Lausch and Herzog 2002). While many of these metrics are correlatated, others are not as they quantify different aspects of the land-use pattern, while others are strongly correlated (Riiters et al., 1995; McGarigal et al., 2009). Therefore, it is not clear which metric, or which combination of metrics, would be most appropriate for the assessment of land-use models (Peng et al., 2010). Moreover, the fact that a metric can be used to characterize some property of the landscape does not necessarily indicate that it is suitable as a measure for the assessment of land-use models. Hence, further research is needed to find out what landscape metric or combination of metrics best characterizes the process accuracy of a land-use model.

Developing a land-use model application requires several decisions that relate to the amount of detail, in the number of land-use classes, in the representation of land-use processes, and in the temporal and spatial resolution (Van Delden et al., 2011b). These decisions partly relate to practical issues and model requirements, such as the data availability or the type of questions that will be addressed using the land-use model. However, these decisions are also partly a subjective choice of the modeller and these choices will probably influence the model performance. Increased availability of high resolution data as well as the availability of increase computational power seem to inspire the current development towards more detailed models. Examples include models that represent large areas on a fine resolution (Verburg and Overmars, 2009; Van Delden et al., 2010) and models representing land-use change processes in more detail, including agent-based approaches (Valbuena et al., 2010; Robinson et al., 2012; Brady et al., 2012). However, detail should not be confused with accuracy, as it is not clear if more detailed models actually generate more accurate results. Consequently, the influence of detail, in terms of process thematic, spatial or temporal resolutions as well as process representation, on the predictive and process accuracy of land-use models are an important subject for further investigation.

Lastly, Land-use models as used in this study are mainly models that yield raster maps representing the predominant land use on a location. Consequently, the approach and methods presented in this thesis are applicable to these land-use models and their results. However, recent developments in land-use modelling demand for new calibration and validation methods, as their results are no longer restricted to categorical raster maps. This is particularly true for models that generate numerical raster maps, indicating quantitative information such as development density (Loibl et al., 2007) and population density (Van Vliet et al., 2012; White et al., 2012). Another recent development is the use of agent-based models to simulate land-use changes (An, 2012). While their results typically include raster based land-use maps, the emphasis on the agents' decision making process requires procedures for model testing that go beyond the assessment of simulated land-use changes only (Robinson et al., 2007).

References

Agarwal, C., Green, G.L., Grove, M., Evans, T., and Schweik, C. (2000). A review and assessment of land use change models: dynamics of space, time and human choice. US Forest Service and the centre of Study of Institutions, Populations, and Environmental Change (CIPEC). Bloomington, USA.

Albrecht, J. (2005). A new age for geosimulation. Transactions in geographic Information Science 30: 249-261.

Allouche, O., Tsoar, A, and Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43: 1223-1232.

Alonso, W. (1964) Location and land use. Harvard university press, Cambridge, MA.

An, L. (2012) Modelling human decisions in coupled human and natural systems: Review of agent based models. Ecological modelling 229: 25-36.

Anas, A., Arnott, R., and Small, K.A. (1998). Urban spatial structure. Journal of Economic Literature 36, 1426-1464.

Andersson, C., Lindgren, K., Rasmussen, S. and White, R. (2002a). Urban growth from “first principles”. Physical Review E 66 (026204): 1-9.

Andersson, C., Rasmussen, S., and White, R. (2002b). Urban settlement transitions. Environment and Planning B 29: 841-865.

Arai, T., and Akiyama, T. (2004). Empirical analysis for estimating land use transition potential functions – case study in the Tokyo metropolitan region. Computers Environment and Urban Systems 28: 65-84.

Arthur, W.B. (1990). Positive feedbacks in the economy. Scientific American 262: 92-99.

Arthur, W.B. (1999). Complexity and the economy. Science 284: 107-109.

Aumann, C.A. (2007). A method for developing simulation models of complex system. Ecological Modelling 202: 385-396.

Balci, O. (1997). Verification, validation and accreditation of simulation models. In: Andrádottir, S., Healy, K.J., Withers, D.H., and Nelson, B.L. Proceedings of the 1997 Winter Simulation Conference: 135-141.

Barredo, J., Demichelli, L., Lavalle, C., Kasanko, M., and McCormick, N. (2004). Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. Environment and Planning B 32: 65-84.

Barredo, J., Kasanko, M., McCormick, N., and Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and Urban Planning 64: 145-160.

Barredo, J., Lavalle, C., Demichelli, L., Kasanko, M., and McCormick, N. (2003). Sustainable urban and regional planning: The MOLAND activities on urban scenario modelling and forecast. Office for Official Publications of the European Communities, Luxembourg.

Batty, M. (2005). Cities and Complexity. The MIT press, Cambridge, MA.

Batty, M. and Longley, P. (1994). Fractal cities: a geometry of form and function. Academic, London.

Batty, M., and Torrens, P.M. (2005). Modelling and prediction in a complex world. Futures 37: 745-766.Batty, M., and Xie, Y. (1994). From cells to cities. Environment and Planning B 21: 31-48.

Batty, M. and Xie, Y. (1994). From cells to cities. Environment and Planning B 21: 31-48.

BC STATS (2006). Population projections 2006 – 2031. URL: http://www.gvrd.bc.ca/growth/keyfacts/popproj.htm (last accessed on March 23, 2007).

Bishop, Y., Fienberg, S., and Holland, P. (1975). Discrete multivariate analysis: theory and practice: 393-400. The MIT Press, Cambridge, MA.

Brady, M., Sahrbacher, C., Kellerman, K., and Happe, K. (2012). An agent-based approach to modelling impacts of agricultural policy on land use, biodiversity and ecosystem services. Landscape Ecology. In press, DOI: 10.1007/s10980-012-9787-3.

Brown, D.G., Page, S., Riolo, R., Zellner, M., and Rand, W. (2005). Path dependency and the validation of agent-based spatial models of land use. International Journal of Geographic Information Science 19 (2): 153-174.

Brown, D.G., Robinson, D.T., An, L. Nassauer, J.I., Zellner, M., Rand, M., Riolo, R., Page, S.E., Low, B., and Wang, Z. (2008). Exurbia from the bottom-up: Confronting empirical challenges to characterizing a complex system. Geoforum 39 (2): 805-818.

Chen, H., and Pontius, R. G. Jr. (2010). Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable. Landscape Ecology 25: 1319-1331.

Chen, Y. (2011). Derivation of the functional relations between fractal dimension of and shape index of urban form. Computers, Environment and Urban Systems 35: 442-451.

Cheng, T., and Molenaar, M. (1999). Objects with fuzzy spatial extend. Photogrammetric Engineering and Remote Sensing 65 (7): 797-801.

Claessens, L., Schoorl, J.M., Verburg, P.H., Geraedts, L., and Veldkamp, A. (2009). Modelling interactions and feedback mechanisms between land use change and landscape processes. Agriculture, Ecosystems and Environment 129 (1-3): 157-170.

Clarke, K.C. (2004). The limits of simplicity: toward geocomputational honesty in urban modelling. In GeoDynamics Atkinson, P., Foody, G., Darby, S. and Wu, F (eds.) GeoDynamics: 213-232. CRC Press, Boca Raton, USA.

Clarke, K.C., Gaydos, L. and Hoppen, S. (1996). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and Planning B 24: 247-261.

Cohen, J.A. (1960). Coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20: 37-46.

Constanza, R. (1989). Model goodness of fit: A multiple resolution procedure. Ecological Modelling 47: 199-215.

Cordoba, J-C. (2008). On the distribution of city sizes. Journal of Urban Economics 63: 177-197.

Couclelis H. (1997). From cellular automata to urban models: new principles for model development and implementation. Environment and Planning B 24: 165-174.

Couto, P. (2003). Assessing the accuracy of spatial simulation models. Ecological Modelling 167: 181-198.

Dendoncker, N., Rounsevell, M., and Bogaert, P. (2007). Spatial analysis and modelling of land use distributions in Belgium. Computers, Environment and Urban Systems, 31(2): 188-205.

De Nijs, A.C.M., De Niet, R., and Crommentuijn, L. (2004). Constructing land-use maps of the Netherlands in 2030. Journal of Environmental Management 72: 35-42.

De Nijs, A.C.M., and Pebesma, E.J. (2010). Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands. Environment and Planning B 37: 21-41.

Dietzel, C., and Clarke, K.C. (2007). Toward optimal calibration of the SLEUTH land use change model. Transactions in GIS 11(1): 29-45.

Doswell, C.A. III, Davies-Jones, R. and Keller, D.L. (1990). On summary measures of skill in rare event forecasting based on contingency tables. Weather and Forecasting 5: 576-585.

Engelen, G., White, R., and De Nijs, A.C.M. (2003). Environment Explorer: spatial support system for the integrated assessment of socio-economic and environmental policies in the Netherlands. Integrated Assessment 4 (2): 97-105.

Epstein, J.M. (1999). Agent-based computational models and generative social science. Complexity 4: 41-60.

Epstein, J.M. (2008). Why model? Journal of Artificial Societies and Social Simulation 11(4): 12.

Environmental Protection Agency U.S. (2000). Projecting land use change: A summary of models for assessing the effects of community growth and change on land use patterns. Cincinatty, USA. U.S. Environmental Protection Agency, Office of research and development 264.

Feranec, J., Hazeu, G., Christensen, S., and Jaffrain, G. (2007). Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia). Land Use Policy 24 (1): 234-247.

Fisher, P. (2000). Sorites paradox and vague geographies. Fuzzy Sets and Systems 113: 7-18.

Fisher, P., Arnot, C., Wadsworth, R., and Wellens, J. (2006). Detecting change in vague interpretations of landscapes. Ecological Informatics 1: 163 – 178.

Foody, G.M. (1996). Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. International Journal of Remote Sensing 17: 1317 – 1340.

Foody, G.M. (2002). Status of land cover classification accuracy assessment. Remote Sensing of Environment 80: 185–201.

Foody, G.M. (2006). What is the difference between two maps? A remote senser's view. Journal of Geographical Systems 8: 119-130.

Foody, G.M. (2008). Harshness in image classification accuracy assessment. International Journal of Remote Sensing 29 (11): 3137 – 3158.

Frankhauser, P. (1994). La fractilite des structures urbaines. Economica, Paris.

Frankhauser, P. (2004). Comparing the morphology of urban patterns in Europe – a fractal approach. European Cities – Insights on outskirts 2: 79-105.

Fritz, S., and See, L. (2005). Comparison of land cover maps using fuzzy agreement. International Journal of Geographical Information Science 19 (7): 787-807.

Furtado, B. Ettema, D., Ruiz, R.M., Hurkens, J. and Van Delden, H. (2012). A cellular automata intraurban model with prices and income differentiated actors. Environment and Planning B 39 (5): 897-924.

Gabaix, X. (1999). Zipf's law for cities: an explanation. The Quarterly Journal of Economics 114 (3): 739-767.

Gallego, J., and Peedell, S. (2001). Using CORINE land cover to map population density. In: European Environment Agency (2001). Towards agri-environmental indicators; integrating statistical and administrative data with land cover information. EEA, Copenhagen: 92 – 104.

Gardner, R.H., and Urban, D.L. (2005). Model validation and testing: Past lessons, present concerns and future prospects. In: Canham, C.D. (Ed.) Models of Ecosystems Science: 184-203. Princeton University press, Princeton, USA.

Grimm, V. (1999). Ten years of individual based modelling in ecology: what have we learned and what could we learn in the future? Ecological Modelling 115: 129-148.

Geertman, S., Hagoort, M., and Ottens, H. (2007). Spatial-temporal specific neighbourhood rules for cellular automata land-use modelling. International Journal of Geographical Information Science 21(5): 547-568.

Geertman, S., and Stillwell, J. (2004). Planning support systems: an inventory of current practice. Computers, Environment and Urban Systems 28: 291-310.

GVRD (1999). Livable Region Strategic Plan. URL: http://www.gvrd.bc.ca/growth/lrsp/LRSP.pdf (Last accessed on March 23, 2007).

Haase, D., Lautenbach, S., Seppelt, R. (2010). Modelling and simulating residential mobility in a shrinking city using an agent based approach. Environmental Modelling and Software 25: 1225-1240.

Haines-Young, R. and Weber, J.L. (2006). Land accounts for Europe 1990-2000. EEA report 11/2006. European Environment Agency, Copenhagen.

Hagen, A. (2002). Multi method assessment of map similarity. Proceedings of the 5th AGILE Conference on Geographic Information Science, 25-27 April 2002, Palma: 171-182. Edited by M. Ruiz, M. Gould, J. Ramon, Universitat de les Illes Balears, Palma, Spain.

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

Hagen-Zanker, A. (2009). An improved fuzzy kappa statistic that accounts for spatial autocorrelation. International Journal of Geographical Information Science 23(1): 61-73.

Hagen-Zanker, A. and Lajoie, G. (2008). Neutral models of landscape change as benchmarks in the assessment of model performance. Landscape and Urban Planning 86 (3-4): 284-296.

Hagen-Zanker, A., and Martens, P. (2008). Map comparison methods for comprehensive assessment of geosimulation models. In: Gervasi, O., Murgante, B., Lagana, A., Taniar, D., Mun, Y. and Gavrilova, M. (eds.) Lecture Notes in Computer Science, Proceedings of the 2008 International Conference on Computational Science and Its Applications ICCSA 2008: 194-209, Springer, Berlin, Germany.

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

Hägerstrand, T. (1967). Innovation diffusion as a spatial process. University of Chicago Press, Chicago, IL.

Hagoort, M., Geertman, S., and Ottens, H. (2008). Spatial externalities, neighbourhood rules and CA land-use modelling. Annals of Regional Science 42: 39-56.

Hansen, H.S. (2007). An adaptive land-use simulation model for integrated coastal zone planning. In: Fabrikant, S.I., and Wachowicz, M. (Eds.), The European Information Society: Leading the Way with Geo-information. Springer-Verlag, Berlin: 35–53. Lecture Notes in Geoinformation and Cartography.

Hansen, H.S. (2012). Empirically derived neighbourhood rules for urban land-use modelling. Environment and Planning B 39 (2): 213-228.

Hartmann, S. (1996). The world as a process: Simulations in the natural and social sciences. In: Hegselmann, R., Mueller, U., and Troitsch, K. (eds.) Modelling and simulation in the social sciences from the philosophy of science point of view: 77-100. Kluwer, Dordrecht, The Netherlands.

Harvey, J.T. (2002). Estimating census district populations from satellite imagery: some approaches and limitations. International Journal of Remote Sensing 23 (10): 2071-2095.

Heidke, P. (1926). Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst. Geografiske Annaler 8: 301-349.

Hellman, F., and Verburg, P.H. (2010). Impact assessment of the European biofuel directive on land use and biodiversity. Journal of Environmental Management 91(6): 1389-1398.

Hellmann, F. and Verburg, P.H. (2011). Spatially explicit modelling of biofuel crops in Europe. Biomass and Bioenergy 35 (6): 2411-2424.

Huang, J., Lu, X.X., and Sellers, J.M. (2007). A global comparative analysis of urban form: Applying spatial metrics and remote sensing. Landscape and Urban Planning 82: 182-197.

Irwin, E.G., and Bockstael, N.E. (2002). Interacting agents, spatial externalities and the evolution of residential land use patterns. Journal of Economic Geography 2: 31-54.

Jakeman, A.J., Letcher, R.A., and Norton, J.P. (2006). Ten iterative steps in the development and evaluation of environmental models. Environmental Modelling and Software 21: 602-614.

Janssen, M., and Ostrom, E. (2006). Empirically based, agent based models. Ecology and Society 11 (2): 37.

Jantz, C.A., Goetz, S.J., and Shelley, M.K. (2004). Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. Environment and Planning B: 251-271.

Jenerette, G.D. and Wu, J. (2001). Analysis and simulation of land-use change in the central Arizona – Phoenix region, USA. Landscape Ecology 16: 611, 626.

Kleindorfer, G.B., O'Neill, L., and Ganeshan, R. (1998). Validation in simulation: various positions in the philosophy of science. Management Science 44 (8): 1087-1099.

Kocabas, V., and Dragisevic, S. (2006). Assessing cellular automata model behaviour using a sensitivity analysis approach. Computers, Environment and Urban Systems 30 (6): 921-953.

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

Konikow, L.F., and Bredehoeft, J.D. (1992). Ground-water models cannot be validated. Advances in Water Resources 15: 75-83.

Koomen, E., Stillwell, J., Bakema, A., and Scholten H.J. (Editors) (2007). Modelling land use change: Progress and applications. Geojournal Library 90. Springer, Dordrecht, the Netherlands.

Krugman, P. (1991). Increasing returns and economic geography. Journal of Political Economy 99 (3): 483-499.

Krugman, P. (1996). Confronting the mystery of urban hierarchy. Journal of the Japanese and International Economies 10 (4): 399-418.

Krugman, P. (1998). What's new about the new economic geography? Oxford review of economic policy 14(2): 7-17.

Lambin, E.F., Turner, B.L., Geist, H.J., Agloba, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemand, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skanes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., and Xu, J. (2001). The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change 11: 261-269.

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

Li, X., and Yeh, A.G.O. (2004). Data mining of cellular automata's transition rules. International Journal of Geographical Information Science 18 (8): 723-744.

Loibl, W., and Tötzer, T. (2003). Modelling growth and densification processes in suburban regions – simulations of landscape transition with spatial agents. Environmental Modelling and Software 18: 553-563.

Loibl, W., Tötzer, T., Köstl, M., and Steinnocher, K., (2007) Simulation of polycentric urban growth dynamics through agents. In: Koomen, E., Stillwell, J., Bakema, A., and Scholten, H-J. Modelling land-use change. The Geojournal Library 90 (IV): 219-236.

Luck, G.W. (2007). A review of the relationship between human population density and biodiversity. Biological Review 82: 607-645.

Luoto, M., Pöyry, J., Heikkinnen, R.K., and Saarinen, K. (2005). uncertainty of bioclimate envelope models based on the geographical distribution of species. Global Ecology and Biogeography 14: 575-584.

Ma, Z., and Redmond, R.L. (1994). Tau coefficients for accuracy assessment of classification of remote sensing data. Photogrammetric Engineering and Remote Sensing, 61 (4): 435 – 439.

Makridakis, S., and Taleb, N.N. (2009). Living in a world of low levels of predictability. International Journal of Forecasting 25: 840-844.

Manson, S.M. (2001). Simplifying complexity: a review of complexity theory. Geoforum 32: 405-414.

Manson, S.M. (2007). Challenges in evaluating model of geographic complexity. Environment and Planning B 34: 245-260.

McGarigal, K., Cushman, S.A., Neel, M.C., and Ene, E. (2002). FRAGSTATS v3: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst.

McGarigal, K. and Marks, B.J. (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure, General technical report PNW-GTR- 351. Portland, OR: USDA Forest Service, Pacific Northwest Research Station.

McGarigal, K., Tagil, S., and Cushman, S. (2009). Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landscape Ecology 24: 433-450.

Menard, A., and Marceau, D.J. (2005). Exploration of spatial scale sensitivity in geographic cellular automata. Environment and Planning B 32: 693-714.

Monserud R.A., and Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. Ecological Modelling 62: 275-293.

Moss, S. (2002). Policy analysis from first principles. Proceedings of the US National Academy of Sciences 99 (3): 7267-7274.

Oreskes, N., Shrader-Frechette, K., and Belitz, K. (1994). Verification, validation and confirmation of numerical models in the earth sciences. Science 263 (5147): 641-646.

Parker, D.C., Hessl, A., and Davis, S.C. (2008). Complexity, land-use modelling, and the human dimension: Fundamental challenges for mapping unknown outcome spaces. Geoforum 39: 789-804.

Parker, D.C., Manson, S.M., Janssen, M.A., Hoffman, M.J., and Deadman, P. (2003). Multi-Agent systems for the simulation of land use and land cover change: A review. Annals of the association of American Geographers 92 (2): 314-337.

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

Pawson, R., Wong, G., and Owen, L. (2011). Known knowns, known unknowns, and unknown unknowns: the predicament of evidence-based policy. American Journal of evaluation 32 (4): 518-546.

Peng, J., Wang, Y., Zhanh, Y., Wu, J., Li, W., and Li Y. (2010) Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. Ecological Indicators 10: 217 – 223.

Poelmans, L., and van Rompay, A. (2010). Complexity and performance of urban expansion models. Computers, Environment and Urban Systems 34 (1): 17-27.

Pontius Jr., R.G. (2002). Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogrammetric Engineering and Remote Sensing 68 (10): 1041-1049.

Pontius Jr, R.G., Boersma, W., Castella, J.C., Clarke, K., de Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Mohdsood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T.N., Veldkamp, A.T., and Verburg, P.H. (2008). Comparing the input, output, and validation maps for several models of land change. Annals of Regional Science, 42 (1): 11-47.

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

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

Pontius Jr., R.G., and Millones, M. (2011). Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing 32 (15): 4407-4429.

Pontius Jr., R.G., and Neeti, N. (2010). Uncertainty in the difference between maps of future land change scenarios. Sustainability Science 5 (1): 39-50.

Pontius Jr., R.G., Shusas, E., and McEachren, M. (2004b). Detecting important categorical land changes while accounting for persistence. Agriculture, Ecosystems and Environment 101: 251-268.

Power, C., Simms, A. and White, R. (2001). Hierarchical fuzzy pattern matching for regional comparison of land-use maps. International Journal of Geographical Information Science 15 (1): 77-100.

Power, M (1993). The predictive validation of ecological and environmental models. Ecological modelling 68 (1-2): 33-50.

Reed, W.J. (2002). On the rank-size distribution for human settlements. Journal of Regional Science 42 (1): 1-17.

Refsgaard. J.C. and Henriksen, H.J. (2004). Modelling guidelines – terminology and guiding principles. Advances in Water Resources 27: 71 – 82.

Remmel, T., and Csillag, F. (2003). When are two landscape pattern indices significantly different? Journal of Geographical Systems 5: 331-351.

Riitters, K., O'Neill, R., Hunsaker, C., Wickham, J, Yankee, D., Timmins, S., Jones, K., and Jackson, B. (1995). A factor analysis of landscape pattern and structure metrics. Landscape Ecology 10 (1): 23-39.

Riiters, K.H., Wickham, J.D., and Wade, T.G. (2009). An indicator of forest dynamics using a shifting landscape mosaic. Ecological Indicators 9: 107-117.

Robinson, D.T., Brown, D.G., Parker, D.C., Schreinemachers, P., Janssen, M.A., Huigens, M., Wittmer, H., Gotts, N., Promburom, P., Irwin, E., Berger, T., Gatzweiler, F., and Barnaud, C. (2007). Comparison of empirical methods for building agent-based models in land use science. Journal of Land Use Science 2 (1): 31-55.

Rutherford, G.N., Bebi, P., Edwards, P.J. and Zimmerman, N.E. (2008). Assessing land-use statistics to model land cover change in a mountainous landscape in the European Alps. Ecological Modelling 212: 460-471.

Rykiel, E.J.Jr. (1996). Testing ecological models: the meaning of validation. Ecological Modelling 90: 229 – 244.

Sargent, R.G. (1998). Verification and validation of simulation models. In Medeiros, D.J., Watson, E.F. Carson, J.S., and Manivannan, M.S. Proceedings of the 1998 Winter Simulation Conference: 121-130.

Schwarz, N. (2010). Urban form revisited – Selecting indicators for characterising European cities. Landscape and Urban Planning 96: 29-47.

Schulp, C.J.E., Nabuurs, G.J., and Verburg, P.H., (2008). Future carbon sequestration in Europe: Effects of land-use change. Agriculture, Ecosystems and Environment, 127 (3-4): 251-264.

Sieber, S., Zander, P., Verburg, P.H., Van Ittersum, M., (2010). Model-based systems to support impact assessment – Methods, tools and applications. Ecological Modelling 221(18): 2133-2135.

Silva, E.A., and Clarke, K.C., (2002). Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Computers, Environment and Ecosystems 26: 525-552.

Stanilov, K. and Batty, M. (2011). Exploring the historical determinants of urban growth patterns through cellular automata. Transactions in GIS 15(3): 253-271.

Straatman, B., White, R., and Engelen, G. (2004). Towards an automatic calibration procedure for constrained cellular automata. Computers, Environment and Urban Systems 28: 149-170.

Tang, G, (2008). A new metric for evaluating the correspondence of spatial patterns in vegetation models. Global Ecology and Biogeography 17 (4): 465-478.

Thomlinson, J.R., Bolstad, P.V., and Cohen, W.B. (1999). Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products. Remote Sensing of Environment 70: 16-28.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. Economic Geography 46 (2): 234-240.

Torrens, P.M. (2011). Calibrating and validating cellular automata models of urbanization. In: Yang, X. (Ed.) Urban remote sensing: monitoring, synthesis and modelling in the urban environment: 335-345. John Wiley and Sons, Chichester.

Townsend, P.A. (2000). A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications. Remote Sensing of Environment 72: 253-267.

Turner, M.G., Constanza, R. and Sklar, F. (1989). Methods to evaluate the performance of spatial simulation models. Ecological Modelling 48 (1 – 2): 1-18.

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

Van Delden, H., Gutiérrez, E.R., Van Vliet, J. and Hurkens, J. (2008). Xplorah, a multi-scale integrated land use model. In: Sànchez-Marrè, M., Béjar, J., Comas, J., Rizzoli, A., Guariso, G. (Eds.) Proceedings of the iEMSs Fourth Biennial Meeting: "Integrating Sciences and Information Technology for Environmental Assessment and Decision Making". International Environmental Modelling and Software Society, Barcelona, Spain.

Van Delden, H., and Hurkens, J. (2011). A generic Integrated Spatial Decision Support System for urban and regional planning. In Chan, F., Marinova, D. and Anderssen, R.S. (eds) MODSIM2011, 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011: 127-139.

Van Delden, H., Luja, P., and Engelen, G. (2007). Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management. Environmental Modelling and Software 22: 223-238.

Van Delden, H., Seppelt, R., White, R., and Jakeman, A.J. (2011a) A methodology for the design and development of integrated models for policy support. Environmental Modelling & Software 26:266-279.

Van Delden, H., Stuczynski, T., Ciaian, P., Paracchini, M.L., Hurkens, J., Lopatka, A., Gomez, O., Calvo, S., Shi, Y., van Vliet, J., and Vanhout, R. (2010). Integrated assessment of agricultural policies with dynamic land use change modelling. Ecological Modelling 221: 2153-2166.

Van Delden, H., Van Vliet, J., Rutledge, D.T., and Kirkby, M.J. (2011b). Comparison of scale and scaling issues in integrated land-use models for policy support. Agriculture, Ecosystems and Environment, volume 142, issue 1-2: 18-28

Van Vliet, J., Bregt, A.K., and Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land-use change models. Ecological Modelling 222 (8): 1367-1375.

Van Vliet, J., Dragicevic, S., and White, R (2009). Modelling urban growth in a variable grid cellular automaton. Computers, Environment and Urban systems 33 (1): 35-43.

Van Vliet, J. and Van Delden, H. (2008). An activity based cellular automaton model to simulate land use changes. Proceedings of the International Congress on Environmental Modelling and Software, July 7 – 10 2008, Barcelona, Spain.

Veldkamp, A. and Lambin, A.F. (2001). Predicting land use changes, Agriculture, Ecosystems and Environment 85: 1-6.

Verburg, P.H., De Nijs, T., Ritsema van Eck, J., Visser, H. and Jong, K. (2004a). A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems 28: 667-690.

Verburg P.H., and Overmars K.P. (2009). Combining top-down and bottom-up dynamics in land use modelling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. Landscape Ecology 24 (9): 1167-1181.

Verburg, P.H., Ritsema van Eck, J., De Nijs, T., Dijst, M. and Schot, P. (2004b). Determinants of land use patterns in the Netherlands. Environment and Planning B 31: 125-150.

Verburg, P.H., Schot, P.P., Dijst, M.J., and Veldkamp, A. (2004c). Land use change modelling: current practice and research priorities. GeoJournal 61: 309 – 324.

Verburg, P.H., Schulp, C.J.E., Witte, N., and Veldkamp, A. (2006). Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agriculture, Ecosystems & Environment (1): 39-56.

Verburg, P.H., Tabeau, A. and Hatna, E. (In press). Assessing spatial uncertainties in land allocation using a scenario approach and sensitivity analysis: A study for land use in Europe. Journal of Environmental Management.

Visser, H., and De Nijs, A.C.M. (2006). The Map Comparison Kit. Environmental Modelling and Software 21 (3): 346-358.

Von Thünen, J. H. (1826). Der Isolierte Staat in Beziehung der Landwirtschaft und Nationalökonomie. English translation: P. Hall (ed.) Von Thünen’s Isolated State. Pergamon, Oxford.

Walker, R. (2003). Evaluating the performance of spatially explicit models. Photogrammetric Engineering & Remote Sensing 69(11): 1271-1278.

Wang, F., Hasbani, J-G., Wang, X., and Marceau, D. (2011). Identifying dominant factors for the calibration of a land use cellular automata model using Rough Set Theory. Computers, Environment and Urban Systems 35 (2): 116-125.

Wear, D.N., Turner, M.G., and Nadman, R.J., 1998. Land cover along an urban gradient: Implications for water quality. Ecological Applications 8 (3), pp. 619-630.

Wickramasuriya, R.C., Bregt, A.K., Van Delden, H., and Hagen-Zanker, A. (2009). The dynamics of shifting cultivation captured in an extende Constrained Cellular Automata land use model. Ecological Modelling 220: 2302-2309.

Wilkinson, G.G. (2005). Results and implications of a study of fifteen years of satellite image classification experiments. IEEE Transactions on Geoscience and Remote Sensing 43: 433-440.

White, R. (2005). Modelling Multi-scale processes in a cellular automata framework. In J. Portugali, ed.: Complex Artificial Environments, Springer-Verlag: 165-178.

White, R. (2006). Pattern based map comparisons. Journal of Geographical Systems 8: 145-164.

White, R., and Engelen, G. (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of land use patterns. Environment and Planning A 25: 1175-1199.

White, R., and Engelen, G. (1997). Cellular automata as the basis of integrated dynamic regional modelling. Environment and Planning B 24: 235-246

White, R., and Engelen, G. (2000). High resolution integrated modelling of the spatial dynamics of urban and regional systems. Computers, Environment and Urban Systems 24: 383-400.

White, R., and Engelen, G. (2003). A calibration procedure for constrained large neigh-bourhood cellular automata based land use models. Paper presented at the 13th European Colloquium on Theoretical and Quantitative Geography, Lucca, Italy.

White, R., Engelen, G., and Uljee, I (1997). The use of constrained cellular automata for high-resolution modelling of urban land use dynamics. Environment and Planning B 24: 323-343.

White, R., Uljee, I., and Engelen, G. (2012) Integrated modelling of population, employment, and land-use change with a multiple-activity based variable grid cellular automaton. International Journal of Geographical Information Science 26 (7): 1251-1280.

Wilkinson, G.G. (2005). Results and implications of a study of fifteen years of satellite image classification experiments, IEEE Transactions on Geoscience and Remote Sensing 43: 433-440.

Wolfram, S. (1984). Universality and complexity in cellular automata. Physica D 10: 1-35.

Wu, C., and Murray, A.T. (2007). Population estimation using landsat enhanced thematic mapper imagery. Geographical Analysis 39 (1): 26-43.

Wu, F. (1998). SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. International Journal of Geographical Information Systems 12 (1): 63-82.

Wu, F. (1998). An experiment on the generic polycentricity of urban growth in a cellular automatic city. Environment and Planning B 25: 731-752.

Wu F., and Webster C.J. (1998). Simulation of land development through the integration of cellular automata and multi-criteria evaluation. Environment and Planning B: 103-126.

Wu, F., and Webster C.J. (2000). Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. International Journal of Geographical Information Science 17 (7): 625-648.

Wyman, M.S., and Stein, T.V. (2010). Modelling social and land-use/land-cover change data to assess smallholder deforestation in Belize. Applied Geography 30 (30): 329-342.

Yeh, A.G.O., and Li, X. (2002). A cellular automata model to simulate development density for urban planning. Environment and Planning B 29: 431-450.

Yang, X., and Lo, C.P. (2003). Modelling Urban growth and landscape changes in the Atlanta metropolitan area. International journal of Geographical Information Science 17(5): 463-488.

Zhang, Q, Ban, Y, Liu, J., and Hu, Y. (2011). Simulating and analysis of urban growth scenarios of the Greater Shanghai Area, China. Computers, Environment and urban Systems 35 (2): 126-139.

Summary

Land use is constantly changing. For example, urban areas expand as a result of population growth, cropping patterns change to fulfil the demand for bioenergy and natural vegetation recovers in locations where farmers cease to farm. Understanding these changes is pivotal to explore future land-use scenarios and to design spatial policies. Land-use models are increasingly being used for these purposes. They function as virtual laboratories in which scientists or policy analysts can conduct experiments. In order to reliably apply models for these purposes, they need to be calibrated, where calibration is the adjustment of parameters to improve the model's performance. Consequently, the value of modelled land-use scenarios and policy assessments depends on the quality of the calibration. Assessment of the quality of the calibration is termed validation, and is ideally performed independently in the sense that the data that is used for validation has not been used for calibration.

The development of a land-use model can be described by four sequential phases: analysis and conceptual modelling, computer programming of the conceptual model, calibration of the computerized model, and experimentation with the calibrated model. As described in chapter 1, the first three phases have their own evaluation procedures: conceptual validation, code verification and operational validation, respectively. The operational validation provides insights into the strengths and weaknesses of a particular model application, and sometimes it can suggest directions for improvement. However, available assessment methods have limitations for their application in land-use modelling. Therefore, there is a demand to develop and apply more appropriate methods. The work presented in this thesis addresses this challenge in four research questions:

1. What characteristics of land-use models are important for assessing these models?

2. How can the predictive accuracy of a land-use model be assessed?
3. How can the process accuracy of a land-use model be assessed?
4. How can the neighbourhood rules in cellular automata land-use models be calibrated and validated?

All four questions are dealt with using or building on the Metronamica land-use model, a constrained cellular automata land-use model. The first three questions are general, meaning that their answers are relevant to a wide range of land-use models. The fourth question is more specific to the land-use model applied in this research: neighbourhood rules, which represent the influence of the existing land-use distribution on the location of future land-use changes, are the defining characteristic of cellular automata based land-use models.

Many land-use changes are directly or indirectly the result of human decisions. However, human decisions are inherently uncertain, and therefore land-use models cannot be expected to simulate these land-use changes exactly. This is acknowledged by many land-use models, including Metronamica, as they use a stochastic component to simulate land-use changes. Therefore, land-use models should not only be validated on their capacity to accurately allocate land-use changes on the map, but also on their capacity to realistically simulate land-use change processes. Moreover, many models start from an initial land-use map and simulate changes relative to this map. The amount of change for a simulation is typically small relative to the entire map, which means that a large part of the result is caused by persistence. For this reason, a benchmark, such as a naive predictor, is required to properly asses the accuracy of simulation results. This benchmark can be implicit, that is, the assessment method accounts for the information available from the initial map, or explicit, i.e. simulation results are compared with results from another land-use model that serves as a reference. Such benchmarks have been applied throughout this thesis as a basis for model assessment. One exception to this are synthetic model applications. Chapter 6 presents a new land-use model, which is assessed using a synthetic application. Because no observed data was available, the accuracy of changes cannot be assessed, and no benchmark model was applied for reference.

The predictive accuracy of a land-use model is typically assessed by comparing a simulation result with the actual land-use map at the end of a simulation. A common method for this is the Kappa statistic, which expresses the agreement between two land-use maps corrected for the expected agreement from a random allocation given the distribution of class sizes. However, this is not an appropriate reference level to assess the predictive accuracy of land-use models, because it does not account for the amount of change. Chapter 2 presents Kappa Simulation, a new map comparison method that is identical in form to Kappa, but which instead applies a more appropriate reference model based on random allocation of class transitions relative to the initial map. This implicitly accounts for the amount of change, which truly allows gauging the predictive accuracy of changes in land-use models. However, Kappa Simulation cannot differentiate between near-hits and complete misses, while this distinction is often very relevant for land-use modellers. Chapter 3 therefore presents Fuzzy Kappa Simulation. This statistic is an improvement of Kappa Simulation, as it applies a fuzzy interpretation of class transitions and their locations. This means that it can account for near-hits, which makes it arguably the most suitable map comparison method to assess the predictive accuracy of land-use models.

Because of the intrinsic uncertainty underlying land-use change processes, a realistic land-use model does not necessarily generate a high predictive accuracy. Therefore, it is worth assessing process accuracy and predictive accuracy separately. Ideally, process accuracy is assessed directly from the values of model parameters. However, it is often impossible to observe real-world values for these parameters because drivers for land-use changes are correlated or they cannot be measured. Therefore, the process accuracy is typically assessed indirectly from the land-use patterns generated by the model. Two groups of methods exist to characterize land-use patterns: landscape metrics and fractal metrics. Landscape metrics are a collection of algorithms that have been applied in landscape ecology to characterize land-use patterns. In this thesis, landscape metrics have been used to compare the simulated land-use map with the observed land-use map instead. Similar characteristics indicate that a land-use model generates a realistic land-use pattern, as is demonstrated in chapter 4. Fractal metrics, which have their origin in complexity science, are another type of measures to characterize regularities in (urban) land-use patterns. Examples are power-law distributions for urban clusters and fractal dimensions of patches of urban land. This thesis applies fractal metrics to compare simulated land-use patterns with observed land-use patterns, as is demonstrated in chapter 5. Moreover, fractal metrics can be interpreted in absolute terms since they represent general regularities in urban systems for which values are known from literature. Therefore, fractal metrics also allow evaluation of the process accuracy of a synthetic application for which no observed land-use pattern is available for comparison, as is shown in chapter 6.

Neighbourhood rules represent the influence of the existing land-use distribution on the location of land-use changes. This includes the persistence, conversion and attraction/repulsion of land uses in the neighbourhood of a location. Because neighbourhood rules cannot be estimated directly from data, they need to be set in a calibration procedure. Chapter 4 assesses whether larger neighbourhoods improve the accuracy of a cellular automata land-use model.

Results indicate that agents consider their neighbourhood at different spatial scales: the direct vicinity of a location has a strong influence on the allocation of new urban land, but neighbourhood rules over larger distance – typically the size of urban regions – also improve the model performance. Chapter 5 addresses the calibration of neighbourhood rules using measurements of the over- or underrepresentation of land uses in the neighbourhood of land-use changes to compare simulated and observed land-use changes. This study shows that this calibration method considerably improves the process accuracy of the applied model. Chapter 6 discusses a special type of neighbourhood rules: rules describing the influence of the existing activity distribution on the allocation of activity changes, where activities denote a quantity or density related to a land use, such as inhabitants for residential land use. This study shows that relatively simple rules can grow a realistic urban settlement structure, which also confirms that neighbourhood rules improve the process accuracy of a land-use model.

The research described this thesis contributes to the calibration and validation of land-use models by introducing and applying several methods to objectively assess the predictive accuracy and the process accuracy of land-use models. In both types of assessments it is important to include a benchmark to interpret the results, either implicitly in the method or explicitly by applying a reference model. Outperforming the benchmark can be considered a minimum threshold to pass; however, it cannot directly inform whether a model is acceptable as this depends on the purpose of the model, the requirements of the study and the application domain. In this context, it should be noted that while the methods presented and applied in this thesis are objective, the selection of assessment methods remains subjective. Moreover, because no method is yet capable of describing land-use patterns satisfactorily, more subjective methods such as visual assessment of simulation results or interpretation of parameter values remain of added value in the calibration and validation of land-use models.

Samenvatting

Landgebruik verandert steeds: stedelijke gebieden breiden uit omdat de bevolking groeit, landbouw gewassen veranderen om aan de vraag naar biobrandstof te voldoen en natuurlijke vegetatie komt terug op plaatsen waar boeren het land verlaten. Om scenarios voor toekomstig landgebruik te verkennen, of om ruimtelijk beleid te maken, is het essentieel om deze veranderingen te begrijpen. In toenemende mate worden hiervoor landgebruikmodellen ingezet: deze dienen als virtuele laboratoria, waar onderzoekers en beleidsanalisten kunnen experimenteren. Om landgebruikmodellen op een betrouwbare manier toe te passen, moeten ze gekalibreerd worden, waar kalibreren is gedefinieerd als het aanpassen van parameters om de modelresultaten te verbeteren. De waarde van gemodelleerde landgebruikscenario's of beleidsevaluaties is daarom afhankelijk van de kwaliteit van de kalibratie. Het beoordelen van de prestatie van een model, en dus de kwaliteit van de kalibratie, heet validatie. Validatie gebeurt bij voorkeur onafhankelijk, in die zin dat de gegevens die gebruikt worden voor kalibratie niet ook gebruikt worden voor validatie.

Het ontwikkelen van een landgebruikmodel kan worden beschreven in vier aparte fases: analyse en conceptueel modeleren, computer programmeren, kalibratie van het computermodel en experimenten met het gekalibreerde model. De eerste drie fases hebben ieder eigen evaluatie procedures: respectievelijk conceptuele validatie, verificatie van de code, en operationele validatie. De operationele validatie geeft inzicht in de sterke en zwakkere kanten van een model toepassing of model kalibratie, en geeft soms aanwijzingen voor verbeteringen. Echter, bestaande methodes hiervoor hebben hun beperkingen voor het beoordelen van landgebruikmodellen. Het is daarom een uitdaging om beter passende methodes te ontwikkelen en toe te passen. Deze uitdaging is het onderwerp van dit proefschrift, welke wordt behandeld aan de hand van vier onderzoeksvragen:

1. Welke eigenschappen van landgebruikmodellen zijn belangrijk voor het beoordelen van hun resultaten?
2. Hoe kan de voorspellende nauwkeurigheid van een landgebruikmodel worden beoordeeld?
3. Hoe kan de proces nauwkeurigheid van een landgebruikmodel worden beoordeeld?
4. Hoe kunnen de omgevingsregels in een landgebruikmodel gebaseerd op een cellulaire automaat worden gekalibreerd en gevalideerd?

Alle vier deze vragen zijn behandeld met behulp van Metronamica, een landgebruikmodel gebaseerd op een cellulaire automaat. De eerste drie vragen zijn algemene vragen, in die zin dat de resultaten toepasbaar zijn voor een breder scala aan landgebruikmodellen. De vierde vraag is specifiek voor landgebruikmodellen op basis van een cellulaire automaat, omdat omgevingsregels het kenmerkende onderdeel van deze modellen is: Ze vertegenwoordigen de invloed van bestaand landgebruik op de locatie van toekomstige landgebruik veranderingen.

Veel landgebruik veranderingen zijn direct of indirect een gevolg van menselijke beslissingen. Deze beslissingen zijn inherent onzeker. Daarom kunnen landgebruikmodellen deze veranderingen nooit exact simuleren. Veel landgebruikmodellen, onder meer Metronamica, onderkennen dit door een stochastische component te gebruiken voor het simuleren van landgebruik veranderingen. Om die reden moeten landgebruikmodellen niet slechts beoordeeld worden op hun voorspellende nauwkeurigheid - het vermogen landgebruik nauwkeurig te plaatsen op een kaart - maar ook op hun proces nauwkeurigheid – het vermogen landgebruikprocessen realistisch te simuleren. Daarnaast beginnen veel modellen met een initiële landgebruikkaart, en simuleren veranderingen ten opzichte daarvan. De hoeveelheid verandering is meestal klein ten opzichte van de totale kaart, waardoor een groot deel van het eindresultaat het gevolg is van persistentie. Daarom is het nodig een referentiemodel te gebruiken bij het beoordelen van de resultaten van landgebruikmodellen, zoals een naïeve voorspeller. Dit referentiemodel kan impliciet zijn, bijvoorbeeld door een methode die rekening houdt met de informatie welke bekend is uit de initiële kaart, of expliciet, bijvoorbeeld door resultaten te vergelijken met resultaten van een ander model dat als referentie dient. Dergelijke referentiemodellen zijn toegepast als basis voor het beoordelen van landgebruikmodellen in hoofdstuk 2, 3, 4, en 5 van dit proefschrift. Een uitzondering hierop vormen synthetische model toepassingen: omdat hiervoor geen geobserveerde data bestaat is, is het niet mogelijk de nauwkeurigheid van veranderingen te beoordelen. Daarom kan er hiervoor geen referentiemodel worden toegepast. Hoofdstuk 6 beschrijft een nieuw landgebruikmodel, welke beoordeeld is met behulp van een synthetische applicatie, en dus zonder referentiemodel.

De voorspellende nauwkeurigheid van een landgebruikmodel wordt meestal beoordeeld door de gesimuleerde landgebruikkaart te vergelijken met de werkelijke landgebruikkaart aan het einde van de simulatie. Een veelgebruikte methode hiervoor is de Kappa statistiek. Deze geeft de overeenkomst tussen twee landgebruikkaarten, gecorrigeerd voor de verwachte overeenkomst wanneer dezelfde verdeling van landgebruikklassen willekeurig over de kaart was verdeeld. Echter, dit is geen geschikt referentiemodel voor landgebruikmodellen, omdat het geen rekening houdt met de hoeveelheid verandering. Hoofdstuk 2 presenteert Kappa Simulation, een nieuwe vergelijkingsmethode, die dezelfde vorm heeft als Kappa, maar welke in plaats daarvan een geschikter referentiemodel toepast dat gebaseerd is op de verdeling van landgebruiktransities. Dit houdt impliciet rekening met de hoeveelheid verandering waardoor het werkelijk de voorspellende nauwkeurigheid van landgebruikmodellen kan bepalen. Echter, Kappa Simulation kan geen onderscheid maken tussen bijna juist voorspelde veranderingen en compleet onjuiste, terwijl dit voor landgebruikmodelleurs een wezenlijk verschil is. Derhalve presenteert hoofdstuk 3 Fuzzy Kappa Simulation. Deze statistiek is een verbetering van Kappa Simulation, omdat het een ruimere interpretatie van landgebruikveranderingen en hun locaties gebruikt. Daardoor kan deze methode ook bijna juist voorspelde veranderingen waarderen, waardoor het wellicht de meest geschikte methode is om de voorspellende nauwkeurigheid van landgebruikmodellen te bepalen.

Vanwege de intrinsieke onzekerheid die ten grondslag ligt aan landgebruikveranderingen, zal een realistisch landgebruikmodel niet per se in een hoge voorspellende nauwkeurigheid resulteren. Daarom is het zinnig om de voorspellende nauwkeurigheid en de proces nauwkeurigheid apart te beoordelen. Idealiter wordt de process nauwkeurigheid direct beoordeeld aan de hand van parameter waardes, maar het is vaak niet mogelijk om deze waardes in werkelijkheid te observeren omdat de onderliggende factoren gecorreleerd zijn of omdat ze simpelweg niet te meten zijn. Daarom wordt de proces nauwkeurigheid vaak indirect gemeten, aan de hand van de gegenereerde landgebruikpatronen. Er bestaan twee groepen methodes om patronen van landgebruik te karakteriseren: landschapsmaten en fractale maten. Landschapsmaten zijn een verzameling algoritmes die worden toegepast om landschappen te karakteriseren in de ecologie. In dit proefschrift zijn landschapsmaten gebruikt om de gesimuleerde landgebruikkaart te vergelijken met de werkelijke landgebruikkaart. Wanneer beiden dezelfde karakteristieken laten zien is dat een indicatie dat het model een realistisch landgebruikpatroon heeft gegenereerd, zoals getoond in hoofdstuk 4. Fractale maten zijn een ander type maten, welke worden toegepast in onderzoek naar complexiteit, om regelmatigheden in (stedelijke) landgebruikpatronen te karakteriseren, zoals een machtswet verdeling in de grootte van stedelijke clusters of de fractale dimensie van stukken stedelijk landgebruik. In dit proefschrift zijn fractale maten gebruikte om gesimuleerde landgebruik patronen te vergelijken met waargenomen landgebruik patronen, zoals getoond in hoofdstuk 5. Daarnaast kan de absolute waarde van fractale maten geïnterpreteerd worden omdat ze een algemene regelmatigheid in stedelijke systemen aangeven, waarvoor waardes bekend zijn in de literatuur. Daarom kunnen fractale maten gebruikt worden om de proces nauwkeurigheid van een synthetische applicatie te beoordelen waarvoor geen geobserveerde waarden bestaan, zoals laten zien in hoofdstuk 6.

Omgevingsregels beschrijven de invloed van het bestaande landgebruik op landgebruikveranderingen. Dit bevat het voortbestaan van bestaand landgebruik, de conversie naar ander landgebruik en de aantrekking of afstoting van landgebruik in de omgeving van een locatie. Omdat omgevingsregels niet direct afgeleid kunnen worden uit meetgegevens, moeten ze ingesteld worden in een kalibratie procedure. Hoofdstuk 4 bekijkt of een grotere omgeving de nauwkeurigheid van landgebruikmodellen verbeterd. De resultaten geven aan dat agenten de omgeving op twee schaalniveaus bekijken: de directe buurt van een locatie heeft een sterke invloed op de aanleg van nieuwe stedelijke gebieden, maar de omgevingsregels over een grotere afstand – ongeveer de afstand tussen verschillende stadsdelen – verbeteren ook de modelresultaten. Hoofdstuk 5 behandelt de kalibratie van omgevingsregels door de over- of ondervertegenwoordiging van landgebruiken in de omgeving van landgebruikveranderingen te meten, en deze te gebruiken om gesimuleerde en werkelijke landgebruikveranderingen te vergelijken. Deze studie laat zien dat deze methode de proces nauwkeurigheid van de modeltoepassingen aanzienlijk verbetert. Hoofdstuk 6 behandelt een speciaal type omgevingsregels: regels welke de invloed van bestaande activiteit op de locatie van veranderingen in activiteit beschrijven. Met activiteit wordt een hoeveelheid of dichtheid bedoeld die gerelateerd is aan een landgebruik, zoals inwoners gerelateerd zijn aan woongebieden. Deze studie toont aan dat relatief simpele regels een realistisch stedenpatroon kan laten groeien, wat wederom bevestigt dat omgevingsregels de proces nauwkeurigheid van een landgebruik model verbeteren.

Het onderzoek dat in dit proefschrift beschreven is, draagt bij aan de kalibratie en validatie van landgebruikmodellen door verschillende methodes te introduceren en gebruiken welke op een objectieve manier de voorspellende nauwkeurigheid en de proces nauwkeurigheid van landgebruik modellen beoordelen. In beide beoordelingen is het belangrijk om impliciet of expliciet een referentiemodel te gebruiken om de resultaten te kunnen beoordelen. Een landgebruikmodel moet ten minste een beter resultaat genereren dan dit referentie model. Tegelijkertijd betekent een beter resultaat dat een referentie model niet automatisch dat een landgebruik model ook acceptabel is, aangezien dat afhangt van het doel van het model of de toepassing. Daarnaast moet vermeld worden dat de methodes die in dit proefschrift gepresenteerd en gebruikt zijn objectief zijn, maar dat de keuze voor een bepaalde methode een subjectieve keuze blijft. Daarnaast is het tot op heden nog niet mogelijk om landgebruikpatronen op een bevredigende manier te beschrijven, daarom blijven meer subjectieve methodes zoals visuele interpretatie van modelresultaten of interpretatie van parameterwaarden van toegevoegde waarde in de kalibratie en validatie van landgebruikmodellen.

Acknowledgements

Doing a PhD, like life, is similar to a game: you'll experience setbacks and find support to overcome these, you'll face problems and meet friends to solve them, you'll arrive at points you never wanted to visit only to discover there are people that help you out. These are the rules:

6 The bridge – move to number 12: After graduation you arrive at your first job, the bridge from study to science. Hedwig, I thank you very much for the challenges and opportunities you provided, it really felt like a leap forward!

19 The tavern – wait one turn: You decide to obtain a PhD, but you're not in a university. Instead you take the long road, spending many evenings and weekends to get the job done. A road that would never lead anywhere without the help of Arnold: your support has been invaluable. This includes your help improving manuscripts and providing organisational support, but certainly also the moral support at moments when the pile of work in front of me was too high too see the other side.

39: The well – wait until someone passes: Waiting is the comfortable activity that allows putting things in perspective, reflecting on the path you've taken, or simply rest to recharge the battery. It feels much, much better when with friends. Tom, Hannes, Corsi, Viet, Wouter, Marije, PJ, Maaike, Jils, Ronald, Bart, Inge, Tom, Wessels, and other Kogelaers, Fletters, and many others, it has been a pleasure to share a drink, discuss things; I hope we will spend many more occasions like these.

42: The maze – move to 39: The road to a PhD is not a straight line; it includes cul-de-sacs, detours and shortcuts. But in every detour you find a colleague to point you back on track, and in every cul-de-sac there's a colleague to put you in the right direction. Roel, Jelle, Christy, Edith, Bernhard, Alex, Roger, Inge, Maarten, Mark, Mohammed, Ron, and Nick I really enjoyed working together. In addition, all my new IVM and AGCI colleagues: thank you for the support in the last stage of my PhD.

52: The prison – wait until someone passes: Well, actually, not just someone, but a very special someone: Maartje, I am so happy to have someone that shares the same interest, ambition and goals. I hope to support you as much as you did support my in the PhD track.

58 Death – go back to the start: Although some think otherwise, this rule doesn't exist in real life; neither does it exist in the PhD game as far as I know.

5, 9, 14, 18, 23, 27, 32, 36, 41, 45, 50, 54, 59: A goose – move the same amount of steps again: Geese are like family, they are everywhere along the road, to help you whenever needed, asked or unasked, expected or unexpected. Mama, Papa, Marieke and Joppe I am truly grateful for the support you provided along the road towards this PhD, but more generally during the game that's life.

63: The end – you win: You are a Dr., and a very happy man. Thanks again for all the help along the way, it was really appreciated.

About the author

Curriculum vitae

Jasper van Vliet was born on July 23, 1980 in Hoorn, The Netherlands. After finishing secondary school (gymnasium) in 1998, he spent one year in Norway at Numedal Folkehøgskole. In 2006 het graduated from Wageningen University as an Agricultural Engineer in Soil, Water, and Atmosphere, and as a Master of Science in Geo-information Science. For his thesis and internship he visited the Dutch Geological Survey (TNO-NITG), the Spatial Analysis and Modelling Laboratory at Simon Fraser University, Canada, and the Department of Geography at Memorial University of Newfoundland, Canada.

After graduation Jasper started working as a scientist and consultant at the Research Institute for Knowledge Systems bv (RIKS) in Maastricht, The Neterlands, where he specialized in the development and application of land-use models for spatial decision support. In addition he has been enrolled as an external PhD candidate at the Laboratory of Geo-information Science and Remote Sensing at Wageningen University since 2009 using the research conducted at RIKS as the basis for his PhD thesis. Since January 2012 Jasper has been employed as a postdoctoral researcher at the Institute for Environmental Studies (IVM) and the Amsterdam Global Change Institute (AGCI) of the VU University Amsterdam, The Netherlands.

Peer-reviewed journal publications

Van Vliet, J., Hagen-Zanker, A., Hurkens, J., and Van Delden, H. (submitted) A fuzzy set approach to assess the predictive accuracy of land-use models. Ecological Modelling.

Van Vliet, J., Naus, N., Van Lammeren, R.J.A., Bregt, A.K., Hurkens, J. and Van Delden, H. (Submitted). Measuring the neighbourhood effect to calibrate land-use models. Computers, Environment and Urban Systems.

Aljoufie, M., Zuidgeest, M., Brussel, M., Van Vliet, J., and van Maarseveen, M. (Submitted) A cellular-automata based land-use and transport interaction model applied to Jeddah, Saudi Arabia. Landscape and Urban Planning

Van Vliet, J., Hurkens, J., White, R., and Van Delden, H. (2012). An activity based cellular automaton model to simulate land use dynamics. Environment and Planning B 39: 198-212.

Van Delden, H., van Vliet, J., Rutledge, D. and Kirkby, M.J. (2011). Scaling issues in integrated models for policy support: lessons learnt and current challenges. Agriculture, Environment and Ecosystems 142 (1-2): 18 – 28.

Van Vliet, J., A.K. Bregt, and Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land-use models. Ecological Modelling 222 (8): 1367-1375.

Van Delden, H., Stuczynski, T., Ciaian, P., Paracchini, M.L., Hurkens, J., Lopatka, A., Gomez, O., Calvo, S., Shi, Y., Van Vliet, J. and Vanhout, R. (2010). Integrated assessment of agricultural policies with dynamic land use change modelling. Ecological Modelling 221 (18): 2153 – 2166.

Van Vliet, J., Dragicevic, S. and White R. (2009). Modelling urban growth using a variable grid cellular automaton. Computers Environment and Urban Systems 33(1): 35 – 43.

Conference proceedings

Van Delden, H., van Vliet, J., Petersen, J-E., Hoogeveen, Y., Elbersen, B., Helming, J., Helming, K., Kuhlman, T., Nowicki, P., Pallière, C., Pérez Domínguez, I.,Pérez-Soba, M., Verburg, P.H., and Woltjer, G. (2012). Exploring land use trends in Europe. Proceedings of the 6th International Congress on Environmental Modelling and Software, July 1-5, 2012, Leipzig, Germany.

Van Vliet, J., and Van Delden, H. (2011) Calibration of the neighbourhood effect of cellular automata land use models. Proceedings of the 17th European Colloquium for Quantitative and Theoretical Geography, September 2-5, 2011, Athens, Greece.

Van Delden, H., McDonald, G., Shi, Y-E., Hurkens, J., Van Vliet, J., Thornton, M., and Van Den Belt, M. (2011). Integrating land use and socio-economics in support of natural resource management. Proceedings of the 17th European

Colloquium for Quantitative and Theoretical Geography, September 2-5 2011, Athens, Greece.

Van Vliet, J., and Van Delden, H., Simulating land-use change in Portugal using an activity-based model (2011). Proceedings of 14th AGILE conference, April 18-21, 2011, Utrecht, The Netherlands.

Van Delden, H., McDonald, G., Shi, Y. Hurkens, J., Van Vliet. J. and Van den Belt, M. (2011). Integrating socio-economic and land-use models to support urban and regional planning. Proceedings of the 14th AGILE conference, April 18-21, 2011, Utrecht, The Netherlands.

Van Delden, H., Van Vliet, J., Navarro, C. and Gutierrez, E. (2010). The Xplorah SDSS – Supporting integrated planning on the island of Puerto Rico. Proceedings of the 10th International Conference on Design & Decision Support systems in Architecture and Urban Planning, July 19-22, 2010, Eindhoven, the Netherlands.

Van Vliet, J., Navarro, C., Gutiérrez, E.R., Hurkens, J., and van Delden, H. (2009). Xplorah Municipio, a multi resolution spatial support system. Proceedings of the 12th AGILE conference, June 2-5, 2009, Hannover, Germany.

Van Vliet, J. (2009). Assessing the accuracy of changes in spatial explicit land-use change models. Proceedings of the 12th AGILE conference, June 2-5, 2009, Hannover, Germany.

Alves, C., de Rosario, L., Van Vliet, J.. and van Delden, H. (2008). Usability assessment of an integrated land use model in Portugal. Proceedings of the 4th International Congress on Environmental Modelling and Software, July 7-10, 2008, Barcelona, Spain.

Van Delden, H., Gutiérrez, E.R., Van Vliet, J., and Hurkens, J. (2008). Xplorah, a multi scale integrated land use model. Proceedings of the 4th International Congress on Environmental Modelling and Software, July 7-10, 2008, Barcelona, Spain.

Van Vliet, J. and van Delden, H. (2008) An activity based cellular automaton model to simulate land use changes. Proceedings of the 4th International Congress on Environmental Modelling and Software, July 7-10, 2008, Barcelona, Spain.

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

The C.T. De Wit Graduate School PE&RC PRODUCTION ECOLOGY & RESOURCE CONSERVATION

Review of literature (6 ECTS)

- Calibration and validation of spatial explicit land-use change models (2011)

Writing of project proposal (4.5 ECTS)

- Calibration and validation of spatial explicit land-use change models (2011)

Post-graduate courses (3 ECTS)

- Agent based models for spatial systems in social sciences & economic science with heterogeneous interacting agents; S4 / ESHIA (2007)
- Uncertainty propagation in spatial and environmental modelling; PE&RC (2011)

Invited review of (unpublished) journal manuscripts (8 ECTS)

- Computes, Environment and Urban Systems (2); International Journal for Geographical Information Science; Environmental Modelling and Software (2); Journal of Environmental Informatics; Environmental Modelling and Assessment; Journal of Applied Earth Observation and GeoInformation: several topics (2009-2012)

Competence strengthening / skills courses (2.7 ECTS)

- Training how-to-write FP7 proposals (cooperation); EG-Liaison (2010)
- Writing a scientific article; Talencentrum VU (2012)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.2 ECTS)

- PE&RC Weekend (2011)
- GRS PhD Day (substituting the PE&RC day) (2012)

Discussion groups / local seminars / other scientific meetings (4.5 ECTS)

- Irregular meetings and discussions with visiting scientist at RIKS (2007-2011)
- RegioResources – a cross disciplinary dialogue on sustainable development of regional resources; Dresden, Germany (2011)

International symposia, workshops and conferences (18.2 ECTS)

- International Congress on Environmental Modelling and Software; Barcelona, Spain (2008)
- International Conference on Impact Assessment of Land-Use Changes; Berlin, Germany (2008)
- 12th Agile Conference; Hannover, Germany (2009)
- 10th International Conference on Design & Decision Support Systems in Architecture and Urban Planning; Eindhoven, the Netherlands (2010)
- International Conference on Managing the Urban Rural Interface; Copenhagen, Denmark (2010)
- 17th European Colloquium for Quantitative and Theoretical Geography; Athens, Greece (2011)
- 14th AGILE Conference; Utrecht, the Netherlands (2011)

Lecturing / supervision of practical's / tutorials (15 ECTS)

- Cellar automata land-use modelling. ITC, Department of urban and regional planning and geo-information management (2009, 2010, 2011)
- Guest lectures at Maastricht University (2007), Wageningen University (2008), University of Puerto Rico (2009), Technical University Delft (2010), and Radboud University (2011)
- Various workshops for professional planners: Puerto Rico Planning Board; ESA scientists; Planners and professionals in the Tisa Catchment Area (2008-2011)
- Several introduction and advanced calibration courses for the Metronamica land-use modelling system (2008-2011)

Supervision of 2 MSc student (6 ECTS)

- Usefulness and usability of a spatial decision support system – A case study with Scendes in Portugal. Carin Alves, CIHEAM, 2008.
- Understanding the neighbourhood effects of land-use change to improve the calibration procedure of a CA based land-use model. Nick Naus, WUR, 2009.