

Chapter 8

Markov–Cellular Automata in Geospatial Analysis

Courage Kamusoko

8.1 Introduction

Spatial simulation models are indispensable for modeling land use/cover changes (Wu and Webster 1998; Messina and Walsh 2001; Soares-Filho et al. 2002), deforestation and land degradation (Lambin 1994; Lambin 1997; Etter et al. 2006; Moreno et al. 2007), urban growth (Clarke et al. 1997; Couclelis 1989; Cheng and Masser 2004; Gar-On Yeh and Li 2009), climate change (Dale 1997) and hydrology (Matheussen et al. 2000). For land use/cover change studies, spatial simulation models are critical for understanding the driving forces of change, as well as to produce “what if” scenarios that can be used to gain insights into future land use/cover changes (Pijanowski et al. 2002; Eastman et al. 2005; Torrens 2006). Recently, the knowledge domain of spatial simulation modeling has advanced owing to the rapid developments in computer technology, coupled with the decrease in the cost of computer hardware. In addition, developments in geospatial, natural and social sciences concerning bottom-up, dynamic and flexible self-organizing modeling systems, complemented by theories that emphasize the way in which decisions made locally give rise to global patterns, have enriched spatial simulation models (Tobler 1979; Wolfram 1984; Couclelis 1985; Engelen 1988; Wu and Webster 1998; Batty 1998). To date, numerous spatial simulation models have been developed and applied, particularly for land use/cover modeling (Clarke et al. 1997; Kaimowitz and Angelsen 1998; Messina and Walsh 2001; Soares-Filho et al. 2002; Walsh et al. 2006).

C. Kamusoko (✉)
Overseas Operations Division, Asia Air Survey Co. Ltd.,
Tokyo, Japan

Formerly in Integrative Environmental Sciences Program, University of Tsukuba,
Tsukuba, Japan
e-mail: cou.kamusoko@ajiko.co.jp

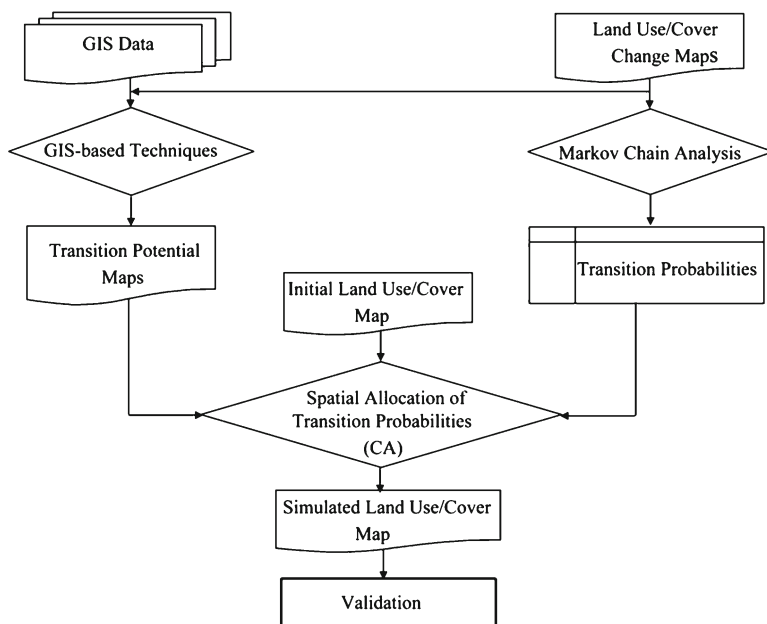


Fig. 8.1 Conceptual framework of the Markov–cellular automata (MCA) model

While a literature review reveals a plethora of spatial simulation models based on different modeling techniques and traditions (Parker et al. 2003; Verburg et al. 2004), in this chapter we focus on the Markov–cellular automata (MCA) model that integrates cellular automata (CA) procedures, Markov chains and geographical information science (GIS)-based techniques such as weight of evidence (WofE) and multi-criteria evaluation (MCE) (Fig. 8.1). The objective of this chapter is to review the methodological developments of the MCA model. The chapter is organised into five sections. Section 8.2 focuses briefly on the conceptual framework of the MCA model, paying special attention to the basics of Markov chains, GIS-based techniques such as WofE and MCE, and CA models. The application of MCA models in previous studies is described in Sect. 8.3, while Sect. 8.4 focuses on the current status and future prospects of the MCA modeling framework. Finally, Sect. 8.5 gives a summary and conclusions.

8.2 Conceptual Framework of the MCA Model

A MCA model is a spatial model for simulating land use/cover changes in landscapes where land use/cover is viewed as a mosaic of discrete states, and changes are multi-directional (e.g. forest to non-forest or vice versa) (Silverton et al. 1992; Li and Reynolds 1997). In order to gain insights into land use/cover changes in a

given landscape, modeling approaches that can adequately represent state-and-transition systems should be used. The MCA model that combines CA with Markov chain analysis and GIS-based techniques (Fig. 8.1) can be used for modeling land use/cover changes since it can effectively represent state-and-transition systems. The Markov chain process uses transition probabilities to control temporal dynamics among the land use/cover classes. Spatial dynamics are controlled by local rules determined either by the CA mechanism (neighborhood configuration) or by its association with the transition potential maps computed from WofE and MCE techniques. The MCA model allows the transition probabilities of one pixel to be a function of neighboring pixels because the CA model consists of a regular grid of cells, each of which can be in one of a finite number of possible states which are updated synchronously in discrete time steps according to a local interaction rule (Messina and Walsh 2001). The transition probabilities of the CA model depend on the state of a cell, the state of its surrounding cells, and the weights associated with the neighborhood context of the cell (White and Engelen 1997). The MCA model (Li and Reynolds 1997) can be expressed as

$$\begin{aligned}
 C(i, j) &= m \rightarrow k, \text{ if } R > P_{m,k} \cdot N_k / 4 \\
 &\text{No change,} \quad \text{if } R \leq P_{m,k} \cdot N_k / 4
 \end{aligned}
 \tag{8.1}$$

where $C(i, j)$ is the use/cover class of cell (i, j) , R is a random number with a uniform distribution, $P_{m,k}$ is the transition probability from one land use/cover class m to k , and N_k is the number of neighboring cells of land use/cover k , which includes the evaluation score of land use/cover transition potential at location i, j . The weight of four is used in (8.1) because each cell is assumed to have four neighbors.

Land use/cover change modeling approaches such as MCA generally consist of three major components: (1) a change demand submodel, (2) a transition potential submodel, and (3) a change allocation submodel (Eastman et al. 2005). The change demand submodel estimates the rate of change between two land use/cover maps from different periods. The results are summarised in a transition probability matrix that expresses the rate of conversion from one land use/cover class to another (Table 8.1). The transition potential submodel determines the likelihood (which can also be expressed as suitability or probability) that land would change from one land use/cover class to another based on biophysical and socio-economic factors (Fig. 8.2). Specifically, it establishes the degree to which locations might potentially change in a future period of time (Eastman et al. 2005). Finally, the change allocation submodel is concerned with the decisions by which specific areas will change, given

Table 8.1 Example of land use/cover change transition probabilities (1993–2000)

| | | 2000 | | |
|------|------------------|----------------|------------------|-------------|
| | | Current forest | Unstocked forest | Non-forest |
| 1993 | Current forest | 0.85 | 0.13 | 0.02 |
| | Unstocked forest | 0.01 | 0.94 | 0.05 |
| | Non-forest | 0.05 | 0.65 | 0.30 |

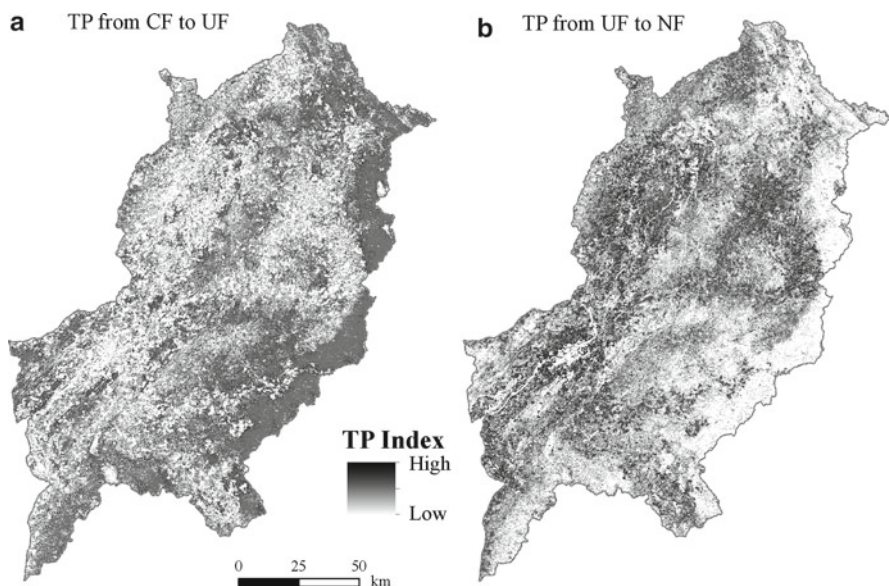


Fig. 8.2 Example of transition potential (TP) maps of Luangprabang province, Lao Peoples' Democratic Republic: (a) current forest (CF) to unstocked forest (UF), and (b) unstocked forest (UF) to non-forest (NF)

the demand and potential surfaces. In the MCA modeling approach, the change demand model is represented by the Markov chains, while the transition potential and change allocation submodels are represented by GIS-based techniques such as the WofE and MCE models and the CA models, respectively.

In the following subsections, we focus briefly on (1) the computation of transition probabilities using Markov chains, (2) the computation of land use/cover transition potential maps based on WofE and MCE techniques, (3) the spatial allocation of simulated land use/cover probabilities based on a CA model, and (4) the mechanism of the MCA model.

8.2.1 Markov Chain Modeling of Land Use/Cover Changes

Markov chains have been widely used to model land use/cover changes (Drewett 1969; Bell 1975; Bell and Hinojosa 1977; Robinson 1978; Jahan 1986; Muller and Middleton 1994; Wood et al. 2004). A Markov chain is a stochastic model based on transition probabilities which describes a process that moves in a sequence of steps through a set of states (Wu et al. 2006). In essence, the probability that the system will be in a given state at a given time (t_2) is derived from the knowledge of

its state at any earlier time (t_1), and does not depend on the history of the system before time t_1 (Petit et al. 2001). This is known as a first-order Markov chain process. The Markov chain can be characterized as stationary or homogeneous in time if the transition probabilities depend only on the time-interval t (i.e. $\Delta t = t_2 - t_1$), and if the time period at which the process is examined is of no relevance (Karlin and Taylor 1975). To model land use/cover change using the stationary and first-order Markov chain, the land use/cover distribution at t_2 is calculated from the initial land use/cover distribution at t_1 based on the transition matrix (Lambin 1994; Petit et al. 2001). The Markov chains can be expressed as

$$v_{t_2} = M * v_{t_1} \quad (8.2)$$

where v_{t_2} is the output land use/cover proportion column vector, v_{t_1} is the input land use/cover class proportion column vector, and M is an $m * m$ transition matrix for the time interval $\Delta t = t_2 - t_1$.

In order to model land use/cover changes using Markov chains, it is essential to understand their basic assumptions and limitations. First, land use/cover changes are considered as a stochastic process where the transition probabilities are stationary (homogeneity) and land use/cover classes are in different states of the Markov chain (Wu et al. 2006). However, it is difficult to expect stationarity in transition probabilities because land use/cover changes are the result of the complex dynamics of socio-economic, political and biophysical factors that change over time (Lambin et al. 2000). While the departure from the simple assumptions of stationary, first-order Markov chains is conceptually possible, analytical and computational difficulties emerge. Nonetheless, it might be practical to assume transition probabilities to be stationary if the time span is not too long (Weng 2002). Second, the Markov chains that handle stationary processes are not appropriate for incorporating human activities (Boerner et al. 1996; Weng 2002). Third, the available land use/cover data may be insufficient to estimate reliable transition probabilities (Pastor et al. 1993), particularly in landscapes experiencing rapid land use/cover changes (Wood et al. 2004). Finally, a stochastic Markov chain model does not consider spatial knowledge within each land use/cover class (Boerner et al. 1996).

While the Markov chains have some limitations, they are relatively easy to derive (or infer) from land use/cover data (Wood et al. 2004). Despite the fact that the Markov chains do not reveal the underlying land use/cover change processes, they give the direction and magnitude of change, and that is potentially of use for simulating land use/cover changes (Weng 2002). In addition, the computational requirements of Markov chain models are quite modest.

Table 8.1 shows the forest cover transition probabilities between 1993 and 2000, calculated on the basis of the frequency distribution of the observations. The diagonal of the transition probability matrix represents the self-replacement probabilities, i.e. the probability of a forest cover class remaining the same (shown in bold in Table 8.1), whereas the off-diagonal values indicate the probability of a change occurring from one forest cover class to another.

8.2.2 *Computation of Transition Potential Maps Using GIS-Based Techniques*

Transition potential (suitability) maps represent the likelihood or probability that the landscape will change from one land use/cover class to another (e.g. forest to non-forest). The basic prerequisite for computing transition potential maps is the derivation of weights representing the relative importance of each factor in relationship to a given land use/cover change (Hosseinali and Alesheikh 2008). Generally, weighting methods are classified into data-driven (e.g. WofE, logistic regression and artificial neural networks) and knowledge-driven (analytical hierarchy processes, ratio estimation etc.) groups (Hosseinali and Alesheikh 2008; Liu and Mason 2009). Although both data-driven and knowledge-driven methods have been used for computing transition potential maps (Eastman et al. 2005), the former has the advantage of reducing the problems of biased or incorrect decisions that knowledge-driven methods have (Hosseinali and Alesheikh 2008). While there are many data- and knowledge-driven techniques (Liu and Mason 2009), in this chapter we limit our discussion to the WofE and MCE techniques.

The WofE algorithm uses Bayes' theorem of conditional probability to compute transition potential maps based on the statistical relationship between each land use/cover change (e.g. forest to non-forest) and predictors (i.e. independent variables) such as distance to roads, soil and elevation. This method employs prior and posterior probabilities. The prior probability is defined as the probability of occurrence of a specific land use/cover change, which is calculated by dividing the number of samples (of that land use/cover change) with the total land use/cover change in the study area. The posterior probability is the conditional probability of the existence of a specific land use/cover change (e.g. forest to non-forest) when a predictor variable exists. For example, the conditional probability of the change from forest to non-forest given the presence of predictor variables such as distance to roads, soil and elevation can be expressed as (Almeida et al. 2005; Levine and Block 2011).

$$P\{R|S\} = P\{S|R\} * P\{R\} / P\{S\} \quad (8.3)$$

where $P\{R|S\}$ is the conditional (posterior) probability of the change from forest to non-forest given the presence of the predictor variables, $P\{S|R\}$ is a likelihood function that gives the probability that the predictor variables (data) would be obtained given that R is true, $P\{R\}$ is the prior probability, $P\{S\}$ is the marginal probability of the predictor variables (that is the probability of obtaining the predictor variables under all possible scenarios), R is the change from forest to non-forest, and S is the predictor variables (e.g. distance to roads, soil and slope).

The advantage of WofE is its simplicity and straightforward interpretation of weights (Agterberg and Cheng 2002). However, the basic assumption of the WofE is that predictor variables should be independent. Therefore, the predictor variables should be tested for independence using methods such as the Crammer coefficient (Bonham-Carter et al. 1988; Bonham-Carter 1994; Agterberg and Cheng 2002). Generally, a predictor variable with a Crammer coefficient of more than 0.5 should

be removed since it would be highly correlated with other variables (Bonham-Carter et al. 1988; Bonham-Carter 1994).

MCE is a technique for combining data according to its importance in making a decision (Liu and Mason 2009). Many researchers have integrated MCE and GIS (Carver 1991; Jankowski and Richard 1994; Jankowski 1995; Eastman et al. 1995; Wu and Webster 1998). Conceptually, the MCE technique involves qualitative or quantitative weighting, scoring or ranking of criteria to reflect their importance to either single or multiple sets of objectives (Eastman et al. 1995). In essence, the MCE technique uses numerical algorithms that define the “suitability” of a particular solution on the basis of the input criteria and weights, together with some mathematical and logical means of determining trade-offs when conflicts arise (Heywood et al. 1998). Two of the most common procedures for MCE are weighted linear combinations and concordance–discordance analysis (Voogd 1993; Carver 1991). In the former, each factor is multiplied by a weight and then summed to arrive at a final transition potential index. In the latter, each pair of alternatives is analyzed for the degree to which one out-ranks the other in the specified criteria (Eastman et al. 1995). The concordance–discordance analysis is computationally impractical when a large number of alternatives is present (e.g. raster data where every pixel is an alternative), while the weighted linear combination is very straightforward in a raster GIS.

The weighted linear combination (Voogd 1993) combines factors by applying a weight to each factor, followed by a summation of the results to yield a transition potential (suitability) map, i.e.

$$S = \text{SUM}(w_i * x_i) \quad (8.4)$$

where S is suitability (transition potential), w_i is the weight of factor i , and x_i is the criterion score of factor i .

In a case where constraints apply, the procedure can be modified by multiplying the suitability calculated from the factors by the product of the constraints, i.e.

$$S = \text{SUM}(w_i * x_i) * \Pi c_j \quad (8.5)$$

Where c_j is the criterion score of constraint j , and Π is the product.

Most GIS software, such as IDRISI, provides a MCE module developed to compute transition potential maps (Eastman et al. 1995). The primary issues in the computation of transition potential (suitability) maps are the standardization of criteria scores and the development of the factor weights using methods such as the analytic hierarchy process (Saaty 1977; Saaty and Vargas 2001).

8.2.3 Cellular Automata (CA) Models

Cellular automata (CA) are bottom-up, individual-based dynamic models that were originally conceptualized by Ulam and von Neumann in the 1940s in order to understand the behavior of complex systems (Moreno et al. 2010). The CA model consists

of an array of cells wherein each cell can assume one of i discrete states at any one time (Tobler 1979; Couclelis 1985; White and Engelen 1997). Time progresses in discrete steps, and all cells change their state simultaneously as a function of their own state, together with the state of the cells in their neighborhood, in accordance with a specified set of transition rules (Engelen et al. 1995). In essence, CA encompasses five major components (Wolfram 1984; White and Engelen 1997):

1. A space composed of a regular grid in one or two dimensions
2. A finite set of possible states associated with every cell (e.g. forest or non-forest)
3. A neighborhood composed of adjacent cells whose states influence the central cell
4. Transition rules applied uniformly through time and space
5. A discrete time at which the state of the system is updated

According to Moreno et al. (2010), circular and extended neighborhoods are commonly used to reduce directional bias and capture the spatial influence of surrounding cells on the central one. Space is typically represented as a grid of regular cells, while the neighborhood is defined as a collection of cells based on physical adjacency (White and Engelen 1997; Moreno et al. 2010). Distance functions are applied within a neighborhood to take into account the spatial-dependent attractiveness or repulsiveness of one cell state over another (Soares-Filho et al. 2002; Moreno et al. 2010). In addition to deterministic transition rules, stochastic rules are commonly applied to capture the intrinsic variability of natural and human systems (Moreno et al. 2010). The CA model works by simulating the present based on an extrapolation of the past land use/cover maps. This allows the model to iterate to any other selected date (Messina and Walsh 2001).

For land use/cover changes and urban growth studies, CA models have been found to be more effective than the conventional modeling approaches for a number of reasons. First, CA models allow the integration of macro-scale with micro-scale temporal processes as well as the integration of macro-spatial and micro-spatial phenomena (Wolfram 1984; Torrens 2000). As a result, CA models can make the maximum possible use of available spatial and temporal detail, in contrast to conventional approaches, which operate either at the macro- or the micro-level (Briassoulis 2000). Furthermore, CA models offer a flexible platform for the interaction of biophysical and socio-economic driving factors, as well as for the simulation of real-world complex systems based on simple rules (Wolfram 1984; Engelen 1988; White and Engelen 1997). More importantly, theoretical assumptions may be tested and validated in a particular environmental and socio-economic context (Briassoulis 2000; Torrens 2000). While CA models have produced important contributions to modeling, recent studies have revealed that raster-based CA are sensitive to the modifiable spatial units used in the model, and the modeling results vary according to the cell size and neighborhood configuration (Moreno et al. 2010).

8.2.4 How the MCA Model Works

The spatially explicit nature of the CA model and its compatibility with GIS and other modeling frameworks such as Markov chains have resulted in the development of various hybrid CA models (Walsh et al. 2006). This chapter focuses only on the mechanism of the MCA model based on the Dinamica EGO (environment for geo-processing objects) platform. Dinamica EGO was developed by the Center for Remote Sensing of the Federal University of Minas Gerais in Brazil (Maeda et al. 2011). This MCA model employs an *expander* transition function to expand or contract previous land use/cover class patches, while the *patcher* transition function is used to form new patches through a stochastic seeding mechanism (Soares-Filho et al. 2002). Thus, on the one hand the *expander* transition function performs transitions from state i to state j only in the neighboring cells of state j . On the other hand the *patcher* transition function performs transitions from state i to state j only in the neighboring cells of states other than j (Almeida et al. 2003). First, the algorithm scans the initial land use/cover map to sort out the cells with the highest probabilities and then arrange them in a data array (Almeida et al. 2005). Then the cells are selected randomly from top to bottom of the data array. Finally, the land use/cover map is again scanned to perform the selected transitions (Soares-Filho et al. 2002). If the *expander* transition function does not perform the amount of desired transitions after a fixed number of iterations, it then transfers to the *patcher* transition function a residual number of transitions, so that the total number of transitions always amounts to a desired value (Soares-Filho et al. 2002). The desired transitions are obtained from Markov chain-computed transition probabilities. However, the *patcher* transition function will simulate land use/cover change patterns by generating diffused patches, while at the same time preventing the formation of single isolated one-cell patches (Almeida et al. 2003). This function searches for cells around a chosen location for a given transition through the selection of the core cell of the new patch based on a specific number of cells around the core cell according to their transition probabilities (Soares-Filho et al. 2002).

The *expander* and *patcher* transition functions are composed of an allocation mechanism responsible for identifying cells with the highest transition probabilities for each ij transition. As a result, cells are stored and organized for later selection. The two complementary functions (i.e. the *expander* and the *patcher*) consist of mean patch size, patch size variance and isometry parameters, which can be changed to produce various spatial patterns of land use/cover patches according to a log-normal probability distribution function (Soares-Filho et al. 2002). For example, an increase in mean patch size results in a less fragmented landscape, while an increase in the patch size variance results in a more diverse landscape (UFMG 2009). Isometry is a number that varies from 0 to 2, and thus an isometry greater than one results in more isometric (equal) patches (UFMG 2009). Finally, MCA model iterations are specified according to time differences between two land use/cover maps ($\Delta t = t_2 - t_1$).

8.3 Application of MCA Models in Previous Studies

Modeling approaches that integrate CA and Markov chains have been explored for some time (Zhou and Liebhold 1995; Li and Reynolds 1997; Parker et al. 2003; Aspinall 1994). A major advantage of the MCA approach is that GIS and remote sensing data can be incorporated effectively (Li and Reynolds 1997). In particular, biophysical and socio-economic data can be used to define initial conditions, to parameterize the MCA model, to calculate transition probabilities and to determine the neighborhood rules with transition potential maps.

Although the potential of the MCA models has been recognized, few studies have used the MCA models for simulating land use/cover changes. Li and Reynolds (1997) developed a combined Markov and CA model to simulate the effects of spatial pattern, drought and grazing on the rates of rangeland degradation. Although their model was conceptually appealing, it did not account for the variations of transition probabilities due to changes in environmental, socio-economic and political factors. To overcome such limitations, Soares-Filho et al. (2002) incorporated a *saturation value* parameter that is designed to vary the transition rates through a dynamic feedback analysis of landscape changes. Their spatially explicit, multi-scale and dynamic stochastic CA modeling framework successfully simulated land use/cover changes in the Amazonian colonization frontier (Soares-Filho et al. 2002; Soares-Filho et al. 2006). Recently, the Dinamica EGO modeling framework has also introduced a scenario generator model that computes transition rates based on the integration of environmental and socio-economic factors (Almeida et al. 2005; Teixeira et al. 2009).

Pontius and Malanson (2005) applied the MCA to predict land use/cover changes in central Massachusetts. Their model used an MCE technique to compute transition potential maps, and a spatial contiguity rule to determine the location of predicted change. Contemporary legal constraint data were used as an additional driver to calibrate the transition potential (suitability) maps (Pontius and Malanson 2005). Although the MCA model produced good results, it did not incorporate additional constraints and factors that represent socio-economic and urban planning issues. Furthermore, the authors concluded that their MCA model was poor at predicting the location of built to non-built conversions (Pontius and Malanson 2005). Paegelow and Olmedo (2005) also used the MCA model for testing the possibilities and limits of a prospective land cover modeling in France and Spain. Their model used the Markov chain analysis to control temporal dynamics, while MCE, multi-objective evaluation and CA controlled spatial contiguity in order to determine the location of the predicted land cover change. Land cover maps and relevant environmental factors were used to calibrate the transition potential (suitability) maps. While the authors reported an overall accuracy of 75%, they noted the need to analyze prediction residues in order to improve the model.

Myint and Wang (2006) also applied the MCA for projecting land use/cover changes in Norman, Oklahoma, USA. Their model also used an MCE technique to compute transition potential maps, and a spatial contiguity rule to determine the location of predicted change. Ancillary map layers such as roads and drainage were

used as driving factors in order to calibrate the transition potential (suitability) maps. The suitability ratings were based on the authors' personal judgment in consultation with land use planners, which may possibly lead to bias (Hosseinali and Alesheikh 2008). Although their model was effective at projecting future land use/cover changes, as indicated by an overall accuracy of 86.2%, their accuracy assessment procedure only considered accuracy in terms of quantity and not in terms of location (Pontius and Malanson 2005). More recently, Kamusoko et al. (2009) applied a MCA model in rural areas in Zimbabwe. Their model's overall simulation success was 69% for the 2000 simulated land use/cover map, and 83% for the 2005 simulated land use/cover map. However, the authors reported that the model was poor at simulating the location of bare land areas owing to the lack of input spatial data.

8.4 Current Status and Future Prospects

The increasing awareness of the impact of land use/cover changes on global climate change has renewed interest in the application of spatial simulation models (Soares-Filho et al. 2006; Brown et al. 2007). For example, initiatives that are currently being negotiated under the United Nations Framework Convention on Climate Change (UNFCCC) to reduce emissions from deforestation and forest degradation in developing countries requires the development of robust baseline or reference scenarios under the business-as-usual (BAU) scenario (Angelsen et al. 2009). A baseline or reference scenario (under BAU) is the projected deforestation and associated emissions in the absence of a REDD (reducing emissions from deforestation and forest degradation) project (Angelsen 2008). Several approaches for setting baseline or reference scenarios have been suggested, which include among others spatial and non-spatial modeling approaches (Brown et al. 2007; Terrestrial Carbon Group 2008). However, this new interest in spatial simulation models also presents new challenges to researchers and decision makers because the establishment of robust baseline or reference scenarios requires a better understanding of the underlying driving forces in order to capture intrinsic landscape processes at multiple spatial and temporal scales (GOFC-GOLD 2010). In addition, attention should also be focused on new theoretical and methodological developments in the modeling framework. This section highlights the current status and future prospects of MCA models, paying special attention to issues pertaining to (1) theories underpinning model development, (2) data issues, and (3) calibration and validation.

8.4.1 Theories Underpinning Model Development

Current spatial simulation models of land use/cover changes can be broadly divided into those which are based on theory and those which are not (Verburg et al. 2004). The former include mainly economic theory-based models as well as spatial interaction models (Lambin et al. 2000; Irwin and Geoghegan 2001; Verburg et al. 2004;

Soares-Filho et al. 2006), while the latter comprise models that do not include theory explicitly, or those that are based on specific theoretical assumptions (Myint and Wang 2006; Kamosoko et al. 2009). Although theory is critical during model specification and interpretation, the influence of theory and assumptions on the modeling results is not always examined (Verburg et al. 2004). This unfortunately limits the reliability and robustness of the model (Briassoulis 2000). To overcome this limitation, future MCA models will need to incorporate a strong theoretical background which is relevant to the given underlying landscape processes. This requires more collective efforts that focus on developing an integrated and multi-disciplinary research paradigm. The land use/cover change modeling community has been working on a number of multi-disciplinary research programs aimed at improving spatial models (Geoghegan et al. 1998; Irwin and Geoghegan 2001).

8.4.2 Data and Scale Issues

Fundamental to the development of robust MCA models are issues such as the spatial and temporal dimensions, reliability, availability and cost of data collection (Briassoulis 2000). In most cases, the spatial units usually follow administrative boundaries, which, although appropriate for policy implementation, may not be meaningful for all types of data (Verburg et al. 2004). With respect to the temporal dimension, the temporal systems of reference (e.g. time and number of observations) are not always compatible and consistent (Verburg et al. 2004). In other words, different definitions among time periods, especially at lower levels of aggregation, give rise to problems of compatibility and consistency, particularly with historical data. For example, the dates of historical land use/cover maps may not be compatible with the available socio-economic data, which may have been acquired at a different time. Furthermore, MCA models are built on the assumptions of temporal homogeneity and progressive linear trends, despite the fact that land use/cover changes have occurred in the context of long-term instability (e.g. fluctuations in climate, prices or state policies). These issues are important for land use/cover models where the exact time and length of the policy intervention is critical in the modeling framework. In addition, the availability and cost of obtaining proper longitudinal data (e.g. socio-economic data) limit the reliability of land use/cover change models that integrate biophysical and socio-economic data.

With reference to the spatial dimension, past studies have revealed that raster-based CA models are sensitive to the modifiable spatial units used in the model, and that results vary according to the cell size and neighborhood configuration (Veldkamp et al. 2001; Chen and Mynett 2003; Jantz and Goetz 2005). To overcome the sensitivity of raster-based CA models to cell size and neighborhood configurations, novel geographic object-based CA models have been developed (Torrens and Benenson 2005; Moreno et al. 2010). According to Moreno et al. (2010), space is defined as a collection of geographic objects of irregular shape and size corresponding to meaningful real-world features. Furthermore, the neighborhood is dynamic (i.e. it includes the whole geographic space), and the model allows

the geometric transformation of each object according to a transition function that incorporates the influence of its neighbors (Moreno et al. 2010).

8.4.3 Calibration and Validation

Calibration and validation are important components in the development of MCA models. However, validation is the weakest part of land use/cover modeling, since there are no agreed criteria to assess the performance of one land use/cover model versus another, or to compare one run versus another run of the same model (Pontius et al. 2004). In order to assess the model's predictive power, a clear distinction between the procedures for calibration and validation must be made, the failure of which makes the interpretation of any results difficult or misleading (Pontius et al. 2004). In some cases, it is more common to force the prediction to simulate the correct quantity of each land use/cover class, than to assess whether the model predicts the correct location of land use/cover (Kok et al. 2001; Pontius et al. 2001). Any lack of clarity in the methodology to distinguish the calibration information from the validation information causes confusion in land use/cover modeling, which can lead to a misunderstanding of the model's certainty (Pontius et al. 2004).

Calibration is “the estimation and adjustment of model parameters and constraints to improve the agreement between model output and a data set” (Rykiel 1996). The information used for calibration should be at or before some specific point in time (t_1), which is the point in time at which the predictive extrapolation begins. In contrast, validation is the process of comparing the model's prediction for t_2 with a reference map of time t_2 , where the reference map is considered to be a much more accurate portrayal of the landscape at time t_2 (Pontius and Malanson 2005). One set of data should be used to calibrate the model, and a separate set should be used to validate the model (Pontius et al. 2004). In order to enhance the validity of land use/cover modeling, Pontius et al. (2004) suggested that it is helpful to use a validation technique that, (a) takes into account the source of error, (b) compares the model to a null model (a model that predicts pure persistence, i.e. no change between t_1 and t_2), and (c) performs analysis at multiple scales.

8.5 Summary and Conclusions

This chapter has attempted to review the current state-of-the-art operational MCA land use/cover change models. Despite the existence of the many land use/cover change modeling challenges highlighted in this chapter, the land use/cover modeling research community has developed a variety of models, which have been applied with varying success in different regions of the world. Interesting data sets, as well as the functioning of interdisciplinary and multi-disciplinary research teams, have made efforts to improve and develop robust land use/cover change models that can be useful for understanding the functioning of land use/cover systems, and also

to support land use planning and policy (Lambin et al. 2000; Irwin and Geoghegan 2001; Rindfuss et al. 2003; Verburg et al. 2004). However, the review has also exposed the limitations of the current MCA models and modeling practice. Many current MCA land use/cover models are still built on common assumptions of homogeneity and linear trends, which may fail to capture the underlying real-world landscape processes characterized by non-linear trends. While much effort has been spent on model calibration, little attention has been given to the development of robust validation methods (Pontius and Malanson 2005).

Nonetheless, the limitations singled out present opportunities for research into MCA models. Future MCA land use/cover change models will need to be more integrated and more responsive to different environmental, socio-economic and political conditions (Verburg et al. 2004). Given the rapid developments in computer technology (increases in memory and speed of computers), more integrated MCA models should be developed. For example, encouraging research is being done in the area of geographic object-based CA models (Torrens and Benenson 2005; Moreno et al. 2010). These novel geographic object-based CA models should be incorporated in the MCA modeling framework. More research should also be done in developing non-linear Markov chains that can compute non-linear transition probabilities. In addition, research will also have to address the problems of the evaluation of policy impacts as well as issues of household decision-making. Predominantly, aggregate modeling techniques need to be complemented by agent-based methods capable of measuring the influence of individuals and communities on land use/cover changes. The feasibility of such research would be greatly enhanced by the availability of the detailed land use/cover, biophysical and disaggregate socio-economic data required for integrated agent-based MCA models (Berger 2001).

Finally, more efforts should be made to disseminate land use/cover models in general, and MCA models in particular, by including institutions and individuals, particularly in developing countries. This must be supported by the development of user-friendly modeling software such as Dinamica EGO (Soares-Filho et al. 2002) and IDRISI Taiga (Eastman 2009). Although, spatial simulation models have been criticized for failing to adapt to new challenges and problems, researchers and decision makers are collaborating in order to develop robust MCA land use/cover change models (Rindfuss et al. 2003). These models would be useful for understanding the driving forces and underlying processes of land use/cover changes, as well as to simulate future land use/cover changes.

References

- Agterberg FP, Cheng Q (2002) Conditional independence test for weights-of-evidence modeling. *Nat Resour Res* 11(4):249–255
- Almeida CM, Batty M, Monteiro AMV, Camara G, Soares-Filho BS, Cerqueira GC et al (2003) Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. *Comput Environ Urban Syst* 27:481–509

- Almeida CM, Monteiro AMV, Camara G, Soares-Filho BS, Cerqueira GC, Pennachin CL, Batty M (2005) GIS and remote sensing as tools for the simulation of urban land-use change. *Int J Rem Sens* 26(4):759–774
- Angelsen A (2008) REDD models and baselines. *Int For Rev* 10(3):465–475
- Angelsen A, Brockhaus M, Kanninen M, Sills E, Sunderlin WD, Wertz-Kanounnikoff S (eds) (2009) *Realising REDD+: national strategy and policy options*. CIFOR, Bogor
- Aspinall R (1994) Use of GIS for interpreting land-use policy and modeling effects of land use change. In: Haines-Young R, Green DR, Cousins S (eds) *Landscape ecology and geographic information systems*. Taylor and Francis, London, pp 223–236
- Batty M (1998) Urban evolution on the desktop: simulation with the use of extended cellular automata. *Environ Plann Plann Des* 30:1943–1967
- Bell EJ (1975) Stochastic analysis of urban development. *Environ Plann* 7:35–39
- Bell EJ, Hinojosa RC (1977) Markov analysis of land use change: continuous time and stationary processes. *Soc Econ Plann Sci* 11:13–17
- Berger T (2001) Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion. Resource use changes and policy analysis. *Agr Econ* 25:245–260
- Boerner REJ, DeMers MN, Simpson JW, Artigas FJ, Silva A, Berns LA (1996) Markov models of inertia and dynamic on two contiguous Ohio landscapes. *Geogr Anal* 28:56–66
- Bonham-Carter G (1994) *Geographic information systems for geoscientists: modelling with GIS*. Pergamon, New York
- Bonham-Carter GF, Agterberg FP, Wright DF (1988) *Integration of geological data sets for gold exploration in Nova Scotia*. American Society for Photogrammetry and Remote Sensing, Maryland
- Brissoulis H (2000) *Analysis of land use change: theoretical and modeling approaches*. Regional Research Institute, West Virginia University, Morgantown. <http://www.rri.wvu.edu/WebBook/Brissoulis/contents.html>. Accessed 14 May 2005
- Brown S, Hall M, Andrasko K, Ruiz F, Marzoli W, Guerrero G, Masera O, Dushku A, De Jong B, Cornell J (2007) Baselines for land-use change in the tropics: application to avoided deforestation projects. *Mitigation and Adaptation Strategies for Global Change* 12:1001–1026
- Carver S (1991) Integrating multi-criteria evaluation with geographical information systems. *Int J Geogr Inform Syst* 5:321–339
- Chen Q, Mynett AE (2003) Effects of cell size and configuration in cellular automata based prey-predator modeling. *Simul Model Pract Theory* 11(7–8):609–625
- Cheng J, Masser I (2004) Understanding spatial and temporal processes of urban growth: cellular automata modelling. *Environ Plann Plann Des* 31:167–194
- Clarke KC, Hoppen S, Gaydos L (1997) A Self-modifying cellular automaton model of historical urbanization in the San Francisco Bay Area. *Environ Plann* 24:247–261
- Couclelis H (1985) Cellular worlds: a framework for micro–macro dynamics. *Environ Plann* 17:585–596
- Couclelis H (1989) Macrostructure and microbehavior in a metropolitan area. *Environ Plann* 16:141–154
- Dale VH (1997) The relationship between land-use change and climate change. *Ecol Appl* 17(3):753–769
- UFMG (Universidade Federal de Minas Gerais) (2009) *Dinamica EGO*. <http://www.csr.ufmg.br/dinamica/>. Accessed 2 Apr 2009
- Drewett JR (1969) A stochastic model of the land conversion process. *Reg Stud* 3:269–280
- Eastman JR (2009) *Idrisi Taiga, guide to GIS and image processing*. Clark University Edition, p 342
- Eastman JR, Jin W, Kyem PAK, Toledano J (1995) Raster procedures for multi-criteria/multi-objective decisions. *Photogramm Eng Rem Sens* 61(5):539–547
- Eastman JR, Solorzano LA, van Fossen ME (2005) Transition potential modeling for land-cover change. In: Maguire DJ, Batty M, Goodchild MF (eds) *GIS, spatial analysis, and modeling*. ESRI Press, California, pp 357–385

- Engelen G (1988) The theory of self-organization and modeling complex urban systems. *Eur J Oper Res* 37:42–57
- Engelen G, White R, Uljee I, Drazan P (1995) Using cellular automata for integrated modeling of socio-environmental systems. *Environ Monit Assess* 34:203–214
- Etter A, McAlphine C, Wilson K, Phinn S, Possingham H (2006) Regional patterns of agricultural land use and deforestation in Colombia. *Agric Ecosyst Environ* 114:369–386
- Gar-On Yeh A, Li X (2009) Cellular automata, and GIS for urban planning. In: Madden M (ed) *Manual of geographic information systems*. ASPRS, Sacramento, pp 591–619
- Geoghegan J, Pritchard L Jr, Ogneva-Himmelberger Y, Chowdhury RR, Sanderson S, Turner BL II (1998) Socializing the pixel and pixelizing the social in land-use and land-cover change. In: Liverman D, Moran EF, Rindfuss RR, Stern PC (eds) *People and pixels: linking remote sensing and social science*. National Academy Press, Washington, pp 151–169
- GOFC-GOLD (2010) A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests, remaining forests, and forestation. GOFC-GOLD Report version COP16-1. GOFC-GOLD Project Office, Natural Resources Canada, Alberta
- Heywood I, Cornelius S, Carver S (1998) *An introduction to geographical information systems*. Addison Wesley Longman, England, pp 138–142
- Hosseinali F, Alesheikh AA (2008) Weighting spatial information in GIS for copper mining exploration. *Am J Appl Sci* 5(9):1187–1198
- Irwin EG, Geoghegan J (2001) Theory, data, methods: developing spatially explicit economic models of land use change. *Agric Ecosyst Environ* 85:7–23
- Jahan S (1986) The determination of stability and similarity of Markovian land use change processes: a theoretical and empirical analysis. *Soc Econ Plann Sci* 20:243–251
- Jankowski P (1995) Integrating geographical information systems and multiple criteria decision making methods. *Int J Geogr Inf Syst* 9(3):251–273
- Jankowski P, Richard L (1994) Integration of GIS-based suitability analysis and multicriteria evaluation in a spatial decision support system for route selection. *Environ Plann* 21:399–420
- Jantz CA, Goetz SJ (2005) Analysis of scale dependencies in an urban land-use change in the central Arizona–Phoenix region, USA. *Landsch Ecol* 16:611–626
- Kaimowitz D, Angelsen A (1998) *Economic models of tropical deforestation: a review*. Center for International Forestry Research, Bogor
- Kamusoko C, Aniya M, Bongo A, Munyadzi M (2009) Rural sustainability under threat in Zimbabwe – simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. *Appl Geogr* 29(3):435–447
- Karlin S, Taylor HM (1975) *A first course in stochastic processes*. Academic, New York
- Kok K, Farrow A, Veldkamp TA, Verbug PH (2001) A method and application of multi-scale validation in spatial land use models. *Agr Ecosyst Environ* 85(1–3):223–238
- Lambin EF (1994) *Modelling deforestation processes: a review*. TREES Publications Series B: Research Report no.1, European Commission, EUR 15744 EN
- Lambin EF (1997) Modelling and monitoring land-cover change processes in tropical regions. *Prog Phys Geogr* 21:375–393
- Lambin EF, Rounsevell M, Geist H (2000) Are agricultural land-use models able to predict changes in land use intensity? *Agr Ecosyst Environ* 82(1–3):321–331
- Levine N, Block R (2011) Are Bayesian journey-to-crime estimation: an improvement in geographic profiling methodology. *Prof Geogr* 63(2):213–229
- Li H, Reynolds JF (1997) Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: a combined Markov and cellular automaton approach. In: Quattrochi DA, Goodchild MF (eds) *Scale in remote sensing and GIS*. Lewis Publishers, Boca Raton, pp 211–230
- Liu JG, Mason PJ (2009) *Essential image processing and GIS for remote sensing*. Wiley-Blackwell, Oxford
- Maeda EE, Almeida CM, Ximenes AC, Formaggio AR, Shimabukuro YE, Pellikka P (2011) Dynamic modeling of forest conversion: Simulation of past and future scenarios of rural

- activities expansion in the fringes of the Xingu National Park, Brazilian Amazon. *Int J Appl Earth Observation Geoinformation* 13:435–446
- Matheussen B, Kirschbaum RL, Goodman IA, O’Dennel GM, Lettenmaier DP (2000) Effects of land cover change on stream flow in the interior Columbia river basin (USA and Canada). *Hydrolog Process* 14(5):867–885
- Messina J, Walsh S (2001) 2.5D morphogenesis: modeling landuse and landcover dynamics in the Ecuadorian Amazon. *Plant Ecol* 156:75–88
- Moreno N, Quintero R, Ablan F, Barros F, Davila J, Ramirez H, Tonella G, Acevedo MF (2007) Biocomplexity of deforestation in the Caparo tropical forest reserve in Venezuela: an integrated multi-agent and cellular automata model. *Environ Model Softw* 22:664–673
- Moreno N, Wang F, Marceau DJ (2010) A geographic object-based approach in cellular automata modeling. *Photogramm Eng Rem Sens* 76(2):183–191
- Muller MR, Middleton J (1994) A Markov model of land use change dynamics in the Niagara Region, Ontario, Canada. *Landsc Ecol* 9:151–157
- Myint SW, Wang L (2006) Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach. *Can J Remote Sens* 32(6):390–404
- Paegelow M, Olmedo MTC (2005) Possibilities and limits of prospective GIS land cover modeling: a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain). *Int J Geogr Inform Sci* 19(6):697–722
- Parker DC, Manson SM, Janssen MA, Hoffman M, Deadman P (2003) Multi-agent systems for the simulation of land-use and land-cover change: a review. *Ann Assoc Am Geogr* 93(2):314–337
- Pastor J, Bonde J, Johnston C, Naiman RJ (1993) Markovian analysis of the spatially dependent dynamics of beaver ponds. In: Gardner RH (ed) *Predicting spatial effects in ecological systems: lectures on mathematics in the life sciences*, 23. American Mathematical Society, Providence, pp 5–27
- Petit C, Scudder T, Lambin E (2001) Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. *Int J Rem Sens* 22:3435–3456
- Pijanowski BC, Brown DG, Shellito BA, Manik GA (2002) Using neural networks and GIS to forecast land use changes: a land Transformation Model. *Comput Environ Urban Syst* 26:553–575
- Pontius RG Jr, Malanson J (2005) Comparison of the structure and accuracy of two land change models. *Int J Geogr Inf Sci* 19(2):243–265
- Pontius RG Jr, Cornell JD, Hall CAS (2001) Modelling spatial patterns of land use change with GEOMOD: application and validation for Costa Rica. *Agr Ecosyst Environ* 85(1–3):553–575
- Pontius RG Jr, Huffaker D, Denman K (2004) Useful techniques of validation for spatially explicit land-change models. *Ecol Model* 179:445–461
- Rindfuss RR, Walsh SJ, Mishra V, Dolcemascalco GP (2003) Linking household and remotely sensed data: methodological and practical problems. In: Fox J, Rindfuss RR, Walsh SJ, Mishra V (eds) *People and the environment: approaches for linking household and community surveys to remote sensing and GIS*. Kluwer Academic, Boston, pp 1–29
- Robinson VB (1978) Information theory and sequences of land use: an application. *Prof Geogr* 30:174–179
- Rykiel EJ Jr (1996) Testing ecological models: the meaning of validation. *Ecol Model* 90:229–244
- Saaty TL (1977) A scaling method for priorities in hierarchical structures. *J Math Psychol* 15:234–281
- Saaty TL, Vargas LG (2001) *Models, methods, concepts and application of analytic hierarchy process*. Kluwer Academic, Boston
- Silvertown J, Holtier S, Johnson J, Dale P (1992) Cellular automaton models of interspecific competition for space – the effect of pattern on process. *J Ecol* 80:527–534
- Soares-Filho BS, Cerqueira GC, Pennachin CL (2002) DINAMICA: a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecol Model* 154:217–235

- Soares-Filho BS, Nepsta DC, Curran LM, Cerqueira GC, Garcia RA, Ramos CA, Voll E, McDonald A, Lefebvre P, Schlesinger P (2006) Modelling conservation in the Amazon basin. *Nature* 23:520–523
- Teixerira AMG, Soares-Filho BS, Freitas SR, Metzger JP (2009) Modeling landscape dynamics in an Atlantic rainforest region: implications for conservation. *For Ecol Manag* 257:1219–1230
- Terrestrial Carbon Group (2008) How to include terrestrial carbon in developing nations in the overall climate change solution. <http://www.terrestrialcarbon.org/>. Accessed 14 Jan 2011
- Tobler W (1979) Cellular geography. *In: Gale S and Olsson G (eds) Philosophy in geography*. Reidel Publishing Company, Dordrecht, pp 379–386
- Torrens PM (2000) How cellular models of urban systems work (1. Theory). CASA Working Paper Series, 28. Centre for Advanced Spatial Analysis (UCL), London. http://www.casa.ucl.ac.uk/working_papers/paper28.pdf. Accessed 17 Aug 2009
- Torrens PM (2006) Simulating sprawl. *Ann Assoc Am Geogr* 96(2):248–275
- Torrens PM, Benenson I (2005) Geographic automata systems. *Int J Geogr Inform Sci* 19(4):385–412
- Veldkamp A, Verburg PH, Kok K, de Koning GHJ, Priess J, Bergsma AR (2001) The need for scale sensitive approaches in spatially explicit land use change modelling. *Environ Model Assess* 6:111–121
- Verburg PH, Schot PP, Dijst MJ, Veldkamp A (2004) Land use change modelling: current practice and research priorities. *GeoJournal* 61:309–324
- Voogd H (1993) Multicriteria evaluation for urban and regional planning. Pion, London
- Walsh SJ, Entwisle B, Rindfuss RR, Pages PH (2006) Spatial simulation modelling land use/land cover change scenarios in northern Thailand: a cellular automata approach. *J Land Use Sci* 1(1):5–28
- Weng Q (2002) Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *J Environ Manag* 64:273–284
- White R, Engelen G (1997) Cellular automata as the basis of integrated dynamic regional modeling. *Environ Plann* 24:235–246
- Wolfram S (1984) Cellular automata as models of complexity. *Nature* 311:419–424
- Wood EC, Tappan GG, Hadj A (2004) Understanding the drivers of agricultural land use change in south-central Senegal. *J Arid Environ* 59:565–582
- Wu F, Webster CJ (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environ Plann* 25:103–126
- Wu Q, Li H, Wang R, Paulusen J, He Y, Wang M, Wang Z (2006) Monitoring and predicting land use changes in Beijing using remote sensing and GIS. *Landsc Urban Plann* 78(4):322–333
- Zhou G, Liebhold AM (1995) Forecasting the spread of gypsy moth outbreaks using cellular transition models. *Landsc Ecol* 10:177–186