CHAPTER 23

MODELING LAND USE AND LAND COVER CHANGE

DANIEL G. BROWN 1, ROBERT WALKER 2, STEVEN MANSON 3, KAREN SETO 4

1School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109 USA, danbrown@umich.edu
2Department of Geography, Michigan State University, East Lansing, MI 48824
3Department of Geography, University of Minnesota, Minneapolis, MN 55455
4Institute for International Studies, Stanford University, Stanford, CA 94305

1 Introduction

Models are used in a variety of fields, including land change science, to better understand the dynamics of systems, to develop hypotheses that can be tested empirically, and to make predictions and/or evaluate scenarios for use in assessment activities. Modeling is an important component of each of the three faci outlined in the science plan of the Land use and cover change (LUCC) project (Turner et al., 1995) of the International Geosphere-Biosphere Program (IGBP) and the International Human Dimensions Program (IHDP). In Focus 1, on comparative land use dynamics, models are used to help improve our understanding of the dynamics of land use that arise from human decision-making at all levels, households to nations. These models are supported by surveys and interviews of decision makers. Focus 2 emphasizes development of empirical diagnostic models based on aerial and satellite observations of spatial and temporal land cover dynamics. Finally, Focus 3 focuses specifically on the development of models of land use and cover change (LUCC) that can be used for prediction and scenario generation in the context of integrative assessments of global change.

Given space limitations, we focus on spatially explicit models of LUCC. Because the majority of models of this sort are implemented at relatively local scales - sometimes called landscape scales (e.g., 1-100,000km²), we focus on these scales. These models, therefore, may not be appropriate for scaling up to continental and global scales. However, we discuss needs and prospects for models coupling cross-scale dynamics towards the end of the chapter.

In the next section, we summarize existing literature reviews of LUCC models, with a focus on typologies of models. Next, we discuss the importance of and approaches to addressing causation in LUCC processes. We then discuss specific modeling approaches in two broad categories: empirically fitted models and process simulation models. Though there is a good deal of overlap between these two categories, the categories address a fundamental distinction in the ways models have been built. Empirically fitted models are inductive, in that they seek descriptions of processes based on data measuring outcomes from those processes observed at specific places and times. We present an illustration from a NASA LCLUC project in China. Process simulation models are generally deductive, in that they start with a general understanding about processes and seek to build models that simulate outcomes in

395

specific places. In this context, we discuss the substantial challenges of calibration and validation and present an illustration from a NASA LULUC project in Mexico. We conclude with remarks about the needs and prospects in LULCC modeling.

2 Types of LULCC Models

Literally hundreds of models of LULCC have been described in the literature on landscape ecology, geography, urban planning, economics, regional science, computer science, statistics, geographic information science, and other fields. Because of differing disciplinary perspectives, as well as differing methodological approaches, data availabilities, and modeling goals, attempts to categorize models are complicated by a relatively large number of dimensions on which the models vary. A number of reviews of LULCC models have been produced in recent years, each from their own perspective and producing a number of different typologies. What follows is a brief summary of several of these efforts, and an outline of the categories of models discussed in this chapter.

Perhaps the first of these reviews was published by Baker (1989) in the context of landscape ecology, so its focus was on land cover change. Models were grouped according to the goals of the models. Whole landscape models seek to model change in some aggregate attribute or state of the landscape over time. Distributional models describe changes in the proportion of the landscape in each of a number of land cover classes. Spatial landscape models describe the location and configuration of changes in land cover. Focused as it was on landscape ecological processes, the paper by Baker (1989) did not discuss models that included explicit representation of human decision-making.

A pair of publications in the late 1990s reviewed the significant amount of work that had gone into modeling tropical deforestation (Lambin 1997; Kaimowitz and Angelsen, 1998). Whereas the review by Lambin (1997) described models of observed land cover change that used mathematical, empirical/statistical, and spatial simulation models, Kaimowitz and Angelsen (1998) focused on land use change models that were developed to describe micro- regional- or macro-economic aspects of the deforestation process, using similar methodological categories. Later, Irwin and Geoghegan (2001) compared and contrasted non-economic models, which included many of the approaches described by Baker (1989) and Lambin (1997) but also included cellular automata (CA), and economic models of LULCC, which were divided into non-spatially explicit and spatial explicit models.

Another recent review offers an alternative typology of land use change models. Agarwal et al. (2002) described 19 models that were arrayed on dimensions of space, time, and human decision-making. Models were characterized on these three dimensions according to both the scale at which they operated and the degree of complexity in the representation. The attention to the ways in which human decision-making is represented, in both the reviews by Kaimowitz and Angelsen (1998) and Agarwal et al. (2002), draws attention to the need for models that represent human decision-making explicitly. Another recent review (Parker et al., 2003) focused on agent-based models as tools for representing human decision-making and simulating the aggregate outcomes that result from decisions made by many individuals. Agent-based models represent a qualitatively different approach to the mathematical and
statistical approaches that pervaded the earlier reviews, and offer potential for new LUCC models, similar to the approach to ecological modeling offered by individual-based models (DeAngelis and Gross 1992). Parker et al. (2003) argue that combining agent-based models, to represent human level decision-making, with cellular models, to represent biophysical landscape change, offers a promising approach for future model development in LUCC.

An important distinction that is made in only a few of these reviews is between models of land use change and models of land cover change. While the reviews by Kainmowitz and Angelsen (1998), Irwin and Geoghegan (2001), and Agarwal et al. (2002) explicitly focused on models of land use change, those of Baker (1989) and Lambin (1997) are oriented towards models of land cover change. The review by Parker et al. (2003) appears to be the first to address both land use and land cover change, though it has a specific focus on agent-based models. The distinction is important because it affects both the data requirements for simulation and validation and the process representations required. There is not always a one-to-one relationship between the two (Cihlar and Janssen, 2001) and, while human activity defines land use, land cover change can proceed with or without a proximal human driver (e.g., through climate change). If we are to bridge human activity and ecological structure and function, representations of both processes are required, adding greater dimensionality to considerations of models in this field.

For this chapter, we describe two groups of models according to whether the modeling approach is oriented primarily toward (a) fitting data or (b) simulating processes. The first category of models includes a broad range of models that employ social science theory and represent decision-making and biophysical processes to varying degrees. We start by describing a range of fitted models of LUCC that help us understand how much LUCC, and of what types, is occurring where. Because of the large number of studies that have involved land use models within an econometric framework, and because of the ready availability of land cover data from remote sensing, we distinguish between fitted land cover models and econometric land use models. Next we describe process simulation models of LUCC that are built as generative representations of the elemental processes of agent decision-making and/or biophysical landscape change for simulating change outcomes.

3 Causation in LUCC Models

LUCC are responses to social and ecological processes on a landscape. Much has been written about the causes of LUCC, both in specific contexts, like the tropics, and in general (Fujita 1989; Fujita et al., 1999). Because of the complexity of causes of LUCC, a useful and commonly held distinction is between the proximate and ultimate causes (Geist and Lambin, 2001; 2002).

Causal explanation draws on both the social and natural sciences. In the social sciences focus is on the land manager, or the so-called "agent." Agents of interest are often those engaged in agricultural activity (e.g., Lynam 2003) or residential development (e.g., Blockset 1996). Anthropologists, economists, geographers, and sociologists typically take the agent as the point of departure for analysis, reaching out to ever broader contexts to identify the circumstances and conditions affecting the behavior and social processes of interest. Political ecologists (e.g., Blaikie and
Brookfield, 1989) and ecometrics (e.g., Chornitz and Gray, 1996), for example, tend to prefer explanations in which individuals pursue behavioral objectives (e.g., profit maximization, risk minimization) within the constraints of their social and ecological circumstances, and in response to broader-scale forces. Ecologists have focused on the ecological and edaphic constraints on human activity, such as when they affect decisions about placement of nature reserves and agricultural activities (e.g., Huston 1993). Further, ecological processes associated with disturbance, species range expansion, and competition can result in land cover changes, either independently or in interaction with human activity. The proximate human causes of LUCC, namely the actions taken by individuals, are therefore promoted and constrained by underlying factors, such as biophysical variability and feedbacks, social structures and macro-economic circumstances (Geist and Lambin, 2001; 2002).

Types of LUCC differ depending on the types of agents that are active on the landscape. In a large and/or heterogeneous region, different agents may be active in different areas, which adds spatial complexity to any attempt at causal explanation (Walker et al., 2000). Further complicating the matter is that LUCC may be substantially affected by dynamic interactions, among agents or between agents and their environment (Geist and Lambin, 2001). The deforestation literature has long recognized interactions between loggers and follow-on farmers, for example (Walker 1987; Rudel 2002). Further, there is increasing evidence that negative feedbacks associated with prior development play an important role in urban land development decisions (Irwin and Bockstael, 2002).

The agents and their interactions have been mainly described as constituting the proximate causes of LUCC (Geist and Lambin, 2001; 2002), in which case it is necessary to consider the ultimate forces in order to arrive at an encompassing account of causality. While research on the ultimate causes has tended to emphasize so-called "economic" factors (e.g., Kaimowitz and Angelsen, 1999; Irwin and Googhegen, 2001), these are shaped by underlying environmental heterogeneity and variability, demographic change, technological evolution, and institutional intervention (Walker et al., 2002). Thus, there are two primary scales of processes at work in LUCC, a micro-scale, in which individuals seek to achieve their objectives, and a macro-scale, reflecting the context of these decisions. The context of decisions include such processes as population growth and movement, the climate and soil processes that constrain production on the land, the evolution of commodity and labor markets, the continuing progress of technological change, and the actions of government bureaucrats responding to political forces. Changes in the macro-scale are often unpredictable, given that they result from complex interactions of economic, political, and transnational institutions and social processes. It is these macro-scale processes and changes that are usually interpreted as ultimate causes in LUCC research.

Finally, it would be a mistake to conclude that there is a one-way causal path, running from macro-scale forces to the micro-reponses of individuals. Indeed, individuals and communities show remarkable resilience and ingenuity in adapting to a changing context to pursue their own objectives. Thus, while macro-scale factors no doubt shape the realm of choices available to agents, they by no means predetermine any particular land use or novel symbolism.
4 Empirically Fitted Models

The first type of models that we describe includes those that are empirically fitted, i.e., they are based on statistically matching the temporal trends (in the case of longitudinal data) and/or spatial patterns (in the case of cross-sectional data) with some set of predictor variables. The variables to be predicted may represent land use, land cover, change in either of these or some combination and may be measured over pixels, derived from remotely sensed data, parcels (i.e., irregularly shaped spatial units defining legal ownership) or aggregated over some jurisdictional unit. Predictor variables are factors thought to influence land use, including proximity to roads, proximity to cities and towns, mix of economic activity, demography, income and wealth, and biophysical factors like slope and soils.

Some of the earliest empirically fitted models represented LUCC as Markov random processes, in which the state of the land at some time in the future was a function only of its present state, represented by a transition probability (Burnham 1973; Bell 1974; Muller and Middleton, 1994). Though these simple models have appeal for their minimal data requirements and analytical properties, their assumptions are restrictive and they have little utility for analyzing policy. To address some of the restrictions, later models have used the Markov framework, but added a dependence on neighboring states (e.g., Turner 1987) and introduced spatial and temporal non-stationarity to the transition probabilities (Skirvin and Lofstrom, 1991; Brown et al., 2000). Future applications of Markov modeling to LUCC will need to incorporate these modifications, which means the models are not strictly Markovian, but rather a more general form of stochastic model.

Statistical estimation methods are more commonly used to fit LUCC models. Some of these models have been constructed with only loose connection to theory, while others are constructed in a more rigorous, usually economic, theoretical context. Given the discrete nature of land use and cover categories and changes, the most common approach is to estimate a logistic regression function that describes either the probability of a particular land category occurring or of the location transitioning from one land category to another. This approach has been used for modeling development in urban areas (Landis 1994; Landis and Zhang, 1998a; 1998b), along urban-rural gradients (Wear and Bolstad, 1996), and in tropical forests (Chomitz and Gray, 1996; Mertens and Lambin, 1997).

A number of limitations challenge the utility of simple statistical models, and solutions have been developed or are in development to deal with these challenges. The assumption of log-linear relationships between the predictor variables and the dependent variable can be restrictive. In order to relax this assumption, a number of non-linear fitting methods have been used to describe land use or cover change, including generalized additive modeling (Brown et al., 2002), which is a non-linear statistical method (Hastie and Tibshirani, 1990), and artificial neural networks (Pijanowski et al., 2002), which is a networked learning approach that comes from artificial intelligence research. The assumption of temporal stationarity (i.e., that the model parameters remain constant over time) leads to models with misleading $R^2$ values and $t$-statistics. These problems can be dealt with using econometric methods developed for analysis of panel or time series data (e.g., Hsiao 1986). An illustration of a panel data analysis is presented below.
Spatial autocorrelation and non-stationarity is also likely to affect estimation efforts because of the effects of contiguity in spatial data, together with the diffusory nature of many land change processes (Walker and Solecki, 1999; Walker et al., 2000). It is necessary to account for the influence of spatial autocorrelation in these models because, otherwise untreated, it can lead to erroneous conclusions about the relationships between dependent and independent variables. Wear and Bolstad (1998) included a neighbor-weighted dependent variable in the list of predictor variables, whereas Bell and Bockstael (1997) used a method that incorporates the spatial weight matrix in a generalized-moments estimation. The method presented by Bell and Bockstael (1997) was applied to household level data, but could be extended to pixel- or parcel-level land use and cover data.

Failure to incorporate information (often survey-based) about the household or community structures can create specification bias (Walker et al., 2002), because land use processes may be different for different types of households or communities.

Empirical models can suffer from the ecological fallacy – when the characteristics of an individual agent are inferred, often incorrectly, from estimation based on aggregated observational units – and the modifiable areal unit problem – where the shape and size of data aggregation affects analysis (Chou 1993). Variables differentially co-vari as a function of the scale at which they are measured; scale can therefore act as an independent variable (Bian 1997; Walsh et al., 1999) and affect the apparent magnitude and direction of relationships among causal factors (Kimmer and Shan, 1994).

Many empirical models assume unidirectional causal change, i.e., income or population growth causes deforestation or urban growth, but that there are no bi-directional feedbacks or testing for causality. Endogenous interactions and feedbacks have been incorporated into some models (e.g., Chomitz and Gray, 1996), but many empirical models assume them away. Models that fail to identify the endogeneity of variables will be misspecified and policy recommendations may be incorrect.

4.2 PANEL APPROACHES IN THE PEARL RIVER DELTA

Panel analysis methods represent promising options for developing empirically fitted LUCC models, as this example demonstrates. Among the chief advantages to panel approaches is the ability to allow the relation between the independent variables (drivers of land use change) and the dependent variable (land use change) to vary across time and space (Hsiao 1986). Panel econometric techniques were used to model land use change in the Pearl River (Zhujiang) Delta (PRD) in South China (Seto and Kaufmann, 2003). Two important aspects of this study affected the choice of modeling methodology employed. First, the remote sensing data used to extract land use trajectories consisted of nine consecutive images from 1988 to 1996. Although this time series is longer than many land use change studies, it is a relatively short period that does not allow sufficient degrees of freedom to estimate a statistically reliable model if the entire study area is treated as a whole. Thus, to increase the reliability of the statistical results, the land use data were augmented by cross-sectional data at the
smallest possible administrative unit, the county. Second, the main objective of the study was to evaluate the macro-level causes of land use change, and in particular, the urbanization of agricultural land and natural ecosystems. Because the aim was to understand the relation between policy reforms and land use, a modeling approach appropriate to address this level of analysis was required.

As with all fitting techniques, the specification of the model is an important determinant of the quality of the results. Model misspecification can lead to spurious results. This is especially the case with panel data, where model coefficients can vary both temporally and spatially. In the PRD case study, panel statistical tests were used to identify the appropriate model based on the characteristics of the data (see Hsiao 1986). A result of this analysis was that the socioeconomic factors correlated with land use change were shown to vary by county. This indicated that a random coefficient model was most suitable. Following estimation of the models, causation among the dependent and independent variables was explored using the method of Granger causality (Granger 1969; Granger and Huang, 1997), which examines lagged correlations in time-series data to identify chains of causation. Results from the analysis of urbanization of natural ecosystems revealed complex interactions and feedbacks among the various factors (Table 1). Two of the independent variables, investment in capital construction and relative rates of labor productivity, appeared to 'Granger cause' the conversion of natural ecosystems to urban uses, i.e. they were correlated with the conversion in space and preceded it in time. This type of land use change also appears to have a feedback on capital construction, suggesting that high rates of urbanization attract additional investments in construction. These results suggested that land use change occurred on a nested-hierarchy of scales, and that the underlying driving forces were not simple, nor necessarily unidirectional. The techniques used in the PRD illustrate some of the ways in which advances in theoretical and applied econometrics are being adapted for land use change modeling.

Table 1. Analysis of causal order for the natural ecosystem–urban model (Seto and Kaufmann, 2003).

<table>
<thead>
<tr>
<th>Causal Variable</th>
<th>Dependent Variable</th>
<th>natural–urban</th>
<th>Capital Investment</th>
<th>Ag/Urb Land Productivity</th>
<th>Ag/Urb Labor Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>natural–urban</td>
<td></td>
<td>-</td>
<td>32</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Capital Investment</td>
<td></td>
<td>21</td>
<td>-</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Ag/Urb Land Productivity</td>
<td></td>
<td>8</td>
<td>14</td>
<td>-</td>
<td>23</td>
</tr>
<tr>
<td>Ag/Urb Labor Productivity</td>
<td></td>
<td>21</td>
<td>17</td>
<td>14</td>
<td>-</td>
</tr>
</tbody>
</table>

Cell values refer to the number of occurrences, out of 165 sub-samples, where the causal variable "Granger caused" the dependent variable. Values in bold are statistically significant at p<0.05.
5 Dynamic Process Models

In contrast to empirically fitted models, dynamic process models of LUCC seek to represent the most important interactions between agents, organisms, and their environment as computer code. Though fitted models can be used to generate simulations (e.g., Brown et al., 2002), simulation is clearly more central to the use of process models. Although these models need to be calibrated and validated, they deemphasize the fitting of data and emphasize the fidelity of model elements and processes to what is known about the processes. Here we describe three types of dynamic process models that have been used for LUCC: process flow models, cellular automata (CA), and agent-based models.

Models that track material and energy flows through a landscape and how they are transformed and/or transmitted at each location (e.g., Fitz et al., 1996), work well for some natural processes and to represent stable, stylized socioeconomic systems. Models of this sort have been used in a variety of settings, including the Everglades (Wu et al., 1996) and the Patuxent Watershed (Vosniou et al., 1999). These modeling approaches often include enough stochasticity that they become simple simulation platforms to evaluate empirical models as described above. Their primary limitation for modeling LUCC, however, is an inability to represent the complex ecological or social interactions, including competition and cooperation, the role of institutions, interest groups, and limits to rational behavior, that give rise to behaviors and decisions that affect land use and cover.

Cellular automata (CA) and other cellular models are fairly common in the land use modeling literature (e.g., Betty and Xie, 1994; Betty et al., 1997; Clarke and Gaydos, 1990). In some CA models, the cells are considered simple actors with fixed neighborhood relations and update rules. In other models, the CA represents the state and dynamics of the environment. The cells can represent parcels of land, each with their own characteristics and each changing as a result of some fixed rules based on the cell's state and the state of its neighbors. CA models, coupled with data on the heterogeneity of the environment, can capture important dynamics of LUCC, including diffusive, road-driven, and spatially random changes (Clarke and Gaydos, 1998). Further, because they are dynamic and iterative, CA models can represent endogenous interactions and feedbacks. An important challenge in the development of CA models is how to establish the rules that govern system behavior, and incorporation of heterogeneity and dynamism into these rules. Also, the simplified decision rules of CA models make policy interventions difficult to explore at the level of individual decision makers, social groups, or institutions.

Agent-based models (ABMs) are defined in terms of entities and dynamics at a micro-level, i.e., at the level of individual actors and their interactions with each other and with their environment (Epstein and Axtell, 1996; Kohler et al., 2000; Gimblett 2002; Janssen 2003a; Parker et al., 2003). ABMs consist of one or more types of agents, as well as an environment in which the agents are embedded. Agents in ABMs may be individuals (e.g., householders, farmers, developers) or institutions (e.g., townships, NGOs, firms; Boussequet et al., 1998; Gimblett 2002; Janssen 2003b). Therefore, systems can be studied at many scales and parts (specified at different scales) can be integrated into a coherent whole. Agent specification requires defining their state (e.g., preferences, memory of events, and social connections) and their decision-making rules, heuristics and other mechanisms to perform particular
behaviors. The agents generate their individual behaviors in response to inputs from other agents and from the environment. Agents may also adapt, or change their behavior, through learning or evolution based on their experience. Adaptive behaviors are usually represented in the models using some sort of learning algorithm, such as genetic algorithms. As the model is run, agent behavior is generated as agents use their rules to determine which other agents to interact with, what to do when they interact, and how to interact with the environment. Representing human decision-making in ABMs remains an active area of research (e.g., Balzani and Happe, 2002; Berger 2002; Hoffman et al., 2003; Lynam 2003).

The environment of an ABM typically represents the physical environment, e.g., land, water, roads or other infrastructure. The environment at any location has associated states, e.g., land cover type, soil quality, and aesthetic quality. The environmental entities in a model may have their own dynamics, describing how they change in time both independent of and as a result of agent behavior, e.g., to represent soil erosion processes, forest growth and other aspects of environmental change (e.g., Janssen et al., 2003; Lynam 2003). These dynamics are often represented by coupling with CA-based models.

The agents' behaviors affect each other and the environment. The environment changes in response to agent behaviors, but also by following its own dynamics. Thus ABMs embody complex interlaced feedback relationships, leading to the nonlinear, path-dependent dynamics often observed in complex systems. Note that the model's output is both the micro-behavior of agents and the environment, as well as the emergent macro-level structures, relationships and dynamics which result from the micro-level activity. Spatial relationships can be incorporated into agent behavior in order to create more realistic interaction with neighboring agents, such as when a farmer is more likely to learn from a nearby neighbor than from a farmer located far away (e.g., Polhill et al., 2001). However, the models also can include social networks of various kinds, each defining an interaction topology based on, for example, membership in groups, business contacts, and common information sources. These characteristics facilitate the investigation of three particularly important properties that define complex systems: emergence, scaling, and feedbacks. Parker et al. (2001) argued that spatial landscape metrics can be used as measurable map properties to "emerge" from individual-level decision-making and interaction. These measures can serve as scaled outputs from an agent-based model, which has inputs at the level of individual decision-making. Recently developed agent-based models have been used to examine feedbacks between land use change and transportation networks (Batty and Torrens, 2001; Waddell 2002) and environmental characteristics (Rand et al., 2002).

5.1 SOUTHERN YUCATAN PENINSULAR REGION INTEGRATED ASSESSMENT

An agent-based process model was developed to project trends and develop scenarios of tropical deforestation and cultivation in Mexico (Manson 2002). Termd SYPRIA (Southern Yucatan Peninsula Region Integrated Assessment), the modeling effort was part of a larger project in the Southern Yucatan Peninsula Region that has drawn on historical analysis, ecological and field studies, and remote sensing (Turner et al., 2001).

SYPRIA uses a three-component actor-institution-environment conceptual framework to model decision-making of farming households in the context of
socioeconomic institutions and the biophysical environment. Agents represent households whose characteristics are specified by a survey. The survey was aimed at determining household attributes such as food demand and labor availability. Remotely sensed imagery was used to map the outcomes of past household land use decisions. Agent agricultural location decisions were specified using a boundedly rational utility function that guides a multicriteria evaluation. Agents make decisions that optimize their utility and also adapt to experience through an agent-specific population of genetic programs (cf., Chattoe 1998; Dawid 1999).

Land tenure and market forces, such as returns to commercial agriculture and market accessibility, were included factors driving agent decision-making. Institutions were represented by simple agents with largely pre-scripted behavior. The environment, represented by a cellular model, affects agent decision-making by changing such factors as soil quality and cover type according to rules derived from ecological research (Lawrence and Foster, 2003; Read et al., 2003).

SVPRRA has been used to examine a number of scenarios, including the effects of agricultural commercialization, alternative land tenure policies, and different secondary succession regimes. A suite of quantitative and qualitative validation techniques was employed to test the model and examine scenario outcomes. Results to date indicate that genetic programming holds promise for representing bounded rationality and that institutions and population growth are key to understanding the quantity and patterning of LUCC in the study region (Manson 2002, 2003).

6 Calibration and Validation

There are substantial challenges to integrating empirical observations with models, especially process-based models. Some of these challenges are related to scalar dynamics. Social variables tend to act at fine scales while biophysical variables have greater influence at larger scales (McConnell and Moran, 2001). Scale mismatch can occur when multiple processes operate at different scales, such as when household decision-making occurs at a scale different from pertinent biophysical processes (Fresco and Kroonenberg, 1992). A model may be able to represent different processes at different scales (Veldkamp and Fresco, 1996). Given the challenges of multiple scales of analysis, most integrated datasets are available within local projects or regionally integrated networks, such as the Hindu Kush-Himalaya project, LUCC Moombo, and START/Southeast Asia.

For model validation purposes it is important to clearly identify the objectives of a model. Where accurate predictions are the ultimate goal, measures of the accuracy of spatial outcomes are reasonable for validation. LUCC models are commonly evaluated by comparing predicted outcomes to data using measures of map agreement, like the kappa statistic and the receiver operating characteristic (ROC). Pontius (2002) has presented methods to, in addition, identify, separately, the degree to which a model gets the correct amount of a given land use or cover versus getting their locations correct. Where the goal of a model is to represent the process accurately, however, for example so that policy interventions can be evaluated, validation might require evaluating how well a model reproduces certain critical system properties (e.g., a clustered pattern of change or a cyclical land use dynamic). This requires tests of spatial patterns and temporal dynamics (Turner et al., 1989; Pontius and Schneider, 2001), for
example using pattern and texture metrics (Giles and Trani, 1999). Researchers developing agent-based models have not yet identified a canonical set of approaches to calibration and validation of these models (Benson 1995; Battu and Torrens, 2001; Polhill et al., 2001; Waddell 2002). There remains a need for new measures of fit between the model and data that go beyond spatial matching to focus on variability of outcomes and dynamics.

Another important approach to evaluating process simulation models is sensitivity analysis that explores the relationships between model parameters and the state or time path of variables endogenous to the modeled system (Klepper 1997). Other work examines the effects of interdependency, nonlinear behavior and sensitivity to initial conditions (Miller 1998; Manson 2001). Such tests address issues of error propagation and uncertainty, a topic too often left unconsidered in LUCC modeling (Robinson 1994). Cogent sensitivity, error, and uncertainty handling is essential to vet data and ensure a model behaves in ways that are consistent with relevant data.

7 Prospects

Recent developments and applications of LUCC models have addressed a number of key conceptual and methodological issues that have limited these models in the past. These developments hold promise for future use of these models in addressing the pressing issues of which LUCC are an integral part, e.g., human health and welfare, global change, and ecosystem dynamics and sustainability. Nonetheless, work remains to improve the efficacy of methods for model development, testing, validation, and application. This future work will further improve our ability to evaluate the effects and interactions of LUCC on and with people and their environment.

Research on statistical, econometric, and computational methodologies has improved the ability of fitted models to account for non-linear relationships and endogeneity, spatial dependence and temporal non-stationarity. Routine use of some of these methods will take time and effort, including in the development of accessible software. Future research on fitted models should continue to emphasize developing and applying methods for testing for causality, measuring and incorporating spatial and temporal non-stationarity, modeling across scales, and incorporating endogenous factors and feedbacks. Tests for causality can help, for example, to develop appropriate and effective policy. Regardless of these advances, empirically fitted models will always be limited in their ability to explain and predict by the specific data and contexts within which they are calibrated and fitted. While they are useful for identifying correlations and causal chains within a particular conceptual framework and projecting trends within a relatively narrow range of conditions defined by the data, these models are limited in their ability to generate surprising results or to predict in situations where the proximal and ultimate causes are different from those operating in the empirical situation.

Computational process models have shown great promise for LUCC research. These methods will benefit from further development and testing in a range of LUCC settings and under a variety of scenarios. For example, work is needed to establish appropriate rules to govern the behavior of CA models and to incorporate heterogeneity and dynamism into these rules. Agent-based models are among the newest entries into the LUCC field and work remains to develop and test various alternative approaches to
representing human decision-making. As with other modeling approaches, further improvements in agent-based modeling will come through adoption of methods applied in other fields, especially in the general area of complex systems. Development and common acceptance of a set of methods for calibration and validation of the model processes and their outcomes is needed.

References
Bell, E.J. 1974. Metrics analysis of land-use change: application of stochastic processes to remotely sensed data. Socioecological Planning Sciences, 8(60), 311-316.


