Chapter 13 - The SYPR Integrative Assessment Model: Complexity in Development

Steven M. Manson

Department of Geography, 414 Social Sciences, 267 19th Avenue South, University of Minnesota, Minneapolis, MN 55455

Citation: Manson, S. M. (2004). The SYPR integrative assessment model: complexity in development. In *Integrated Land-Change Science and Tropical Deforestation in the Southern Yucatán: Final Frontiers*. B. L. Turner, D. Foster and J. Geoghegan (eds). Oxford, United Kingdom: Clarendon Press of Oxford University Press, pp. 271-291.

PREPRESS VERSION

Chapter 13 - The SYPR Integrative Assessment Model: Complexity in Development

1 Integrated Assessment

Be it global environmental change or environment and development, land-use and land-cover change is central to the dynamics and consequences in question in the southern Yucatan peninsular region. Designing policies to address these impacts is hampered by the difficulty of projecting land use and land cover, not only because the dynamics are complex but also because consequences are strongly place-based. This chapter describes an *integrated assessment* modeling framework that builds on the research detailed in earlier chapters in order to project land-use and land-cover change in a geographically explicit way.

Integrated assessment is a term that describes holistic treatments of complex problems to assess both science and policy endeavors in global environmental change (Rotmans and Dowlatabadi 1998). The most common form of integrated assessment is computer modeling that combines socioeconomic and biogeophysical factors to predict global climate. Advanced in part by the successes of these global-scale models, integrated assessment has expanded to structure knowledge and set research priorities for a large range of coupled human-environment problems. Increasing recognition is given to the need for integrated assessment models to address regional-scale problems that are masked by global-scale assessments (Walker 1994). Such models must address two issues to project successfully land use/cover change at the regional scale. First, change occurs incrementally in spatially distinct patterns that have different implications for global change (Lambin 1994). Second, a model must account for the complexity of, and relationships among, socioeconomic and environmental factors (Turner et al. 1995). The SYPR integrated assessment model therefore has a fine temporal and spatial grain and it places land-use/cover change at the intersection of land-manager decision making, the environment, and socioeconomic institutions. What follows is a description of an ongoing integrated assessment modeling endeavor of the SYPR project (henceforth, SYPR IA model).

2 Conceptual framework: actors, institutions, and environment

The depth and breadth of the SYPR project poses a challenge to the integrated assessment modeling effort since some unifying framework must reconcile a broad array of issues, theories, and data. The global change research community offers a general conception of how environmental change results from infrastructure development, population pressure, market opportunities, resource institutions, and environmental or resource policies (Stern, Young and Drukman 1992). When applied to land-use/cover change, these causes and effects are apportioned among the conceptual foci of social systems, ecological systems, and land managers (Turner et al. 1995). The SYPR IA model recasts these foci as a three-component actor-institution-environment conceptual framework. The first part focuses on the decision making of households and other actors that directly contribute to change in land use/cover. The second component concerns socioeconomic institutions that affect actor decision making. Both actors and institutions interact with the third component, the biophysical environment.

Before examining the conceptual framework in detail, it is useful to understand its context. Use of separate, yet linked, actor and institutional components reconciles views of decision making that range from narrowly defined individual rational choice to perspectives that focus on social and cultural structures. The actor component has roots in rational choice theory. Although criticized for its shortcomings, this framework has accommodated criticism and remains an analytically defensible technique when used in a manner that recognizes it limitations (Myers and Papageorgiou 1991). An extension to rational choice is bounded rationality, which relaxes assumptions regarding the generation of choice alternatives, such as perfect access to information, and attempts to account for factors such as learning over time (Tversky and Kahneman 1990; Simon 1997). Bounded rationality can allow for external influences in a manner consistent with social and cultural perspectives on decision making (Jager et al. 1997), embodied here in the institutional component of the conceptual framework.

A variety of theoretical frameworks speaks to actor decision making in a land-use/cover change context (see also Lambin 1994; Kaimowitz and Angelsen 1998; Agarwal et al. 2000). The actor component of the conceptual framework concerns agrarian decision making within the context of bounded rationality and resources available to actors. This view is based on agrarian and development literature that posits household behavior in terms of rational choice theory, where activity ranges from subsistence to market production (Singh, Squire and Strauss 1986). An actor chooses production activities that maximize utility; under bounded rationality the actor satisfices utility. The utility afforded to the actor by production activities is subject to external factors and actor attributes. The latter can be treated as a resource profile that includes measures of access to information, ability to learn, and personal characteristics (Sen 1981).

Actor decision making is affected by socioeconomic institutions (Chpt. 7). These are conceived as 'shared concepts used by humans in repetitive situations organized by rules, norms, and strategies...' (Ostrom 1999: 37). Rules are shared prescriptions that are predictably enforced while norms are those that are generally enforced through costs and inducements. Strategies are the regularized plans that individuals make given their understanding of rules, norms, and the likely behavior of others. Institutions are characterized by their members, geographical extent, degree of formalization, regulatory mechanisms, and administrative structure (Young 1994). Institutions are important to the SYPR IA model primarily because they influence actor decision making and they are a useful proxy for the external large-scale political economy.

Completion of the conceptual framework hinges on how the environment relates to institutions and actors. Actor production strategies in forests center on agriculture, timber, and nontimber forest products (Galletti 1998). These strategies are subject to biophysical configurations of precipitation, soil characteristics, secondary growth, pest infestations, and disturbances such as fire or hurricanes. Complex relationships, laden with feedback, result when these biophysical forces are in turn affected by production strategies that lead to changes such as cover conversion, soil erosion, or nutrient depletion (Uhl 1987, Chpts. 4 & 5).

The actor-institution-environment framework offers a general view on human-environment situations given its derivation from global environmental change research. Its key focus is how actor decision making affects the environment and how this decision making is influenced by the environment and institutions. The framework is sufficiently flexible, however, to allow for the fact that actors are active participants in institutions. Similarly, institutions and the environment can be seen as affecting each other through the intermediation of agent decision making. Finally, the conceptual framework complements work presented in previous chapters, particularly those concerning models of decision making (Chpts. 10 & 11). The dynamic linkage between actors and institutions is grounded in the rational choice framework that underlies SYPR's economic research.

3 Simulating land-use and land-cover change

While there are a number of spatiotemporally explicit modeling techniques, no single approach can implement the actor-institution-environment framework in order to create the SYPR IA model. It has been necessary, therefore, to create a unified software system that can combine a variety of approaches. While this simulation system has primarily been used to create the SYPR IA model, it can be treated as an entity unto itself.

The simulation system relies on scripts, text files written in a simulation language. Each script is a blueprint for building a simulation model from basic building blocks such as control structures, variables, functions and software objects. The simulation language provides three groups of software functions. The first are specialized simulation routines that allow the user to automate tasks and incorporate the functionality of other software applications. The second set of functions form an agent-based model, a system of software objects that semiautonomously manipulate information and perform actions. The third set of functions comprise a cellular automata, a two-dimensional grid where cell values change over time according to rules based on the value of adjacent cells or external inputs. Following sections examine these specialized functions within the context of the SYPR IA modeling effort.

Just as an assortment of tools and materials can be used to construct various buildings, the simulation system is a general toolbox that can be used to construct an array of models. When constructing a building, a general vision of the structure is embodied in a blueprint that in turn guides the application of tools and materials. The conceptual framework and integrated assessment principles provide the general intent and form of the SYPR IA model. A first draft blueprint is provided by the technique of dynamic spatial simulation as it has been applied to modeling deforestation (e.g., Southworth, Dale and O'Neill 1991; Gilruth, Marsh and Itami 1995). This method portrays the landscape as a two-dimensional grid where rules based on factors such as agricultural suitability determine how cells change. In an exhaustive review of deforestation models, this method is considered the ‘most advanced modeling approach for a complex, dynamic and spatial problem such as tropical deforestation’ (Lambin 1994: 92).

This basic architecture has been expanded by using cellular automata functions to model the environmental component of the conceptual framework and agent-based modeling is used for the actor and conceptual components. It is important to emphasize the difference between the uses to which the agent and cellular models are applied. Cellular automata models are used to represent actor decision making when the actors are distributed in a manner consistent with the assumptions underlying cellular automata, such as tessellated actors reliant on local interactions (Benenson 1999). In cases where these assumptions do not hold, there is much to be gained by agent-based approaches instead of cellular automata (Torrens and O'Sullivan 2001).

Before continuing, it is important to review some terminology. Key is the distinction among the terms ‘agent’, ‘actor’, and ‘institution’. The latter two refer to components of the conceptual framework in that ‘actors’ are land managers and ‘institutions’ are socioeconomic and political frameworks while an ‘agent’ is a software object used to represent individual actors and institutions. The correct term for an agent representing an actor or institution, respectively, is ‘actor-agent’ or ‘institution-agent’ but when examining obviously computational or methodological details of a model, the -agent suffix is assumed. Also important is the use to which the terms ‘model’ and ‘simulation’ are applied. The model is a framework of concepts and data that is embodied in simulation language scripts to become a ‘simulation model’. In this sense, it is most appropriate to refer to the ‘SYPR integrated assessment simulation model’ but the ‘SYPR IA model’ suffices given the specificity of its application.

4. SYPR IA model overview

When set in motion, the SYPR IA model creates an artificial world, a simulated southern Yucatán peninsular region as it was in 1970. Time then starts in this simulated world, measured as a series of simulated years over forty model ‘years’ (1970 to 2010). Within each year, four general processes occur in the simulated study region. First, parameters exogenous to the simulated world are introduced or changed. Second, institutions change actor resource profiles or decision variables. Third, the environment, expressed primarily as variables related to land cover, evolves according to ecological rules and the impact of previous actor decision making. Finally, each actor in the simulated southern Yucatán peninsular region makes land-use decisions that take into account its resource profile, exogenous parameters, institutional imperatives and environmental factors. The results of these decisions are registered in the actor’s resource profile and as an environmental impact when appropriate. Upon completion, the simulation results are validated against observed data, as described below. The next six subsections explore the simulation procedure in detail.

Before continuing, however, a quick overview of data is in order. Much of the modern history of land-use/cover change in the region is accessible through government documentation, scientific research, the local populace, remotely sensed data, and the SYPR project. As described throughout this volume, the SYPR project provides spatial data and an actor survey of households (Chpts. 8 & 10). Taken across ten *ejidos*, the survey details an extensive set of socioeconomic characteristics that bear on decision making, including household statistics, production activities, labor factors, institutional affiliations, and spatiotemporally explicit land-use histories. In addition to these land-use histories, the project’s spatial data includes satellite imagery, land-use and land-cover maps, aerial photography, and geographical information system layers for a host of socioeconomic and biogeophysical phenomena.

4.1 User model inputs

Within each iteration, the model first initializes or updates information that feeds into later simulation stages. These user-specified simulation variables are stored in database tables and spatial layers where it may be accessed by model components. The bulk of user variables concern production activities, such as the rate at which land may be cleared or the effort required to maintain land under cultivation. This information is derived from research literature and project personnel. Scripts calculate production characteristics every year as function of household characteristics. The rate at which land can be cleared, for instance, depends on labor and technology, such as number of family members or possession of a chainsaw. Miscellaneous data, such as the amount of food necessary to support an individual, affects agent decision making. These are determined through consultation with SYPR project members, government publications, and research literature. The extent and depth of variables is typically dependent on the degree to which these data are necessary to other parts of the model.

4.2 Institutions

Critical institutions in the region include state-owned forests, forest concessions, the biosphere reserve, government subsidy programs, nongovernmental organizations, and *ejidos* (Alcorn and Toledo 1998, Chpts. 3, 7, & 10). The model focuses on large scale, agricultural institutions: land tenure, governmental subsidies, and economic markets. Land tenure includes *ejidos*, the biosphere reserve, state-owned forests, and timber concessions. Government subsidies through the PROCAMPO program offer direct payments to farmers (as described in Chpt. [?]).

The SYPR IA model uses agent-based model functions to represent both institutions and actors. This method is increasingly used to combine empirical and theoretical knowledge of decision making. Agents are software programs that represent adaptive autonomous entities that extract information from their environment and apply it to functions such as perception, planning, and learning (Conte, Hegselmann and Terna 1997). All three institutions have spatial characteristics that can be expressed as geographic information system layers that institution-agents combine with internal rules to determine how and when to influence actor-agent decision making. The land tenure agent guides actor decision making by limiting land use according to location and actor characteristics. Subsidy institutions make available funds to targeted actors by changing their resource profiles. The economic market agent creates a layer for distance to market and makes it available to all agents and sets prices according to user specified variables.

All three institutions act in a deterministic manner and, as such, are less interesting than actors as examples of how an agent-based model can be used. Given the flexibility of the modeling system, however, later iterations of the SYPR IA model can link to econometric research or to larger scale economic models in a manner similar to that employed by other regional scale models (e.g., Riebsame et al. 1994; Yates and Strzepek 1998). It is also possible to develop more intentional and dynamic institution-agents that would have fuller participation of actors (e.g., Seror 1994).

All three institutions in the SYPR IA model (land tenure, subsidies, and the market) are important to land use in the southern Yucatan peninsular region. Land tenure is essential to land use and much of this resides in ejidos. Plot-level data, as gathered in the 1997 and 1998 survey (Chpt. 8), are valuable given its spatial specificity, but acquiring it for the entire study region would prove resource intensive. Market and subsidy information is critical to harness the explanatory power offered by actor decision making models based on economic theory. Finally, Chapter 7 describes how institutions not included in the present model can affect land use, and incorporating them is a long-term goal of the SYPR IA modeling effort.

4.3 Environment

The environmental component of the conceptual framework is based on cellular automata, which have a long history of being applied to physical and biological processes (Smith and Bull 1997). The raw material of cellular automata is a discrete lattice of uniformly shaped cells. This lattice is functionally identical to the geographic information system raster layers in which SYPR stores spatial data, since these are two-dimensional grids of cells. The SYPR model therefore uses these layers as a base for the cellular automata, defined as [S, N, T] where S is a set of finite cell states, N are the set of states of neighboring cells, and T is a set of transition rules that pair input-state tuples to an output state for the cell.

The prime example of states are land-use and land-cover classes, where states are mutually exclusive categories, such as those created by the image processing efforts of the project (Chpt. 6). Examining a forest/nonforest case, cells exist in either in a forest state (s_1) or nonforest state (s_2), $S = \{s_1, s_2\}$. The neighborhood can be any configuration; for this example, the neighborhood is the cell itself and its four non-diagonal neighbors. At regular intervals, each cell in the layer is tested to see what state it will become in the next time step. This future state is determined by T, where the number of possible rules that enumerate every permutation of S and N is S^N . In the case of the example here, there are two states (forest/nonforest) and a cell's five neighbors (four other cells and itself) there are 2^5 different possible rules.

In theory, T is a fully enumerated set, but the simulation language allows users to specify rules that are more general. For example, one can specify that 'if the number of neighbors in a forested state is greater than three then change the state of the cell to forest'. By using compound conditional statements, complex ecological transition rules within a land-cover layer are created. Cellular automata can also refer to cells in other layers in order to incorporate other biogeophysical factors into state transition rules. This mechanism is codified as generalized cellular automata, which allows transition rules that do not reference adjacent neighborhood cells (Takeyama and Couclelis 1997). These other layers are treated as cellular automata where there are no transitions taking place (i.e., $T = \{\varnothing\}$) and the definition of states, S, is relaxed so that cells can take continuously varying values. These states do not need to be discrete since no state transformations take place. Other layers can then refer to these values in their transition rules. The ability to reference other layers is how cells affected by actor-agents are recognized as such and are accommodated by the environment. The simulation language also allows transition rules to probabilistically assess neighborhood states in order to make cell transitions more stochastic.

Despite the flexibility offered by the simulation system, only a few of the range of ecological relationships that exist in the region (Chpts. 2, 4 & 5), are included in this version of the SYPR IA model. Current rule sets focus on modeling secondary succession, pest invasions, and soil fertility degradation. Future directions are suggested by work on rangeland dynamics (Li and Reynolds 1997), species composition (Silvertown et al. 1992), forest succession (Hogeweg 1988), and a host of biological models (Ermentrout and Edelstein-Keshet 1993; Gronewold and Sonnenschein 1998). Generalized cellular automata show potential for modeling large-scale phenomena such as fires or hurricanes as provided by project ecological research.

These potential extensions beg the question: how complex should the model environment be? Having an essentially unchanging environment is in keeping with deforestation models in general, is simple to implement, and may be adequate for short term estimates of deforestation. A static environment is inadequate, however, if ecological processes such as succession are incorporated into observed actor decision making; the extent to which this is true is still unclear. Moreover, ecological processes at the species composition scale are difficult to specify given species-area relationships, recruitment limitations, and succession dynamics.

4.4 Actors

Key actors in the region are land managers, ranchers and smallholders whose production activities include swidden cultivation, agroforestry, modest logging, and market cultivation. A submodel probabilistically inserts actor-agents into the simulated study site at one of several hundred settlement locations according to exogenously given population densities. Actors are created in a given location, corresponding to a spatial layer cell, and make annual production decisions in the area around their respective locations. Actor-agents exist in a landscape represented by a collection of spatial layers contained either in the environment component or separately within the model. Actor-agents ‘know’ the surrounding landscape insofar as the agent software processes have access to layers representing spatial attributes such as land-cover type or distance to roads.

An actor makes decisions about production activities in order to maximize its utility. In the simplest case, the actor is limited to a single strategy, such as milpa (swidden agriculture), and the actor’s utility is tied to the amount and location of land under this form of cultivation. In a computational setting, this choice is played out on the landscape by the actor choosing which cells in the land-use layer will host production, a multicriteria evaluation problem (Eastman et al. 1995). An actor-agent determines the suitability, S of each grid cell c, for a given production activity l, S_{cl}:

$$S_{cl} = \sum_{i=1}^{m} w_{il} v_{il} * \prod_{j=1}^{n} b_{jl} \qquad (1)$$

as a function of a set of continuously varying spatial factors $V = \{v_1, \ldots, v_m\}$, a set of factor weights $W = \{w_1, \ldots, w_m\}$, and set of Boolean constraints $B = \{b_1, \ldots, b_n\}$. Factors, V are spatial layers of landscape attributes such as elevation. The importance for each use l of any given factor v_{il} is considered as a weight, w_{il}, which is determined by each agent, usually as a function of household resource profiles, in a manner considered below. Constraints B are layers in which cells have the value of either one or zero. When calculating S_{cl} for a given cell, if the corresponding cell in any Boolean layer is zero then S_{cl} equals zero, regardless of the values of the other constraints and factors. A land tenure institution-agent, for instance, can force an actor-agent to effectively ignore a whole swath of cells by marking their corresponding cells in a Boolean layer with the value of zero. Another use of Booleans is to limit the cells in which an agent can be interested in order to reflect some realistic situation, such as an agent not considering far-off cells, or in order to limit the actor's knowledge in keeping with the precepts of bounded rationality.

After determining S_{cl} for all cells C**,** an actor a selects a subset C^a to maximize the sum of cell suitabilities for the production strategy. Selecting these cells is an intensive integer programming problem. Heuristic solutions either rank cells by suitability and select a given number of cells for the agent or lie in compactness and contiguity algorithms. Finally, when multiple production strategies are possible, multicriteria evaluation is used to pick cells for competing objectives and then these cells are allocated according to a multiobjective land allocation process (Eastman et al. 1995).

Multicriteria evaluation and multiobjective land allocation are heuristic processes that approximate the utility offered to the actor by production activities. Ideally, an actor considers all permutations of cells in light of all pertinent decision making variables in order to maximize or satisfice utility. The present model, however, employs two computational short cuts. First, the number of cells selected for C^a is a stochastic function of demographic characteristics, agent production characteristics, or area under cultivation by households per the actor survey. Second, area of land subject to different forms of production is either chosen directly by the user in order to force the actor to consider a fixed set of alternatives or via an actor-agent decision rule that chooses a mix of productive activities based on their past utilities (after Arthur 1993). Better estimation of the amount and distribution of land dedicated to production in later versions of the SYPR IA model should be possible through tighter integration with the econometric modeling and understanding of social processes explored by other portions of this volume. A heuristic that better reconciles actor utility with measures of suitability is also feasible given the similarly of this problem to linear programming approaches to multiobjective land allocation (Diamond and Wright 1988).

The SYPR IA model focuses on several methods of determining suitability (S_{cl}) for actors. The first is a rule-based or heuristic process. For example, an agent can determine suitability as a probabilistic function of road proximity and then use a plot for three years before returning it to fallow. This rule is similar to those used by other dynamic spatial simulations (e.g., Southworth, Dale and O'Neill 1991; Gilruth, Marsh and Itami 1995). Heuristics are applicable to situations where data on few landscape attributes are available. When remotely sensed imagery is the only source of information, for instance, then heuristics based on cells classified as roads can work well. At the same time, heuristics are dependent on the spatial extent of their underlying data and areas where roads missing due to misclassification or cloud cover will not be modeled well.

Second, an agent can approximate S_{cl}, the best example being estimated parameter models programmed into agents. Actors calculate cell suitability in this manner with the simulation system's ability to process equations that combine both spatial and aspatial data to produce agent-specific suitability rankings. Examples of estimated parameter models are offered in Chapter 12, by other research informed by economics (Chomitz and Gray 1996; Kaimowitz and Angelsen 1998), or by efforts based on other inductive alternative-oriented methods (Ludeke, Maggio and Reid 1990; Mertens and Lambin 2000). The key value of this approach is the flexibility it offers to modeling actor decision making and the opportunity to draw on theoretical frameworks tied to empirical data. A wide variety of estimated models may be accommodated methodologically and the conceptual framework can host a variety of theoretical backdrops.

The third means of determining suitability is with genetic programs. These are software programs that evolve in a Darwinian fashion according to their fitness in meeting measurable goals (Koza 1992; Banzhaf et al. 1998). Genetic programs are also bounded rationality representations of decision making (Dawid 1999) and are therefore in keeping with the conceptual framework. Genetic programs evolve over simulation time to match surveyed actor land-use histories to spatial data and an array of decision variables from the actor survey. Alternatively, they can be used to empirically parameterize an equation linking a dependent land-use variable against independent variables in a manner functionally similar to that used to develop estimated parameter models noted above.

Genetic programs also offer a means of exploring agent communication because individual programs can be shared among agents. Similarly, genetic program evolution approximates learning over time (Beckenbach 1999). Genetic programming is a parallel means of searching complex solution spaces that is resistant to the pitfalls of local maxima and minima, discontinuities, and high dimensionality. This power comes at the cost of difficult implementation, interpretation, and the emerging nature of research linking social science perspectives on decision making to genetic programming representations (Chattoe 1998).

The SYPR IA model thus far has focused largely on smallholders. Introduction of other kinds of actors offers greater variety in patterning and verisimilitude. Use of varied agents may also be useful for modeling communication among actors at the cost of over-specification and higher data needs. Nonetheless, the actions of ranchers or the role of 'coyotes' in encouraging chile cultivation, for instance, would be interesting model extensions (Chpt. 10).

4.5 Results and Validation

The SYPR IA model runs as a series of Monte Carlo simulations whereby the model executes many runs, each one creating a simulated environment for each model year. For any given year, a set of simulated land-cover layers may be superimposed to create a single probabilistic image. Stochasticity enters the simulation through random queuing and placement of agents, heterogeneous actors, varying agent processes, and cellular automata transition rules. Snap-shots of ecological and human-induced land-use/cover changes are validated against observed data where the latter exists. They are used to understand the ramifications of land-use/cover change.

The simulation model creates a variety of outcomes that vary according to model configuration. Figure 1 illustrates the difference in spatial outcomes between three simulations of land-use/cover change in a small part of the study area (Manson 2000). Each simulation has an identical initial configuration save for different models of actor decision making. The environment is limited to cellular automata rules that mimic secondary succession in the absence of actor influence. Institutions, except for land-tenure, are ignored except as noted below.

The key difference between the three models lies in the agents' derivation of S_{cl} in terms of Eqn 1, and the following values are used in this calculation:
1) land use l, where actors are interested in a generic land use termed 'deforestation' that approximates deforestation for the purpose of cultivation;
2) factors V include a) distance to roads; b) distance to forest; and c) and initial land use (three forest types, agriculture, infrastructure);
3) constraints B denoting prohibited areas, which include a) land-tenure; and b) current infrastructure, such as roads and current settlement. Cells in each constraint layer are initialized to either zero or one, where zero denotes areas that are off limits to agent intervention; and
4) factor weights W are determined by each agent according to one of three methods.

In order to determine factor weights, the first set of actors uses a heuristic to probabilistically choose cells as 1) a function of distance to roads, where weights are asymmetrical sigmoidal probabilistic function that has its inflection point at 50 m and becomes asymptotic at 600 m; 2) distance to forest edge, where weights probabilistically decline linearly to zero at 500 m; and 3) time in current land use, where agents randomly leave cultivated land after one to three years and have a random chance of using recent secondary regrowth of 4-7 years in age. The second means of determining W is a logit regression model applied to the three spatial data factors V and accounting for B. This model is methodologically and conceptually similar to, but simpler than, the first decision making model presented by Chpt. 10 since it has fewer number of factors and it is run for separate, random subsets of agents within the modeled area. The area-wide "cutoff threshold" (pp. [*]) is replaced by the number of cells required by each agent to enumerate cells in C^a as specified in section 4.4 above. The third decision-making model is similar to the second in that it uses the same threshold for C^a, factors V, and constraints B. It differs in that it uses genetic programs that evolve over time. Each agents is stochastically imbued with three to twenty programs that represent different strategies to assess suitability. Each genetic program has a tree structure of varying size and shape composed of variables that correspond to factors V and coefficients that correspond to factor weights W. Genetic program trees are reinterpreted by the agents as an equation that each agent solves to create a suitability surface.

< Figure 1 >

While there are obvious visual differences between the results illustrated by the figure, the simulation system offers several means of determining the difference in accuracy between probabilistic projected land-use/cover change and that observed. The first is Kappa Index of Agreement (KIA), which calculates how closely two categorical maps compare in terms of location while accounting for chance in creating seemingly correct outcomes (Pontius 2000). The second and third tests are fractal dimension - a measure of patch complexity – and contagion – the extent to which land uses are clumped (O'Neill et al. 1988; Turner, Costanza and Sklar 1989). Both of these provide a sense of how well the simulated surfaces reproduce landscape pattern. The fourth test is a multi-resolution goodness of fit metric (Costanza 1989). This compares the relative number of correctly classified cells within a moving window at multiple resolutions in order to balance the roles that location and quantity play as a function of scale. Finally, a Monte Carlo uncertainty analysis measures of how confident one can be about the predicted land-use/cover change transition of any given location (Ogneva-Himmelberger 1998).

When a single validation test is applied to different model configurations, some are by definition more correct than others. Over multiple tests, however, a model configuration that scores well on one test can score worse on another, such as when one configuration projects quantity correctly while another better approximates the spatial patterning of observed land use. With the KIA, for instance, the genetic program model has the highest correspondence to observed data compared to the other two models. Under the contagion test, however, the linear model scores best, followed by heuristics and genetic programs.

Differences among model configuration over validation metrics can be an advantage. Consider two example applications that link the regional integrated assessment effort to the larger global environmental change agenda. Carbon release and sequestration are implicit in land-use/cover change transitions over time (Houghton 1993). Each cell in land-use/cover layers experiences one of a finite set of state transitions that has a distinct effect on the amount of carbon sequestered at that location. Since these cell transitions can be characterized in terms of their net effect on carbon sequestration (e.g., Masera, Ordonez and Dirzo 1997) then carbon regimes for the region are established by linking model results to the work of project researchers and ecological literature. Assuming there exists a relationship between cell transitions and quantity of sequestered carbon, model configurations that score well on quantity metrics are appropriate, genetic programs in the case above.

Conversely, the relationship between habitat fragmentation, expressed as cell-transitions, and biotic diversity may be more reliant on the spatiotemporal configuration and patterning of change. Although the relationship between ecological characteristics and landscape measures is not simple, if it is possible to link pattern to biotic diversity (after Li 2000; Plotkin et al. 2000), then the best choice is a model scoring well on fractal dimension or contagion, in this case the estimated parameter model.

4.6 Summary

The SYPR IA model extracts key land-use/cover change interactions in the southern Yucatán peninsular region. Institutions interact with actors by changing resource profiles and other spatial and aspatial variables impinging on actor decision making. The land tenure institution, for instance, manipulates land tenure grids that are referred to by agent processes that represent actor decision making. Actors make production decisions in order to maximize utility in order to change their resource profiles, such as grow food for consumption. These actions also change the state of cells in spatial layers that represents land use. As cellular automata transition rules rely on these cells to represent ecological phenomena, the modeled environment is affected by actors. The reverse also holds, however, as land used by an actor is subject to environmental events such as secondary succession that require continued actor intercession to counteract. As the simulation is iterative, there is constant interplay between actors and institutions, the effects of actor decision making on the environment, and the effects of environmental transitions on actor decision making.

5 Regional integrated assessment

The SYPR IA model addresses the distinct spatiotemporal patterns of land-use/cover change and the complexity of, and relationships among, socioeconomic and environmental factors. By facilitating the management of information and relationships, the SYPR IA model improves understanding of the causes and mechanisms governing land-use/cover change. Being an integrated assessment model, it allows examination of the impacts of policy alternatives that cannot be easily explored in reality.

As illustrated throughout the chapter, there remains much to do in terms of knowledge integration and model refinement. Better linkage of ecological research to simulation outcomes remains an ongoing challenge. This is especially true when linking model results to larger issues such as carbon sequestration and biodiversity. It is also necessary to better encapsulate more nuanced explorations of actors and institutions offered in other chapters. Agent definitions must be expanded to include a broader range of institutions, particularly those acting over less than the full spatial scale of the southern Yucatan peninsular region. A larger array of models can be introduced into the decision-making processes of agents representing actors, particularly in terms of stronger links to economics and econometrics. As the SYPR project moves into its second phase, the SYPR modeling efforts will improve and expand, seeking to offer a complex, integrated assessment model of land change.

Bibliography

Agarwal, C., G. L. Green, J. M. Grove, T. Evans and C. Schweik (2000). A review and assessment of land-use change models: dynamics of space, time, and human choice. Fourth International Conference on Integrating GIS and Environmental Modeling (GIS/EM4), Banff, Canada.

Alcorn, J. B. and V. M. Toledo (1998). Resilient resource management in Mexico's forest ecosystems: The contribution of property rights. Linking Social and Ecological Systems: Management practices and social mechanisms. Ed. F. Berkes and C. Folke. Cambridge, UK, Cambridge University Press: 216-249.

Arthur, W. B. (1993). On designing economic agents that behave like human agents. Journal of Evolutionary Economics 3(1): 1-22.

Banzhaf, W., P. Nordin, R. E. Keller and F. D. Francone (1998). Genetic Programming: An introduction. San Francisco, California, Morgan Kauffman Publishers.

Beckenbach, F. (1999). Learning by Genetic Algorithms in Economics. Computational Techniques for Modelling Learning in Economics. Ed. T. Brenner. Boston MA, Kluwer Academic Publishers: 73-100.

Benenson, I. (1999). Modelling population dynamics in the city: From a regional to a multi-agent approach. Discrete Dynamics in Nature and Society 3: 149-170.

Chattoe, E. (1998). Just how (un)realistic are evolutionary algorithms as representations of social processes? Journal of Artificial Societies and Social Simulation 1(3): www.soc.surrey.ac.uk/JASSS/1/3/2.html.

Chomitz, K. M. and D. A. Gray (1996). Roads, land use, and deforestation: A spatial model applied to Belize. The World Bank Economic Review 10(3): 487-512.

Conte, R., R. Hegselmann and P. Terna, Eds. (1997). Simulating Social Phenomena. Berlin, Springer-Verlag.

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

Dawid, H. (1999). Adaptive Learning by Genetic Algorithms: Analytical results and applications to economic models. Berlin, Springer.

Diamond, J. T. and J. R. Wright (1988). Design of an integrated spatial information system for multi-objective land use planning. Environment and Planning B 15: 205-214.

Eastman, J. R., W. Jin., P. A. K. Kyem and J. Toledano (1995). Raster procedures for Multi-Criteria/Multi-Objective decisions. Photogrammetric Engineering and Remote Sensing 61(6): 539-547.

Ermentrout, G. B. and L. Edelstein-Keshet (1993). Cellular automata approaches to biological modeling. Journal of Theoretical Biology 160(1): 97-113.

Galletti, H. A. (1998). The Maya forest of Quintana Roo: Thirteen years of conservation and community development. Timber, Tourists, and Temples: Conservation and development in the Maya forest of Belize, Guatemala, and Mexico. Ed. R. D. Primack, D. B. Bray, H. A. Galletti and I. Ponciano. Washington, DC, Island Press: 33-46.

Gilruth, P. T., S. E. Marsh and R. Itami (1995). A dynamic spatial model of shifting cultivation in the highlands of Guinea, West Africa. Ecological Modelling 79: 179-197.

Gronewold, A. and M. Sonnenschein (1998). Event-based modelling of ecological systems with asynchronous cellular automata. Ecological Modelling 108(1): 37-52.

Hogeweg, P. (1988). Cellular automata as a paradigm for ecological modelling. Applied Mathematics and Computation 27(1): 81-100.

Houghton, R. A. (1993). The role of the world's forest in global warming. World forests for the future: Their use and conservation. Ed. K. Ramakrshna and G. M. Woodwell. New Haven, CT, Yale University Press: 21-58.
Jager, W., M. B. A. van Asselt, J. Rotmans, C. A. J. Vlek and P. Costerman Boodt (1997). Consumer Behavior: A Modelling Perspective in the Context of Integrated Assessment of Global Change. Amsterdam, Netherlands, Rijksinstituut voor Volksgezondheid en Milieu (Dutch National Institute of Public Health and the Environment.
Kaimowitz, D. and A. Angelsen (1998). Economic Models of Tropical Deforestation: A Review. Jakarta, Indonesia, Centre for International Forestry Research.
Koza, J. (1992). Genetic Programming: On the Programming of Computers by Means of Natural Selection. Cambridge, MA, MIT Press.
Lambin, E. F. (1994). Modelling Deforestation Processes: A Review. Luxemburg, European Commission.
Li, B.-L. (2000). Fractal geometry applications in description and analysis of patch patterns and patch dynamics. Ecological Modelling 132(1 / 2): 33-50.
Li, H. and J. F. Reynolds (1997). Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: A combined Markov and cellular automaton approach. Scale in Remote Sensing and GIS. Ed. D. A. Quattrochi and M. F. Goodchild. New York, Lewis Publishers: 211-230.
Ludeke, A. K., R. C. Maggio and L. M. Reid (1990). An analysis of anthropogenic deforestation using logistic regression and GIS. Journal of Environmental Management 31: 247-259.
Manson, S. M. (2000). Agent-based dynamic spatial simulation of land-use/cover change in the Yucatán peninsula, Mexico. Fourth International Conference on Integrating GIS and Environmental Modeling (GIS/EM4), Banff, Canada.
Masera, O. R., M. J. Ordonez and R. Dirzo (1997). Carbon emission form Mexican forest: Current situation and long term scenarios. Climatic Change 35: 265-295.
Mertens, B. and E. F. Lambin (2000). Land-Cover-Change trajectories in Southern Cameroon. Annals of the Association of American Geographers 90(3): 467-494.
Myers, G. M. and Y. Y. Papageorgiou (1991). *Home economicus* in perspective. The Canadian Geographer 35(4): 380-399.
Ogneva-Himmelberger, Y. A. (1998). Exploring Empirical Diagnostic Modeling of Land-Use/Land-Cover: An example from Southern Yucatan Peninsular Region. Worcester, MA, Clark University.
O'Neill, R. V., J. R. Krummel, R. H. Gardner, G. Sugihara, B. Jackson, D. L. DeAngelis, B. T. Milne, M. G. Turner, B. Zygmunt, S. W. Christensen, et al. (1988). Indices of landscape pattern. Landscape Ecology 1(3): 153-162.
Ostrom, E. (1999). Institutional rational choice: An assessment of the institutional analysis and development framework. Theories of the Policy Process. Ed. P. A. Sabatier. Boulder, CO, Westview Press: 35-71.
Plotkin, J. B., M. D. Potts, D. W. Yu, S. Bunyavejchewin, R. Condit, R. Foster, S. Hubbell, J. LaFrankie, N. Manokaran, L. H. Seng, et al. (2000). Predicting biodiversity in tropical forests. Proceedings of the National Academy of Science 97(10850-54).
Pontius, R. G. J. (2000). Quantification error versus location error in comparison of categorical maps. Photogrammetric Engineering and Remote Sensing 66(8): 1011-1016.
Riebsame, W. E., W. J. Parton, K. A. Galvin, I. C. Burke, L. Bobren, R. Young and E. Knop (1994). Integrated modeling of land-use and cover change. BioScience 44: 350-356.

Rotmans, J. and H. Dowlatabadi (1998). Integrated assessment modeling. Human Choice and Climate Change. Ed. S. Rayner and E. L. Malone. Seattle, WA, Battelle Press. 3: 291-377.
Sen, A. (1981). Poverty and Famine: An Essay on Entitlement and Deprivation. New York, Oxford University Press.
Seror, A. C. (1994). Simulation of complex organizational processes: A review of methods and their epistemological foundations. Simulating Societies: The computer simulation of social phenomena. Ed. N. Gilbert and J. Doran. London, UCL Press: 18-40.
Silvertown, J., S. Holtier, J. Johnson and P. Dale (1992). Cellular automaton models of interspecific competition for space: The effect of pattern on process. Journal of Ecology 80(3): 527-534.
Simon, H. A. (1997). Behavioral economics and bounded rationality. Models of Bounded Rationality. Ed. H. A. Simon. Cambridge, MA, MIT Press. 1: 267-433.
Singh, I., L. Squire and J. Strauss, Eds. (1986). Agricultural household models: Extensions, applications, and policy. Baltimore, MD, Johns Hopkins University Press.
Smith, G. C. and D. S. Bull (1997). Spatial and temporal ordering of events in discrete cellular automata: An overview. Ecological Modelling 96(1/3): 305-324.
Southworth, F., V. H. Dale and R. V. O'Neill (1991). Contrasting patterns of land use in Rondonia, Brazil: Simulating the effects on carbon release. International Social Science Journal 130: 681-798.
Stern, P. C., O. R. Young and D. Drukman, Eds. (1992). Global Environmental Change: Understanding the human dimensions. Washington, DC, National Academy Press.
Takeyama, M. and H. Couclelis (1997). Map dynamics: Integrating cellular automata and GIS through Geo-Algebra. International Journal of Geographical Information Science 11(1): 73-91.
Torrens, P. M. and D. O'Sullivan (2001). Cellular automata and urban simulation: Where do we go from here? Environment and Planning B 28(2): 163-168.
Turner, B. L., II, S. Skole, G. Sanderson, L. Fischer, L. Fresco, R. Leemans and E. al. (1995). Land-Use and Land-Cover Change: Science/Research Plan. Stockholm and Geneva, IGBP/HDP.
Turner, M. G., R. Costanza and F. Sklar (1989). Methods to evaluate the performance of spatial simulation models. Ecological Modelling 48(1/2): 1-18.
Tversky, A. and D. Kahneman (1990). Rational choice and the framing of decisions. The Limits of Rationality. Ed. K. S. Cook and M. Levi. Chicago, IL, University of Chicago Press: 60-89.
Uhl, C. (1987). Factors controlling succession following slash-and-burn agriculture in Amazonia. Journal of Ecology 75(2): 377-407.
Walker, B. H. (1994). Landscape to regional-scale responses of terrestrial ecosystems to global change. Ambio 23(1): 67-73.
Yates, D. N. and K. M. Strzepek (1998). An assessment of integrated climate change impacts on the agricultural economy of Egypt. Climate Change 38(3): 261-287.
Young, O. (1994). International Governances: Protecting the Environment in a Stateless Society. New York, Cornell University Press.

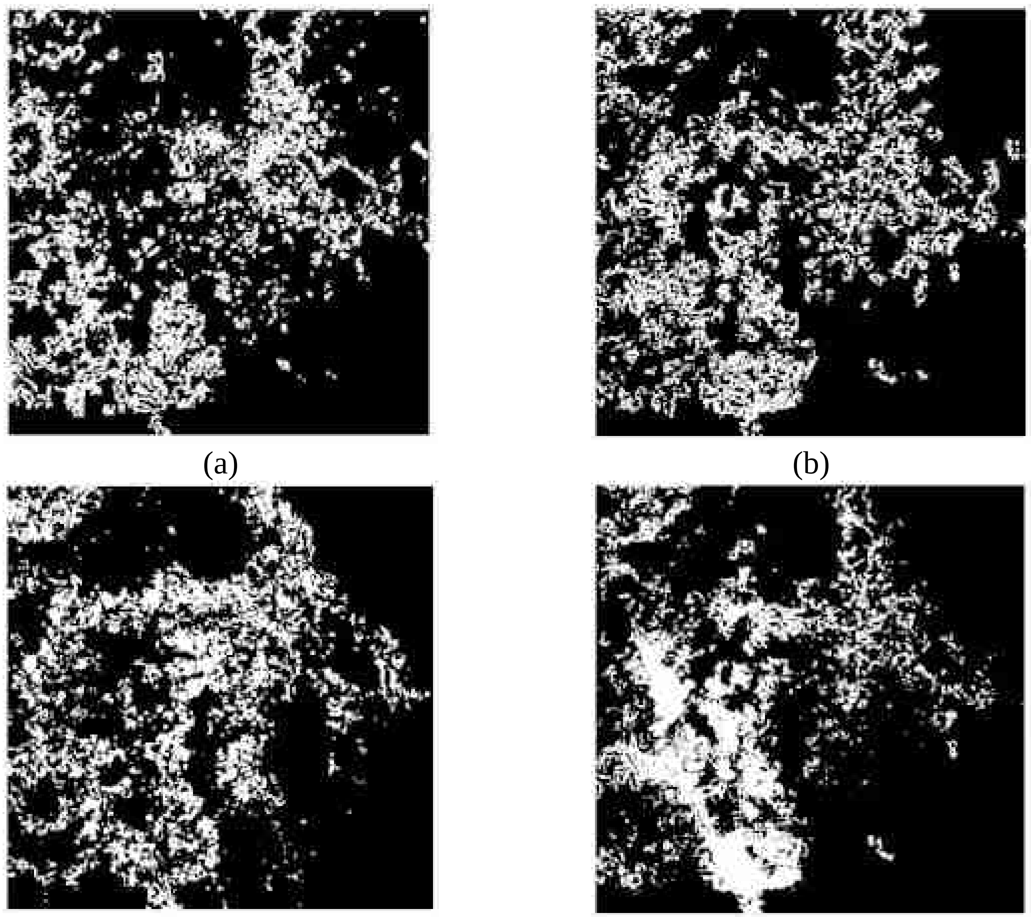

Figure 1. Actual change (a) and change projected by agents using (b) heuristics; (c) linear model; and (d) genetic programming.