In Proceedings of the Sixth International Conference on Cognitive Modeling, 142-147. Mahwah, NJ: Lawrence Earlbaum.

Comparing Agent-Based Learning Models of Land-Use Decision Making

Tei Laine (telaine@cs.indiana.edu)
Indiana University
Computer Science Department and the Cognitive Science Program
150 S. Woodlawn Ave.
Bloomington, IN 47405-7104 USA

Jerome Busemeyer (jbusemey@indiana.edu)
Indiana University
Psychology Department
1101 E. 10th Street
Bloomington, IN 47405-7007 USA

Abstract

An agent-based model, incorporating a small set of primarily agent-based variables, was designed to explain private land-use decision making. Agents are land-owners, who allocate their labor and land for different uses in regular time intervals. The goal is to understand what kind of spatial patterns emerge from different agent characteristics, and decision and learning mechanisms. Landscapes produced by two different learning models are compared to actual land-cover data. By calculating a set of spatial metrics from the simulated land-cover and comparing them to the metrics calculated from the actual land-cover data, the role of agent preferences for different land-uses is explored. The preliminary results suggest that the models capture relatively well the quantitative patterns of land-cover changes but they are poor in predicting the location of changes.

Land-use Decision Making

Land-use decision making is complex, multi-asset, real world decision task. The land-owner has to consider which activities (land-uses) he wants to implement on his land and decide where on that land to implement them. The decision maker's task is to find an effective way of using his assets — size and quality of land, technology, education and experience — in allocation of available resources — labor and land — to different uses. The number of factors to be considered range from the suitability of the land, dictated by various bio-physical variables, to the expected monetary or non-pecuniary returns from the uses. The optimal or good decision does not depend solely on the careful consideration of the above factors, but also on the decisions of neighboring owners and the use of their land.

General goal of modeling land-use decision making is to understand global environmental change and human role in bio-ecological systems. Specific goals are several; for instance, to explain the factors that drive people's land-use decisions, e.g., the role the preferences play in decision making; to understand the nature of spatial patterns emerging from learning and interaction of land-owners; to predict the location and rate of land-cover change (Schneider & Pontius, 2001); finally, the knowledge obtained in modeling studies may be used in environmental policy making (Daalen, Dresen, & Janssen, 2002) or to educate land-owners about sustainable land management practices.

Schneider & Pontius (Schneider & Pontius, 2001) model deforestation caused by residential development in the watershed in North-eastern Massachusetts. They predict the quantity and the location of deforestation between 1985 and 1991 using suitability maps calibrated with the real changes between 1971 and 1985. When comparing the model predictions of the rate of changes to the real changes 1985-1991, the model performs well, but fails to predict the sites of changes, which are relatively evenly spread over the study area. The model predicts suitability maps for residential development in the watershed, but does not account for individual decision makers' behavior.

Ligtenberg *et al.* combine multi-actor approach with cellular automata in modeling spatial planning (Ligtenberg, Bregt, & Lammeren, 2001). They apply the model with several types of decision making agents in a single-use framework to predict the spread of urban areas in Netherlands. Axtell *et al.* propose a multi-agent model of growth and collapse of Anasazi settlement in North-eastern Arizona from 1800 B.C. to 1300 A.D. (Axtell et al., 2002). Agents represent households whose primary source of nourishment is growing crops. By incorporating a number of agent-based variables (e.g., household size, age, nutritional needs, behavioral rules) to the archaeological and environmental data, agents' periodical choices for the locations of their residence and agricultural activities are simulated. The model closely reproduces the actual size of the population, its gradual migration north, and eventual abandonment of the area, while it occassionally fails to predict exact locations of household residences.

Hoffman, Kelley and Evans (Hoffman, Kelley, & Evans, 2002) and Kelley and Evans (Kelley & Evans, Under review) propose an agent-based agricultural land-use model and compare it to the actual data in order to explain land-cover changes in Southern Indiana. Individual agents represent households, who make decisions about how to use their land or whether to seek off-farm employment. They approach the decision process as a expected utility-maximizing behavior. They explore how spatial patterns of land-uses are influenced by land-owners' preferences and land suitabilities. Their finding is that land-owners preference heterogeneity plays at least as important role in emerging land-use patterns as the land suitability.

Figure 1: Changes in forest cover percentage in Indian Creek township from 1940 to 1993.

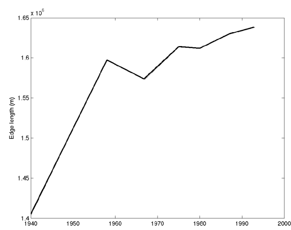

Figure 2: Changes in forest edge length in meters in Indian Creek township from 1940 to 1993.

Study Objective

Although this study concentrates on specific land-uses and a particular geographical area, the general principles used here are applicable to variety of cases of modeling private agricultural land-use decision making. The interest of this study is in the forest cover changes in Indian Creek township in South-central Indiana between the years 1940 and 1993. The available land-cover data indicates that there has been a significant increase in forest cover within the first 15 years of the study period and after that a modest, but gradual increase. The overall increase of forest cover is about 16%. The Figure 1 shows the monotonic increase of the percentage of forest in the area. Another metric that characterizes land-cover configuration is the length of the forest edge; the total length of the border between forest cells and non-forest cells. The non-monotonic increase in forest edge in meters in Indian Creek is shown in the Figure 2.

Theories in land and agricultural economics assume that land-use decision preferences are primarily formed by comparing expected financial benefits from different activities to the potential monetary costs of carrying out these activities. Koontz conducted an interview study among South-central Indiana land-owners in which he tried to explicate their motivations in land-use decision making (Koontz, 2001). The survey results suggest that non-monetary benefits also play a significant role, especially if the land is not the land-owner's primary source of income.

The goal of the current modeling study is to explain spatial patterns in South-central Indiana by postulating a small set of individual characteristics and learning mechanisms for decision makers and the payoff scheme that combines both monetary and non-monetary benefits. Simulating land-owners' yearly land-use decisions model makes predictions about the set of alternative land-uses, from which the changes of forest cover patterns are predicted.

Landscape Data

Three types of actual data are used in the modeling enterprise: forest-cover data, slope data and land ownership data. The time series of land-cover data were acquired from the historical areal photographs of the years 1939, 1958, 1967, 1975, 1980, 1987 and 1993. The slope data was extracted from the topographic maps and the ownership information from the historic parcel maps from the modeled period (Kelley & Evans, Under review). However, since the current model does not account for parcelization, only the parcel boundaries of the year 1957 are used.

These data were encoded into layered raster GIS (Geographic Information System) representation. In this representation the landscape is divided into a grid of cells of equal size and each layer records one type of information for each cell. In the current study the cell size is 50m $\times$ 50m and the landscape consists of 195 $\times$ 189 cells. A group of cells belonging to an individual landowner is called a *parcel*. There are 190 landowners in the modeled area, and their parcel sizes vary between 12 and 1936 cells, the average being 185 cells. For modeling purposes some additional landscape variables were introduced and encoded as layers in the GIS representation: soil quality and the number of years the cell has preserved its current use.

Agent-based Model

An agent-based model (see (Janssen, 2004) and (Parker, Manson, Janssen, Hoffman, & Deadman, 2003) for a review of agent-based models) is proposed that simulates land-owners' land-use decision making. The basic components of the model are the *landscape*, a rectangular area of land divided into cells corresponding to the actual landscape, and *agents*, the landowners whose primary source of income is the land they own. Like in the actual landscape data, the cell determines the resolution in which bio-physical information of the land is encoded. The cell is also the basic decision-making unit. Figure 3 represents the general model architecture.

Since the only available land-cover information is whether the cell is covered in forest or not, the initial landscape is constructed so that it allows more elaborate land-use decisions. This is accomplished by introducing two new land-uses — farming and abandoned land — and initializing the original landscape so that the shallow slope non-forest cells are marked as farmed cells, while the cells with steeper than average slope are marked as abandoned. After the initialization, the agents make two kinds of decisions each year; first, how to allocate their available labor between different activities, and secondly, on which parts of their land to apply these activities. The available activities are farming, off-farm employment, cutting trees and growing trees. The possible land-covers resulting from these activities are farmland, abandoned and forested land.

The specific agent actions are to decide whether to farm their non-forest land or seek off-farm employment, in which case they leave part of the land unused, and to decide whether to harvest trees or let them grow. The only information they use in these decisions is the payoff received from these activities in the past. In the farming decision the agents compare last year's hourly income from farming to the income from off-farm employment; if farming was more profitable, they decide to farm more

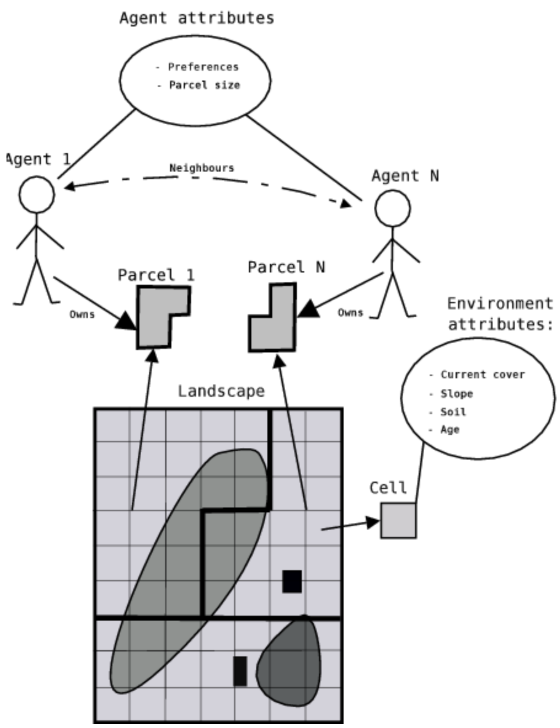

Figure 3: The basic components of the agent-based land-use decision model. Different shades of gray represent different land-uses.

in the future. In case farming was less profitable, they decide to work more off-farm and leave more cells unused. The cells left unused revert back to forest after certain number of years. Similarly, the decision between harvesting more trees or letting them grow depends on the payoffs from past decisions. The harvested areas start growing trees back. The tree growth follows a logistic function that is weighted by the average distance to other cells growing trees.

Agents

Agents represent private land-owner households, who make decisions about the use of their own land. They are assumed to have different individual characteristics; besides that their land varies in size, location and quality, the agents may vary in age, wealth, education, family size, experience, risk attitude, goals and subjective preferences.

Since we are interested in agent heterogeneity and its role in emerging land-use patterns, but most of the above data is not available, we postulate for each agent the subjective preference for farming (α_{farm}) and for growing trees, i.e., aesthetic enjoyment of having trees around (α_{trees}). Instead of estimating these parameters individually for each agent, the property size is used as the primary variable to infer the distribution of preference parameters. In other words, the agent j's preference for the use $i \in \{farm,\ trees\}$ is assumed to be a linear function of the property size s^j: $\alpha_i^j = b_i s^j + c_i$, where parameters b_i and c_i are estimated from the data. For the other land-uses α's are assumed to be 1.

Two additional parameters, fitted to the data, are the strength of externality effects for farming (ϵ_{farm}) and for growing trees (ϵ_{trees}). The externality effect in this context means either a positive or negative effect the cover of the neighboring cells have on the cover of the cell, or the effect of land-uses across parcel borders. For instance, the cell may be more likely to grow trees if, there are trees on the neighboring cells, or the agent may be more likely to farm the part of its land that borders neighbor's farmed land rather than forest. The strength of the externality effect can be conceived as agent's preference for large continuous areas of homogeneous land-use rather than scattered pattern of different uses.

Learning and Decision Strategies

Two different types of decision making agents are implemented and compared in the same land-use decision task. Both types of agents use the reward received from past decisions to infer if the decision was successful or not. The reward combines both monetary and non-pecuniary profits, and depends on the number and location of cells allocated to the activities, the suitability of these cells for the activities, unit returns from the activities and agent's subjective preferences for the activities. The payoff from the land-use i ($\in$ {*farming, off-farm employment, cutting trees, growing trees*}) is calculated by:

$$P_i = \sum_j \alpha_i \rho_i \sigma_i^j \delta_i^j \epsilon_i^j, \qquad (1)$$

where j goes over all the agent's cells that are allocated to use i, α_i is the general preference for the land-use i, ρ_i is the unit return from the use i [1], σ_i^j is the suitability parameter for the use i on the cell j, δ_i^j is the average distance from the cell j to other cells with use i, and finally ϵ_i^j is strength for externality effect of use i on cell j. Finally, the P_i's are summed up to get the agent's total payoff.

The suitability represents a discount factor that penalizes land-uses in unfavorable conditions. For farming and tree harvesting it depends linearly on soil quality and inversely on slope: $\rho = \frac{soil}{1+slope}$, where the soil quality degrades linearly over time with continuing use of the same kind. For other land-uses the suitability is assumed to be 1.

Learning and Resource Allocation Rieskamp *et. al* (Rieskamp, Busemeyer, & Laine, 2003) compared two learning algorithms to data obtained in the resource allocation laboratory experiment, in which subjects repeatedly allocated a fixed amount of money between three assets with either certain or uncertain returns. The payoff function was designed so that it had two maxima;

[1] The crop prices are aggregated from corn and soybean prices per produced unit, while timber prices are aggregated from certain hardwood species commonly harvested in the modeled time period. The minimum wage is used as an indicator of the changes in off-farm income rates in the modeled period.

two local and one global. The goal of the study was to find out, if the subjects were able to learn the location of these maxima from the numerical feedback they received after each allocation.

The two algorithms are based on different learning principles. The local adaptation model (LOCAD) makes small adjustments to the allocation depending on the feedback received from the trial one time step before, while the global search model (GLOS) keeps track of all possible allocations and probabilistically samples the areas in allocation space, which have proved profitable earlier.

The models' performance was compared to individual subjects' performance along various dimensions, e.g., the average allocation across trials, and the success in increasing payoff through learning. The main result is that the LOCAD model explains subjects' learning behavior better than the GLOS.

These promising experimental results give us a guideline to how to approach modeling of decision making and learning behavior in real-world resource allocation task that is much more complex. The proposed models for land-use decisions do not assume expected utility maximizing behavior, but the agents are expected to adapt to changes in the environment, for instance, to degradation of soil quality or changes in market prices, and adjust their decisions accordingly. The two learning models, discussed next, have different adaptation mechanisms.

Local adjustment learning agent considers the payoff received from one time step before and makes small adjustments to labor allocation in the next time step. The amount of labor it allocates to certain use reflects the outcome of the previous year's allocation so that, if the allocation was success, the new allocation is adjusted upwards, and if it was a failure, it is adjusted downwards. The success and failure are defined relative to the payoffs received from other uses. The decision is not deterministic; there is a slight chance that an opposite decision is made. For instance, even if the farming was more profitable than off-farm employment in the last period, the amount of farmed land may be decreased rather than increased in the next period.

The local adjustment learning agent uses *mixed* learning and decision strategy; each agent starts as a *individual learner*, i.e., it only considers its own payoff from the previous decision, but changes its type to *social learner*, if the decision turns out to be poor. This learner type considers the payoffs of all the agents when making decisions. Depending on the success or failure of the subsequent decisions the agent changes its type between these two types. Success is defined as an increase in the total payoff in two consecutive years, while failure is defined as decrease in the payoff.

The local adjustment agent is assumed to track small changes in market prices or environmental conditions relatively well. However, it is prone to get stuck in local minimum or in case of larger yearly changes, to oscillate between two allocations.

Experience-based agent has a perfect memory of all labor allocations it has tried before and the payoffs they have resulted in. Instead of making a small adjustment to the previous labor allocation, it samples its memory of allocations and chooses the one that produced the highest payoff. The choice is stochastic, proportional to the payoff received from the allocation, so that the allocations with higher payoffs are more likely chosen. After finding a good allocation from the memory, in order to assure a some level of exploration, the agent adjusts it slightly, i.e., it either increases or decreases the percentage of labor allocated to one or more activities.

This agent is assumed to repeat the early decision several times in the beginning, but to shift to more radical changes later on. It is not sensitive to changes in the environment, but chooses good allocations from the past, which may not be profitable in a new situation.

Results

The models' prediction accuracy is assessed by fitting the free parameters to the actual data in order to minimize the sum of square error (see equation 2) between the spatial metrics produced for the actual landscapes and the simulated landscapes. The spatial metrics are landscape-level forest cover percentage and forest edge length.

The model fit is assessed with respect to the null model, the landscape before the simulated period, to test if the proposed model predicts the land-cover changes better than the model that assumes no changes occurred (Kelley & Evans, Under review). Since the changes in the actual data are relatively small, it is expected that the initial landscape produces pretty accurate predictions.

$$\begin{aligned} SSE &= \sum_i (y_i - \hat{y}_i)^2 \\ SSE_{null} &= \sum_i (y_i - \hat{y}_{null})^2 \\ R^2 &= 1 - \frac{SSE}{SSE_{null}} \end{aligned} \tag{2}$$

Figures 4 and 5 present forest cover percentage and edge length changes for real and modeled landscapes, when preference parameters and strengths of externality effects are estimated as described above. The experience-based model fails to predict the data both qualitatively and quantitatively. The local adjustment model is relatively accurate in forest cover percentage, but is not able to account for the non-monotonic trend in the increase of the forest edge length.

Since both models use the same tree growth function, the experience-based model either harvests too many trees or farms extensively. In order to test the former hypothesis, one more parameter is added and fitted to the data: the frequency in which the logging decision is made. This parameter does not have any effect in the

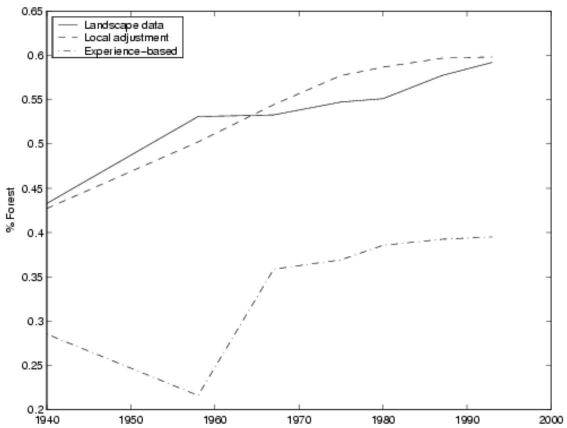

Figure 4: Changes in forest cover percentage over time. The parameters fitted are the linear preference parameters and the strength of externality effects.

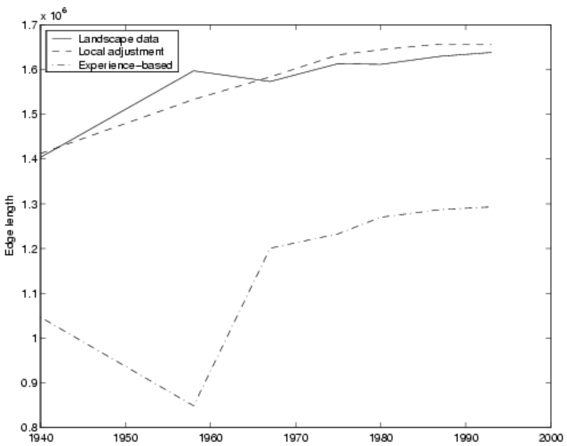

Figure 5: Changes in edge length over time. The parameters fitted are the linear preference parameters and the strength of externality effects.

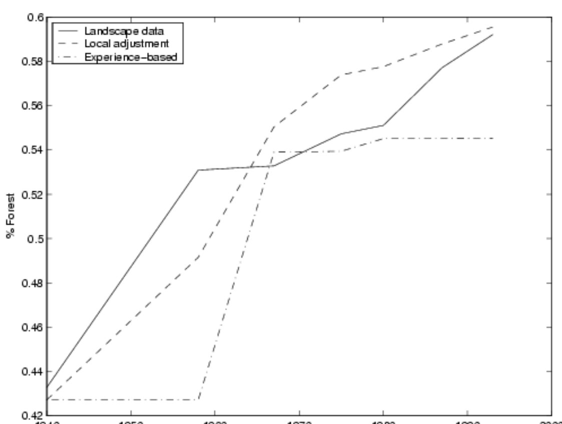

Figure 6: Changes in forest cover percentage over time. In addition to the preference parameters, the frequency of logging decision is fitted to the data.

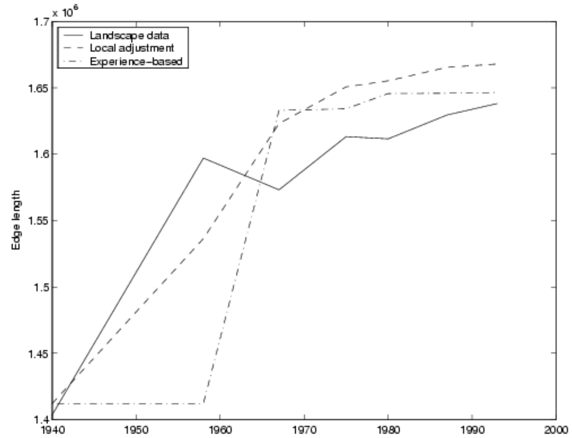

Figure 7: Changes in edge length over time. In addition to the preference parameters, the frequency of logging decision is fitted to the data.

performance of the local adjustment model, but clearly changes the other model's behavior in the beginning. Although it is still unable to predict the early significant increase in the cover, it captures some of the trend later. The results of these simulations are presented in Figures 6 and 7.

All the above results are from simulation runs in which the local learning model uses a mixed learning and communication strategy. In these runs a little more than half of the agents eventually converge to social learners. In order to test the effect of agent communication the possibility of changing the type was disabled and the agents were all set to either individual or social learners.

The trends of forest cover changes for the real landscape and the three types of local adjustment models are presented in the Figure 8[2]. If the agents share the information of everyone's payoffs, the model predicts an excess of forest cover starting early in the period, while the agents learning only from their own experience predict too conservative changes. As expected, the more complex model, with one extra parameter fitted to the data, predicts the data more accurately.

The model performance was also tested by comparing the metrics on the individual decision-maker level, i.e., forest cover percentage and edge length were calculated for each parcel separately. The preliminary results of these runs suggests that neither of the models does better than the null model in predicting individual parcel metrics. Actually, both models do slightly worse. They fail to explain why certain land-owner parcels have regrowth of forest while others do not. This may be because the same parameters are used to predict all the stages of forest-cover increase. However, the parameters get mostly calibrated to the early period in which most of the change occurs and to the parcels that change

[2]The pattern of the edge length change is pretty similar, and therefore not presented here.

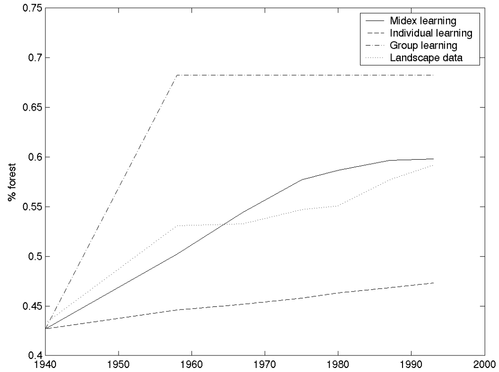

Figure 8: Changes in forest cover percentage over time. Results for different types of local adjustment models.

the most. Consequently, they are unable to capture the smaller changes occurring later.

Discussion

Two agent-based learning models, the local adjustment model and the experience-based model, were proposed to explain private land-owners' agricultural land-use decision making. The agents differ in the amount of information they use in decisions; the local adjustment learning agent utilizes the payoff information of only one time step behind, while the experience-based learning agent uses the information of all past decisions.

The local adjustment model predicts the increase in forest cover percentage relatively accurately, while the experience-based model fails to account for the significant increase in the beginning of the modeled period. This may be a direct effect of the strategy used; it has very little experience at early stages, so it basically repeats the same decision, or almost the same decision, over and over again. Neither of the models is able to predict the non-monotonic increase in the edge length.

The models do not do very good job in predicting the locations of changes on the parcel level. This may be because of the parameters do not capture the essential agent characteristics, i.e., the size of the parcel is not a good indicator of land-use preferences. However, Koontz's survey among Indiana land-owners indicated that there is a relationship between the parcel size and motivations for land-use activities (Koontz, 2001): on larger parcels the activities were associated with monetary benefits, while on smaller parcels the motivations were driven by non-monetary benefits. In the current study the aesthetic enjoyment of having trees around is assumed to be driven by non-monetary benefits. The model does not distinguish between financial and nonpecuniary motivations, and therefore the size of the parcel may not have a real effect in the preferences of different activities.

Finally, unlike the two learning models designed to explain the laboratory resource allocation behavior, the models designed for the real-world resource allocation task make both qualitatively and quantitatively different predictions when the amount of changes is considered. The models do not differ significantly in their accuracy of predicting the location of changes; they both fail equally.

Acknowledgments

This research was supported by NSF grant SES-008351. The authors are grateful for the helpful comments received from the members of the CIPEC, Biocomplexity project throughout the modeling process.

References

Axtell, R., Epstein, J., Dean, J., Gumerman, G., Swedlund, A., Harburger, J., Chakravarty, S., Hammond, R., Parker, J., & Parker, M. (2002). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. In *Proceedings of the National Academy of Sciences of the U.S.A.* (Vol. 99: Suppl. 3, p. 7275-7279).

Daalen, C. E. van, Dresen, L., & Janssen, M. A. (2002). The roles of computer models in the environmental policy life cycle. *Environment Science & Policy*(238), 1-11.

Hoffman, M., Kelley, H., & Evans, T. (2002). Simulating land cover change in South-central Indiana: Complexity and ecosystem management. In M. E. Janssen (Ed.), *The theory and practice of multiagent systems.* Edward Elgar, Cheltenham, U.K.

Janssen, M. A. (2004). Agent-based modelling. In J. Proops & P. Safonov (Eds.), *Modelling in ecological economics.* Edward Elgar.

Kelley, H., & Evans, T. (Under review). The relative influences of land-owner and landscape heterogeneity in an agent-based model of land-use. *Journal of Economic Dynamics and Control.*

Koontz, T. M. (2001). Money talks — But to whom? Financial versus nonmonetary motivations in land use decisions. *Society and Natural Resources*, *14*, 51-65.

Ligtenberg, A., Bregt, A. K., & Lammeren, R. van. (2001). Multi-actor-based land use modelling: spatial planning using agents. *Landscape and Urban Planning*, *56*, 21-33.

Parker, D., Manson, S., Janssen, M., Hoffman, M., & Deadman, P. (2003). Multi-agent system models for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers*, *93*(2), 316-340.

Rieskamp, J., Busemeyer, J., & Laine, T. (2003). How do people learn to allocate resources? Comparing two learning models. *Experimental Psychology: Learning, Memory & Cognition*, *29*(6).

Schneider, L. C., & Pontius, R. G. (2001). Modeling land-use change in the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, *85*, 85-94.