Agent-Based Model of Land-Use Decision Making*

Tei Laine
Indiana University
telaine@cs.indiana.edu

Jerome Busemeyer
Indiana University
jbusemey@indiana.edu

Abstract

An agent-based model, incorporating a small set of primarily agent-based variables, was designed to explain land-use decision making. Agents are land-owners, who allocate their labour 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 the learning agent model are compared to actual land-cover data. By varying the parameter estimation schemes and the spatial metrics calculated from the simulated land-cover and the actual land-cover, the role of agent preferences for different land-uses is explored. The preliminary results suggest that the model captures relatively well the quantitative patterns of land-cover changes but it is poor in predicting the location of changes.

Contact:
Tei Laine
Computer Science Department
Indiana University
Lindley Hall 215
150 S. Woodlawn Ave.
Bloomington, IN 47405-7104

Tel: +1 812 855-8702
Fax: +1 812 855-4829
Email: telaine@cs.indiana.edu

Keywords: Agent-based modeling, land-use, decision making

*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.

Agent-Based Modeling of Land-Use Decision Making

Tei Laine and Jerome Busemeyer

Introduction

Land-use decision making is complex, multi-asset, real world decision task. The land-owner has to consider which activities (land-uses) he/she wants to implement on one's land and decide where on that land to implement them. The decision maker's task is to find an effective way of using his/her assets — size and quality of land, technology, education and experience — in allocation of available resources — labour and land — to different uses. The number of factors to be considered range from the suitability of the land for different uses, 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.

In this paper we propose an agent-based model of land-use decision making for explaining forest cover changes in Indian Creek township in South-central Indiana between the years 1940 and 1993. Although this study concentrates on specific land-uses and a particular geographical area, the general principles used here are applicable to variety of cases. The available land-cover data indicates that in Indian Creek 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 and other land-uses. The non-monotonic increase in forest edge in meters in Indian Creek is shown in the Figure 2.

The goal of the study is to explain these spatial patterns by postulating a small set of individual characteristics and a simple learning mechanism for decision makers. Simulating their yearly land-use decisions model makes predictions about the set of alternative land-uses, from which the changes of forest cover patterns are predicted.

Agent-based Model

The basic components of the agent-based model (see (Janssen, 2004; Parker, Manson, Janssen, Hoffman, & Deadman, 2003) for a review), designed to simulate land-owners' land-use decision making, 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. The cell determines the resolution in which bio-physical information of the land is encoded, and it is also the basic decision-making unit.

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 land-owner maps from the modeled period (Evans & Kelley, In press).

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 — farmland and abandoned land[1] — 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 labour 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 land and forested land.

Agents represent the 190 private land-owner households on the modeled area, who make decisions about the use of their own land. They are assumed to have different individual characteristics; besides that their

[1] The land that is not used for financially profitable activities.

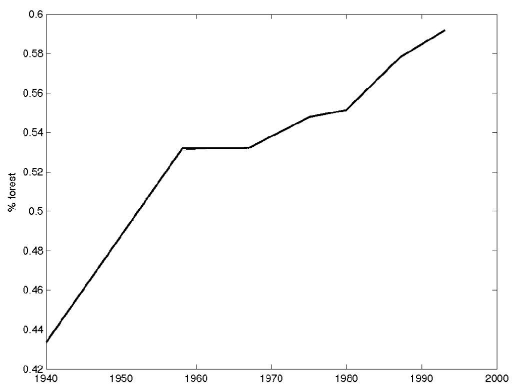

Figure 1: Changes in forest cover percentage in Indian Creek township from 1940 to 1993.

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

land varies in size (between 12 and 1936 cells, the average being 195 cells, in the actual data), location and quality, the agents may vary in age, wealth, education, family size, experience, risk-attitude, goals and subjective preferences.

The only information the agents use in these decisions is the rewards received from the previous activities, which tells if they were 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:

$$P_i = \alpha_i \rho_i \sum_j \sigma_i^j \delta_i^j \epsilon_i^j, \tag{1}$$

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, σ_i^j is the suitability parameter for the use i on the cell j (derived from the slope and soil information), 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 (see explanation later) of use i on cell j. Finally, the P_i's are summed up to get the agent's total payoff.

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. 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 in the future. In case farming was less profitable, they decide to work more off-farm and leave more cells to unused. The cells left abandoned 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.

Since we are interested in agent heterogeneity and its role in emerging land-use patterns, but data on agent characteristics is not available, we postulate for each agent the subjective preference for farming (α_{farm}) and for growing trees (α_{trees}), i.e., for aesthetic enjoyment of having trees around. Three estimation schemes for these parameters are suggested. In the first one, instead of estimating preference parameters individually for each agent, the property size is used as the primary factor from which the distribution of agent preferences are inferred. 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, i.e., $\alpha_i^j = b_i s^j + c_i$, where parameters b_i and c_i are estimated from the data. For the other land-uses α's are assumed to be 1.

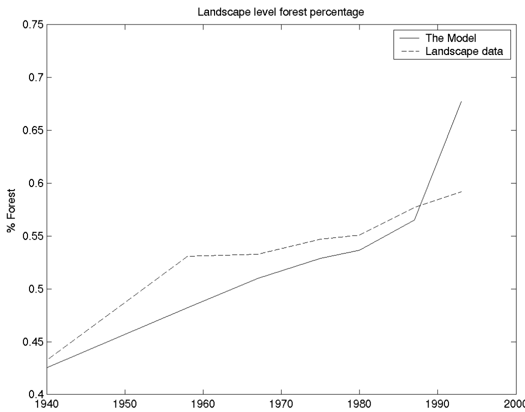

Figure 3: Actual and modeled changes in forest cover percentage over time. No parameters are fitted.

Figure 4: Actual and modeled changes in forest edge length in meters over time. No parameters are fitted.

In the second scheme, when calculating the preference for the land-use i, the proportion of the property assigned to use i is used, rather than the size of the whole property. This scheme assumes that the current land-cover has an effect on agent's preferences, which in turn influence the future decisions and consequent land-covers. The preferences for alternative uses are also assumed complementary, so that $\alpha_{off} = 1 - \alpha_f$ and $\alpha_{cut} = 1 - \alpha_{trees}$.

In the third scheme, the agents are classified into three categories according to their preferences. In this scheme both preference parameters α_{farm} and α_{trees} may take values 0.1, 0.5 and 0.75. The agents change their preference type after each unsuccessful decision, i.e., a decision that has a worse outcome than the previous one. Again, in this scheme the preferences of alternative activities are assumed complementary.

Simulations

Assessing the performance

The spatial metrics used to assess model's prediction accuracy are forest cover percentage and forest edge length. Several tests are run, varying the parameter estimation scheme and the resolution on which these metrics are generated. The free parameters are fitted to the actual data at six data points from 1958 to 1993 in order to minimize the sum of square error ($SSE = \sum_i (y_i - \hat{y}_i)^2$) between the spatial metrics produced for the actual landscapes and the simulated landscapes. The model fit is assessed with respect to the null model, the landscape before the simulated period, by calculating $R^2 = 1 - \frac{SSE}{SSE_{null}}$, where $SSE_{null} = \sum_i (y_i - \hat{y}_{null})^2$, 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.

In addition to the free parameters discussed above, two other parameter sets are fitted: first, *the rate of change* which approximates the number of cells whose use is changed from year to year, and second, 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 property borders.

Results

The baseline case is to run the model without fitting any parameters. The changes in forest cover percentage and forest edge length produced on the landscape level by this model are plotted in the Figures 3 and 4, respectively. The model tracks the trend in forest cover percentage relatively well in the beginning, but

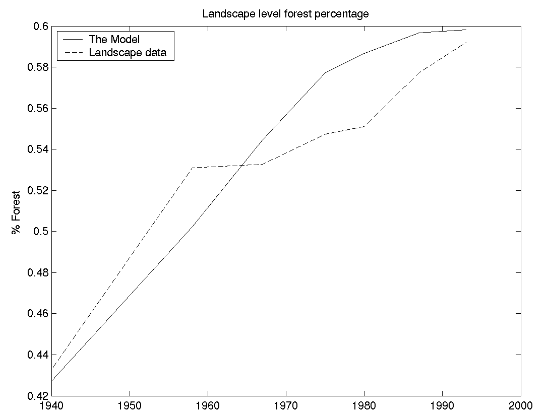

Figure 5: Actual and modeled changes in forest cover percentage over time. Linear preference parameters and strength of externality effects fitted.

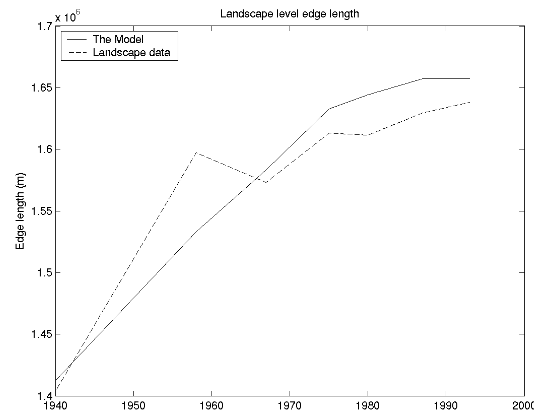

Figure 6: Actual and modeled changes in forest edge length over time. Linear preference parameters and strength of externality effects fitted.

overshoots in the end of the period. However, it cannot account for the non-monotonic increase in the forest edge length.

The next step is to introduce the free parameters. Figures 5 and 6 present landscape level forest cover percentage and edge length changes for real and simulated data, when preference parameters (from property sizes) and strengths of externality effects are estimated. The model is relatively accurate in forest cover percentage, but is still not able to account for the non-monotonic trend in the increase of the forest edge length.

The model predicts relatively well the changes spatial patterns on the landscape level. Next we focus on how the model accounts for changes on individual land-owner level. In order to accomplish this, the spatial metrics are generated for each agent's property separately. Several parameter fitting schemes are compared: first, the preferences for activities are assumed to be a function of the size of property used for the activities, second, the rate of change is introduced to the fitted parameters, and finally, the complementary preference parameters are incorporated. The agent level spatial statistics (R^2's) are presented in the Figure 7 when forest cover percentage was fitted on individual agent level.

It is evident from the R^2 values that the model is not very successful in explaining the land-cover changes on the individual agent level. Therefore, another test is conducted to more rigorously evaluate model's performance in predicting location of changes; the landscape is divided into nine equal-size areas (of 65×63 cells), and the spatial metrics are produced for each of the divisions. Also the third preference parameter scheme is used; since the model is not predicting spatial metrics on agent level, but instead for divisions of the landscape, individual agents characteristics are not assumed important. Instead, the attempt is made to categorize the agents into different categories according to their preferences to find out if land-use preferences coincide with the characteristics of the landscape, and thus lead into different patterns of land-cover changes on different divisions. The results, presented as time series of R^2's are presented in the Figure 8.

The model predicts forest cover changes accurately on three out of nine divisions (the R^2's are close to 1 in the end of simulated era), while the model performs worse than the null model on three divisions (i.e., their R^2's are below zero).

Discussion and Future Work

Hoffman, Kelley and Evans (Hoffman, Kelley, & Evans, 2002) and Kelley and Evans (Kelley & Evans, Under review) propose an agent-based land-use model to explain land-cover changes in Southern Indiana. They approach the decision process as a expected utility-maximizing behavior rather than a learning pro-

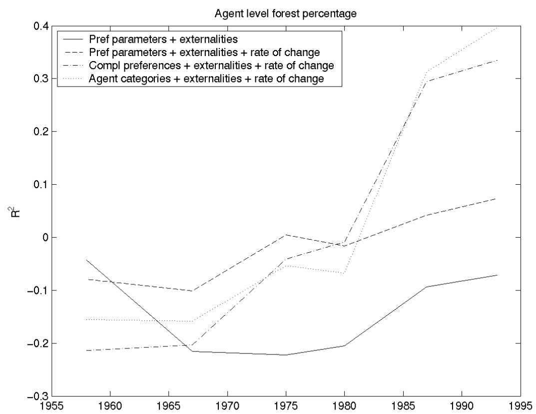

Figure 7: R^2's for different parameter estimation schemes when forest cover percentage is fitted on the individual decision-maker level.

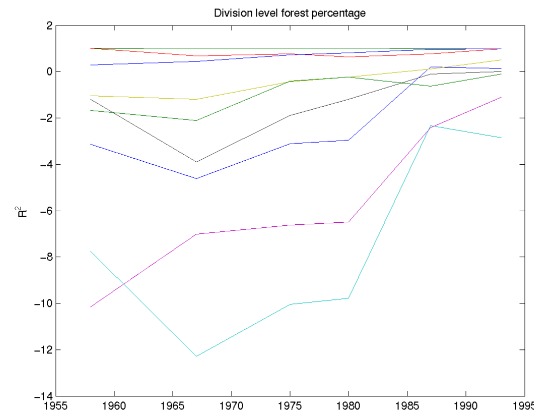

Figure 8: R^2's for nine landscape divisions when forest cover percentage is fitted on the division level.

cess. 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. Evans and Kelley (Evans & Kelley, In press) study the influence of different spatial resolutions on the model's performance. The primary finding is that the coarser the resolution, the worse the fit to the actual landscape.

One purpose of the current study was to test different schemes giving rise to the agent heterogeneity, without undermining the model generality, while the other purpose was to assess the model performance on different levels of resolution. Instead of fitting parameters for each agent, more general trends in agent characteristics were explored. The first main finding was that the property size is not very good indicator of individual preferences for different land-uses. The second finding was that the prediction accuracy (measured in R^2) is worst for the divisions of the landscape that have the least amount of change. Finally, the agents' classification according to their preferences shows slight support for the hypothesis that the agent preferences play an important role in land-use decisions. For the farming preferences $\alpha_{farm} \in \{0.1, 0.5, 0.75\}$, the number of agents in each of the classes are 36, 128 and 26, respectively, and for the aesthetic preferences α_{trees} the numbers are 1, 185 and 4. In the first case, there seems to be a slight tendency not to prefer farming over off-farm employment, while in the latter case, the agents are relatively indifferent between enjoying trees around and making money out of them. The next step in this modeling enterprise is to incorporate more elaborate learning mechanisms and agent interaction.

References

Evans, T. P., & Kelley, H. (In press). Scale issues in agent-based models of landcover change. *Journal of Environmental Management*.

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 multi-agent 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*.

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.