Position paper for UCSB workshop "Agent-based modeling of complex spatial systems"
Dawn Parker, George Mason University
Feb. 19, 2007

Early successful examples of spatial agent-based models were often highly abstract, and were used to demonstrate how particular macro-scale outcomes could emerge as the result of decentralized interactions of autonomous decision-making agents. Examples include the demonstration of segregation provided by Schelling's early models (Schelling 1971) and the emergence of power-law distributions of wealth from Epstein and Axtell's Sugarscape model (Epstein and Axtell 1996). Such models had a very simple role: to extend theory by demonstrating that a particular set of behaviors and interactions could *generate* a particular macro-scale or emergent outcome (Epstein 1999). As the discipline of spatial-agent based modeling matures, an increasing number of highly empirical models (in the style of "Cell 4" models as described by Parker, Berger, and Manson (2002) are being constructed. These models are often developed to support policy analysis, and are motivated by a belief that models that include complex dynamics and interactions are likely to produce "better" results that models that omit those dynamics. While my discussion focuses on models of human-environment interactions (Bousquet and Le Page 2004; Grimm and Railsback 2006; Janssen 2003; Parker et al. 2003), such empirical spatial agent-based models are also being developed in other areas of social science, including epidemiology, sociology, and political science (Castle and Crooks 2006).

Many scholars have voiced concern that such highly empirical models of complex systems are in danger of being black boxes in which the relationships between inputs and outputs, and even the behavior of the model itself, are not well understood. Some argue that for these reasons agent-based modeling is appropriate only for highly abstract demonstration of theoretical outcomes, and that the push to increase complexity in these empirical models may run counter to scientific principles of parsimony in modeling. Fundamental questions are raised about where and how agent-based modeling fits into the scientific method. In the contrary view, others argue that the policy environment is complex and therefore demands development of complex models, and that policy questions we now face, such as potential contributions of land-use and land cover change to global climate change, and the potential transmission paths of pandemics, are so pressing that any and all practical approaches to obtaining answers must be tried. A separate thread argues that spatial agent-based models may contribute to the development of an interdisciplinary theory of spatial social science. Yet, whatever one's view of the appropriate role for ABM in spatial social science, tensions exist between the need to simply build better empirical models (that better represent real-world drivers and complex interactions) and goal of building theoretical frameworks that have a more formal relationship to the scientific method, which could then contribute to building integrated theories of spatial social science. The first goal demands more details and realism; the second demands a high-level abstract framework that is generalizable across case studies and perhaps even realms of social science.

The questions lain out in this position paper attempt to steer the debate from what sorts of spatial agent-based models we should be doing to how we might do modeling—both theoretical and empirical. These questions also aim to focus on how agent-based modeling can be more closely connected to the scientific method, at both theoretical and empirical ends of the spectrum. While empirical models face additional challenges of parameterization, calibration and validation, models constructed at either end of the spectrum share a goal of incorporating as much complexity is needed to represent the problem under study, but not too much. And both types of models ultimately seek to make substantive contributions to science.

Question 1: How to identify a just-sufficient level of detail for the model? The idea that the goal of agent-based models is to produce macro-scale emergent patterns through micro-level behaviors and interactions is well established. This idea can, however, provide insufficient guidance to create a parsimonious model. A slight refinement of the the model question to "How do particular agent behaviors and interactions at a micro-level produce patterns observed at a macro-level?" can help focus efforts. Grimm et al. (2005) suggest that at a minimum, your model must embed processes of sufficient complexity such that it produces the macro-scale patterns of interest. Further, these patterns must of course be non-trivially produced through micro-scale interactions, rather than be obvious consequences of the rules specified at an agent level. Initially, of course, a modeler will introduce micro-level behaviors and interactions that she believes are linked to macro-scale patterns. In relation to land-use modeling, I characterize these behaviors and interactions into spatial, temporal, and behavioral complex drivers (Parker 2007). Grimm et al. (2005) also note that because multiple processes can produce the same single observed pattern, often multiple observed output patterns are needed to distinguish between competing process models. In short, identifying a minimum amount of complexity requires a clear initial hypothesis that links a micro-scale exogenous driver (which could be a model parameter, an agent behavior, or the structure of agent interactions) to a macro-scale emergent pattern outcome. A focus on linkages between hypothetical drivers and outcomes, rather than simply on reproduction of observed patters, could reunify the process of model building with the traditional scientific method. Further, for highly empirical models designed for policy analysis, it brings a focus to policy levers whose values may be modified for scenario or sensitivity analysis, thus encouraging the development of the model to be driven by its intended use.

Question 2: What are potential roles for statistical analysis in spatial agent-based models? In Parker et al. (2003), statistical models were characterized very much as substitutes for agent-based models. I have modified that view substantially since that paper was first written. I now see statistical methods as an essential part of the agent-based modeling process; certainly for highly empirical models, but also for purely theoretical models. In both cases, however, there are yet (to my knowledge) too few statistical tools available specifically for analysis of output from complex spatial systems. I strongly suspect that these tools exist and await only discovery and better communication between currently disconnected groups of researchers.

In addition to the roles for statistical models in building empirical agent decision functions described in Robinson et al (Forthcoming), I see two potential additional roles:

Application 1: Pseudo-inductive analysis. Demonstrating that an ABM can recreate an observed pattern is a first step towards using ABM as a substitute for traditional abstract mathematical models. Yet a modeler is likely to have larger goals. He may want to demonstrate that the outcome holds globally, over a large (and reasonable) range of parameter values. He may also want to understand global relationships between directions of change of parameters and directions of change of outputs. The solution to this problem has been well articulated in the agent-based modeling literature (Axelrod and Tesfatsion 2006): the modeler should create a database of outcomes by sweeping the parameter space, then use statistical methods to analyze the generated data. However, what has been less well articulated (to my knowledge) is what tools are available that are statistically appropriate for analysis of complex systems. By their nature, complex systems are characterized by endogeneity between micro-scale elements and macro-scale outcomes due to cross-scale feedbacks. They are also characterized by non-linear response surfaces and thresholds where abrupt changes occur. These properties violate the assumptions of the mostly-linear regression models in the historical toolkit of social science modelers. Although traditional regression techniques have been applied to the analysis of output from agent-based land-use models by myself (Parker 2005) and Happe et al (2006), both authors acknowledge the limitations of the linear models that they employ.

Application 2: Model validation. The first application focuses on statistical analysis of the relationship between micro-scale drivers and macro-scale measures of pattern outcome. A second potential application (not applied to the author's knowledge) involves statistical analysis of ABM output at the micro level, and comparison of that model output to results from the same statistical model applied to real-world data (which of course would have to lie outside any data used to build the ABM). For example, an empirically-parameterized agent-based model of residential land markets would produce outputs at the micro level similar to those used to estimate spatial econometric models of land-use change, including land-use transitions, transaction prices, distance to transportation networks, and neighborhood composition at multiple spatial scales. Such data could be used to estimate a land-use change model using techniques similar to those described by Bell and Irwin (2002). If the estimated parameters of the simulated and real-world models had the same signs and statistical significance, the model could be said to exhibit qualitative agreement with the real world-data. If the parameter estimates were statistically similar, the model would have hit the proverbial "home run" (although likely other validation methods would probably also be called for). This approach goes beyond asking if the model replicates spatial outputs such as location, pattern, and composition at multiple scales, and rather asks how closely the model replicates the structural relationships found in the real-world data. (Additional thought would be required to account for path-dependence and the distribution of simulated model outcomes that would be possible (Brown et al. 2005).)

Question 3: What is the role for calibration in empirical agent-based models? Calibration (derivation of a set of best-fit model parameters through comparison of outcomes at the micro or macro level) has played multiple roles in spatial empirical ABMs. Both statistical and mathematical programming-based agent decision models have been calibrated using micro-level outcome data (Balmann et al. 2003; Berger 2001; Happe, Kellermann, and Balmann 2006; Schreinemachers and Berger 2006). Parameters of agent decision models have also been calibrated using macro-scale data on land-use composition and pattern (Caruso, Rounsevell, and Cojocaru 2005; Evans and Kelley 2004). Each of these methods implicitly assumes that the chosen structure of the decision function is correct, but that the parameters of that function are uncertain. Thus calibration does not necessarily provide a means to distinguish between competing decision models. If macro-scale pattern outcomes of interest are used for calibration, then separate outcome data must be used to validate the model after calibration, increasing overall data requirements for the model. Decision models often contain many more parameters than the number of observations available on macro-scale outcomes, leading to potential parameter identification problems, and exacerbating the problem of distinguishing between alternative decision models, as there may be cases where different sets of parameter values for the same decision model may lead to equivalent pattern outcomes. Grimm et al (2005), referring to the calibration process described above as "inverse modeling", propose calibration via multiple pattern outcomes as a means of addressing the parameter identification problem. Calibration may have an important role to play for spatial ABM at a later stage of model development and/or in cases where agent rules and behaviors are well known (for example, when they have been determined via field observation or laboratory experiments).

References:

Axelrod, R., and L. Tesfatsion. 2006. A Guide for Newcomers to Agent-Based Modeling in the Social Sciences in L. Tesfatsion and K. Judd, eds. *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*. North-Holland/Elsevier, Amsterdam

Balmann, A., K. Happe, K. Kellermann, and A. Kleingarn. 2003. Adjustment costs of agri-environmental policy switchings: A multi-agent approach in M. A. Janssen, ed. *Complexity and Ecosystem Management: The Theory and Practice of Multi-agent Approaches*. Edward Elgar Publishers, Cheltenham, U.K.; Northampton, MA

Bell, K. P., and E. G. Irwin. 2002. Spatially explicit micro-level modelling of land use change at the rural-urban interface. *Agricultural Economics* 27 (3): 217-232

Berger, T. 2001. Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes, and policy analysis. *Agricultural Economics* 25 (2-3): 245-260

Bousquet, F., and C. Le Page. 2004. Multi-agent simulations and ecosystem management: a review. *Ecological Modelling* 76 (3-4): 313-332

Brown, D. G., S. E. Page, R. Riolo, M. Zellner, and R. W. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of*

Geographic Information Systems 19 (2): 153-174. http://www.pscs.umich.edu/research/projects/sluce/publications/ijgis-sluce-final.pdf.

Caruso, G., M. Rounsevell, and G. Cojocaru. 2005. Exploring a spatio-dynamic neighbourhood-based model of residential behaviour in the Brussels periurban area. *International Journal of Geographical Information Science* 19 (2): 103-123

Castle, C., and A. Crooks. 2006. Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. London: CASA Publication 110. http://www.casa.ucl.ac.uk/publications/workingPaperDetail.asp?ID=110.

Epstein, J. M. 1999. Agent-based models and generative social science. *Complexity* 4 (5): 41-60

Epstein, J. M., and R. Axtell. 1996. *Growing Artificial Societies: Social Science from the Ground Up*. Brookings Institution Press, Washington, D.C.

Evans, T. P., and H. Kelley. 2004. Multi-scale analysis of a household level agent-based model of landcover change. *Journal of Environmental Management* 72: 57–72

Grimm, V., and S. F. Railsback. 2006. Chapter 1: Introduction in V. Grimm and S. F. Railsback, eds. *Individual-based Modeling and Ecology*. Princeton University Press, Princeton, NJ

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis. 2005. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* 310: 987-991

Happe, K., K. Kellermann, and A. Balmann. 2006. Agent-based Analysis of Agricultural Policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its Adaptation and Behavior. *Ecology and Society* 11 (1): 49. http://www.ecologyandsociety.org/vol11/iss1/art49.

Janssen, M. A., ed. 2003. *Complexity and Ecosystem Management: The Theory and Practice of Multi-Agent Approaches*. Edward Elgar Publishers, Cheltenham, U.K.; Northampton, MA

Parker, D. 2007. Class Web Site: Spatial agent-based models of human-environment interactions. http://mason.gmu.edu/~dparker3/spat_abm/spat_abm.html.

Parker, D. C. 2005. Agent-Based Modeling to Explore Linkages Between Preferences for Open Space, Fragmentation at the Urban-Rural Fringe, and Economic Welfare. Paper presented in the The role of open space and green amenities in the residential move from cities, Dec. 14-16, Dijon, France.

Parker, D. C., T. Berger, and S. M. Manson. 2002. Meeting the Challenge of Complexity: Proceedings of the Special Workshop on Agent-Based Models of Land-Use/Land-Cover Change. Santa Barbara: CIPEC/CSISS Publication CCR-3. http://www.csiss.org/maslucc/ABM-LUCC.htm.

Parker, D. C., S. M. Manson, M. A. Janssen, M. Hoffmann, and P. Deadman. 2003. Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. *Annals of the Association of American Geographers* 93 (2). http://www.csiss.org/events/other/agent-based/papers/maslucc_overview.pdf.

Robinson, D. T., D. G. Brown, D. C. Parker, P. Schreinemachers, M. A. Janssen, M. Huigen, H. Wittmer, N. Gotts, P. Promburom, E. Irwin, T. Berger, F. Gatzweiler, and C. Barnaud. Forthcoming. Comparison of empirical methods for building agent-based models in land use science. *Journal of Land-Use Science*

Schelling, T. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143-186

Schreinemachers, P., and T. Berger. 2006. Land use decisions in developing countries and their representation in multi-agent systems. *Journal of Land Use Science* 1 (1): 29-44. http://journalsonline.tandf.co.uk/openurl.asp?genre=article&issn=1747-423X&volume=1&issue=1&spage=29.