Agent-based modeling

Dawn C. Parker and Derek T. Robinson
University of Waterloo, Canada

Agent-based models: definition and context

A definition of agent-based models of geographic systems

Agent-based modeling (ABM) is a computational simulation methodology that represents key system actors as agents who make individualized, decentralized decisions. Agent-based models represent characteristics of real-world actors, items, and processes as virtual agents, objects, and algorithms, most often in computer code. These actors can include, but are not limited to, gas particles, animals, plants, people, or more complex social entities such as households or institutions. When applied in a geographic context, ABMs also generally include a spatially explicit landscape over which agents interact. Thus ABMs can situate agents within a geographic context, concurrently representing spatial aspects of actors, their decisions, and their interactions. This enables the modeling of both direct and indirect spatial interactions, which are often central to the research questions investigated by ABMs of geographical systems.

Identifying the community of ABM users

ABM is increasingly used throughout the social, natural, and information sciences. ABM has been applied in the domain of ecology, where it has been referred to as individual-based modeling (IBM) since the 1970s. Many geographical applications represent coupled human and natural systems (CHANS). In other disciplines ABM is often used to generate solutions to specific problems. For example, in computer science, software agents are designed to accomplish a specific task (e.g., harvest data from the World Wide Web), and in operations research, n-polynomial optimization problems are solved using agents. In many other cases ABM is used for pedagogical purposes (e.g., to illustrate theory or teach computer programming).

How ABMs develop decision models

Often modelers combine theoretical frames from several traditional disciplines to develop agent decision-making, action and reaction, or interaction rules among agents and between agents and their environment. In some research and applications, agent rules, attributes, and behaviors are empirically informed from social surveys, behavioral economics experiments, ethnographic fieldwork, ecological fieldwork, spatial analysis, and other data collection and analysis methods. Often these empirical efforts strive to (i) objectively identify groups of agents within a specific agent type (e.g., profit-driven, risk-averse, or hobbyist farmers), (ii) calibrate agent attributes and population numbers, and (iii) provide empirical patterns for validating model structure and output.

ABMs as representations of complex systems

Since it is not necessary to define and solve a mathematical equilibrium or optimum for

The International Encyclopedia of Geography.
Edited by Douglas Richardson, Noel Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu, and Richard A. Marston.

DOI: 10.1002/9781118786352.wbieg0891

ABMs, as required by many other modeling methods, there is a high degree of flexibility that allows researchers to construct ABMs that represent a wide range of processes and feedbacks. Therefore, ABMs are ideal tools to represent and explore complex systems. While multiple definitions of complex systems are in play, commonly recognized components include: heterogeneity and interaction among agents and their environment, learning and adaptation, aggregation (i.e., emergence), path dependence, and nonlinear dynamics. Many of these components cannot be practically or wholly represented using mathematical (e.g., ordinary differential equations) or statistical (e.g., logistic regression) models.

Modelers incorporating the concepts of complex systems into ABMs are able to make both theoretical and applied advances by constructing a range of models that extend from simple toy models (e.g., flocking behavior in birds) that lack data but offer a proof of existence to high-fidelity models that enable the simulation of real-world physics and processes (e.g., a flight simulator). Thus, ABM is used both as a tool for exploration and to make substantive gains in domain-specific and broader scientific knowledge. In the real world, many phenomena cannot be methodically explored due to lack of data variation, especially given that the real world provides only a single historical trajectory reflecting particular circumstances. In such circumstances, ABMs can be used as "computational laboratories," to explore such phenomena "in silico."

One strong appeal for complex systems modelers using an ABM approach is to demonstrate that the individual actions and interactions of agents can aggregate to create macroscopic patterns of interest at the system or landscape level, which can include measures of spatial pattern and connectivity or equilibrium (i.e., emergence). If, through the validation process, these emergent outcomes are found to be consistent with a sufficient number of real-world data series and/or patterns, the models are said to offer a supportable candidate explanation for its hypothetical causal mechanisms. Again, although an equilibrium may emerge in the ABM, its conditions are not imposed on the model a priori. In addition to the standard notion of equilibrium, ABMs are also used to represent nonequilibrium situations including threshold effects, punctuated equilibrium, quasi-equilibrium, and out-of-equilibrium dynamics.

Elements of ABM

Landscape

An agent-based model of a geographic system generally begins with a representation of the spatial environment over which modeled agents interact. Typically, the data models used to represent the landscape correspond to raster and vector models used in traditional geographic information systems. The majority of spatially explicit ABMs use a raster-based representation of the landscape. This may be due to the ease with which data can be assimilated during model construction, accessed by agents within the model, and compared to remotely sensed data during model evaluation, among other reasons. Vector-based applications are increasingly used but face challenges when the subdivision of a polygon or other topological process requires representation. Other representations of space have also been used (e.g., triangular irregular networks, hexagon tessellations, networks). In all cases, the landscape units (e.g., cell or polygon) can store attribute information that facilitates linkages between stationary or mobile agents and neighborhood stimuli. In land-change models, for example, agents can modify the land use of landscape units, and these units in turn may be used to drive other modeled social or environmental processes or inform the decisions of other agents.

Collection of agents

The second element of an ABM is a defined collection of agents. While many formal "agent" definitions are seen in the literature, an agent is generally a goal-oriented entity, possessing a model of cognition that links goals and behavior, that is capable of sensing and responding to changes in its environment through autonomous action. Agents are generally characterized by both internal resources (including attributes representing knowledge, experience, behavioral parameters, and decision strategies) and external resources, such as assets or position within a social network.

Each agent possesses a decision model – a set of rules that link information about his or her environment, resources, and estimated payoffs to a decision regarding possible actions. In an ABM, decisions often depend on the state of the agent's attributes and the physical environment. They may also depend on the state of the agent's social environment. Specifically, decisions in ABMs are often linked to and influenced by the prior decisions of other agents. A wide variety of models are used to represent decision-making in ABMs, including rule-based and classifier systems, regression and neural network models, and optimization routines such as mathematical programs or genetic algorithms. There is no consensus on the most appropriate way of representing agent decision-making in an ABM, and current decisions are based on the complexity of the actors represented, the skill and domain specialization of the modeler(s) or project team, and the data and time available.

Interaction

The third key element of ABMs is a set of rules governing interaction among agents and between agents and their environment. Rules governing agent–agent interactions may involve a distance parameter that defines the vision of the agent. For example, the implementation of a predator–prey model requires both the predator and the prey to have defined a maximum distance for which they may observe other agents in the landscape. In other cases, interactions may be based on the evolution of a social network, a change in household dynamics based on its life cycle, or observation of the behaviors of spatial neighbors (adjacent or farther away). The result of these interactions can be a change in the flow of information, interdependencies between and impacts of agent actions, and the sequencing of model events.

Event sequencing within an ABM plays an important role. In ABMs, event-sequencing mechanisms can generally be characterized as predetermined and event-driven. Predetermined event-sequencing rules are set by the modeler, and generally operate similarly over model runs. Such rules can be used to simulate synchronous behavior, where each model agent has an opportunity to make a decision, taking the last previous action of all other agents as a starting point. Rules can also be asynchronous, where, for example, each agent has a fixed probability of being active in any time period. In asynchronous implementations of agent behaviors, model results can be dependent on the order in which agents act. To alleviate this modeling artifact, most ABM platforms provide a shuffle routine that rearranges the order of agents within the population that takes actions. Event-driven rules, in contrast, dictate action when a state variable in the system – for a single agent, another agent, or the system as a whole – reaches a certain threshold. For example, a farm may be sold when the owner retires, or a family may seek a larger residence when they have children.

ABMs are also generally constructed with mechanisms to track and record information about model states and flows. Such information can be stored in many formats – GIS layers, text files, formal databases, and others. Software and libraries designed for building agent-based models often have such capabilities built in.

When ABMs are applied to models of geographic systems

Spatially explicit ABMs have been used by researchers in a variety of different disciplines to gain insight into space-dependent problems (e.g., geography, economics, planning, political science, ecology, anthropology). Examples of spatial applications include: evaluating the effects of different types of policies on the patterns of urban development and their subsequent ecological impacts; estimating conflict and improving route scheduling for tourists in national parks; quantifying changes in crop yields due to demographic and life cycle changes in farming households; comparing social and environmental factors affecting the extinction of communities; quantifying human responses to emergencies and assessing alternative designs to affect evacuation patterns; and defining territorial boundaries for nations within states. A large suite of recent applications are presented in Heppenstall *et al.* (2012).

In each of these domains of study, spatially explicit ABMs can be better designed and understood by conceptualizing the sources of spatial, temporal, and behavioral complexity in the system of study and captured in the ABM. The majority of geographic applications include one or more sources of spatial complexity. While not comprehensive, spatial complexity can be broken down into three areas: spatial autocorrelation, zonal spatial relationships, and network relationships.

- *Spatial autocorrelation*, at a general level, follows Tobler's first law of Geography: "All things are related, but nearby things are more related than distant things." Spatial autocorrelation refers to cases in which a spatial state is either more or less likely when in proximity to a similar spatial state. In human systems, spatial autocorrelation can occur for many reasons, including imitative behavior, diffusion processes, synergies between compatible activities, spatial competition, and negative spatial spillovers. It implies a causal spatial relationship between locations.
- *Zonal spatial relationships* occur when locations share a common characteristic or state variable value, for a noncausal reason – for example, similar levels of accessibility, political or zoning designation, soil type, climate, or topography. Often different zonal relationships are non-nested (for example, political and watershed boundaries), creating measurement and modeling challenges.
- *Spatial networks* define spatial relationships between entities that are constrained by an existing structure in the landscape (e.g., roads, pipelines, forest corridors). Unlike measurements that are based on Euclidean distance (i.e., what is commonly referred to as "as the crow flies"), spatial networks provide a realistic account of distance and include impedances that can slow the travel of processes through the network. Networks can be physical, such as transportation or hydrologic networks, or social, such as communication, relationships, and trade and commerce networks. Concepts and algorithms from network analysis are commonly applied in ABMs.

Temporal complexity shares many of the same characteristics as spatial complexity, but instead of quantifying the behavior of a process across space, it is concerned with quantifying the process over time. Temporal autocorrelation measurements are used to quantify the dependency of future measurements of a phenomenon based on past measurements. Geographical systems are often characterized by important temporal growth and decay processes. These processes can be social, such as population growth and decline, social trends, and growth of financial investments. They can also be biophysical, such as soil degradation, carbon sequestration, and erosion. Such processes lead to temporal autocorrelation, where a given time period's state depends on that of the last time period.

When combined with spatial autocorrelation and its various drivers, processes will affect state variables across both time and space, such as species colonization and fire spread. Such processes are often modeled using cellular automaton models, which are implicitly contained in many ABMs. Further, Hagerstrand's time geography approach is increasingly applied to monitoring individual and group behaviors in agent-based models and provides one of many geovisualization techniques to evaluate agent behavior in an ABM.

Behavioral complexity in spatially explicit ABMs refers to the degree of complexity incorporated in representing how agents make decisions as well as the complexity of social and institutional structures represented in the model. With regard to decision-making, a range of approaches are used that span from simple heuristic "if-then-else" algorithms, to utility equations, and learning and adaptation algorithms (e.g., genetic programming). While often omitted from ABMs of geographical systems, forward-looking behavior by humans and other animals plays an important role in decision-making. Examples include food and seed storage, crop rotation, strategic behavior, such as pre-emptive colonization, and financial investments. To account for such behavior, models must build in expectations of the future and possible responses in order to develop models of learning and adaptation.

Behavioral actors in models are often characterized by heterogeneity and interdependencies. Not only are different types of actors present in most systems, even within a type, actors may differ in their goals, expectations, strategies, and motivations, as well as their internal and external resources. Actors are also interconnected in social, economic, and ecological networks. Collections of actors may also have their own rules of behavior.

The representations of the spatial, temporal, and behavioral complexity of a study system in an ABM are each characterized by a notion of scale, and complex processes for each can operate at multiple scales. Spatial scale is characterized by spatial resolution (e.g., minimum mapping unit) and spatial extent (i.e., total area represented), temporal scale by temporal resolution (i.e., length of the smallest unit of time represented, also known as the time step) and temporal extent (duration of time represented), and behavioral scale by the minimal decision-making actors and collectives of such actors.

Processes can be characterized by feedbacks across scales in each case – for example in a spatial system, between neighborhoods and a city, in a temporal system, between fast processes such as weather events and slower processes such as climate change, and in a behavioral system, by interactions between family members and a household. Cross-scale processes can operate both from the bottom up and from the top down, with feedbacks between scales.

The collective result of these forms and sources of complexity is that systems generally modeled using ABM are highly path dependent. A small change in an initial condition or stochastic seed at a given point in time may move the system to a qualitatively different outcome. Modelers need to be aware of this sensitivity, testing models using a wide variety of initial conditions and/or stochastic seeds to trace out distributions of outcome variables that may emerge from the modeled system.

Model development: elements and process

While the development of ABMs often proceeds iteratively, ideally based on scientific and stakeholder feedback, the following template provides a framework that can be usefully applied to the development of most ABMs of geographical systems.

First, the purpose of modeling should be identified. While many researchers find that the process of model construction can be useful to codify existing knowledge and gain consensus as to an appropriate representation of the real-world system, ideally model design should be framed by the clearly defined research questions of what the model is for and what questions its output can be used to answer.

Next, researchers typically identify the real-world spatial, temporal, and behavioral processes they strive to represent. ABMs are structural models, which embed hypotheses about process-based relationships and causal mechanisms. The process of identifying these relationships and mechanisms in the applied problem is critically important. Often, it can be very useful to create simple system diagrams, noting the scale of operation of each process and identifying key causal linkages via arrows and feedback loops. Having done this, researchers can identify what factors and processes they might reasonably assume are exogenous (determined from outside the model – information you have before you run the ABM) and endogenous (determined as part of the model run – information you did not have before you ran the model). It is important within this process to try to distill the model elements to the minimal possible level of complexity that you can practically model. While the first question to ask is, "How can we characterize the complexity of the real-world system?" the second is, "How little complexity can we get away with modeling?"

Having undertaken this process of identifying real-world structures and mechanisms, if building an empirically supported model, the team should identify the ideal spatial, temporal, and behavioral representation (both structure and resolution) for model parameterization, calibration, and assessment data. As with any empirical undertaking, this phase is generally followed by a practical assessment of the data that might actually be available for model construction. This is generally followed by a second round of model structuring during which the modelers simplify the model plan and goals. Empiricizing ABMs can be particularly challenging, as they require data about both actors and the actors' environment, in contrast to some geographic models that rely on spatial data alone, or social science models that use only actor data. The fine-scale, disaggregated nature of ABMs creates further challenges.

Having made high-level decisions regarding what processes to model and what data to use, modelers must turn their attention to detailed decisions regarding model design, model rules, and system architecture. While no standard texts or templates have been available historically, experienced researchers are beginning to offer texts to guide learning and model development (see below for readings). Decisions regarding software implementation must be made. In spite of community consensus that such tools could be useful, to date few standard software packages and model libraries are available. However, two sets of software libraries are commonly used: RePast and Netlogo. Further, the Community Modeling Library hosted by the CoMSES (Computational Modeling in the Social and Ecological Sciences) Network provides code examples of previously implemented models (http://www.openabm.org/models). Journals are also providing links to model code archives with increasing frequency. Along with a push towards code sharing, many scholars have emphasized the need for standard model documentation protocols and metadata. Various approaches have been taken to the problem. Along these lines, research groups should set forth protocols for documenting, archiving, and sharing their model data and code, with access rights within and external to the research group carefully defined before modeling begins. Often such agreement might be dictated and/or constrained by institutional and access agreements. Such points should be considered when choosing data and software options.

The value of ABMs

ABMs offer many advantages. While often seen first and foremost as models, ABMs offer value by encouraging discussion and learning in a number of ways. First, even before an ABM is coded, the design phase can act as a medium for discussion and can formalize a range of assumptions and knowledge about a study system. This process continues as assumptions are formalized in computer code. This modeling process often acts as a bridge across disciplines, facilitating communication and project developments on large interdisciplinary research teams. Second, the ABM offers a proof of concept and a pedagogical framework to demonstrate the effects of different processes in isolation or in combination. Third, because ABMs are built with a focus on representing process, behavior, and interaction, they often cause data to be analyzed and interrogated in new ways. They may also highlight new types of data that have not been collected but are necessary to understand a question being answered.

Beyond their pedagogical and conceptual contributions, ABMs are valued for their ability to provide evidence to support general and overarching advances to our scientific understanding of a broad range of human and natural systems. For example, Schelling's segregation model demonstrates that society may segregate from households having only a small preference for neighbors that are like them (e.g., 25%). New models have represented sheepdog behavior with only two agent rules and offer approaches for corralling crowds. Others have shown that simple behaviors can solve complex problems (e.g., ant-agents searching for food and leaving pheromones can solve n-polynomial optimization problems). These are but few of the many hundreds, if not thousands, of systems to which ABM are (or could be) applied.

Limitations and challenges

In spite of the successful applications described above, the development, analysis, and communication of ABMs present challenges. These challenges are widely recognized by the modeling community, and incremental and steady progress is being made to address them.

On the model development side, standard and easy-to-use software libraries, such as those available for statistical analysis, are not yet generally available for ABMs. This creates a barrier to entry to the field, as generally some programming abilities are needed to construct and modify an ABM. However, increasing availability of software packages such as Netlogo and introductory texts create a fairly accessible entry point to programming and modeling. Even for the proficient programmer, there is a lack of standardized libraries and modeling algorithms available. In spite of a recognition that it may not be the most efficient use of resources at a community level, most projects start from scratch to build their own models, rather than reusing existing model components. Thus, model comparison, evaluation, and replication are hindered.

Due to their fine-scale design, ABMs can be computationally demanding. This can create a particular challenge in terms of scalability, especially in cases where the ABM is intended to provide an input into a higher-scale process model. For example, an ABM may be used to model a local economy, with the input designed to feed into a regional model. One option to address this challenge is to explore methods to fit metamodels, or aggregate functions that represent model behavior, to the output of ABMs, then use the metafunction as an input to the higher scale model. For example, ABMs have been developed that replicate the smooth, regular system dynamics seen in analytical epidemiological models and predator–prey models. Further, conceptualizing the event-sequencing methods and the modes of interaction among actors and processes acting across and between spatial and temporal scales remains an ongoing challenge.

As mentioned above, empirical ABMs can have high data demands, especially when they combine spatial and actor level data. Both cost and confidentiality concerns can be barriers to data access. The need to separate data used for model development from that used for validation creates further data demands. As ABMs become more empirically informed, they have become more site specific and have rarely been applied across multiple case-study sites, limiting their ability to demonstrate broad scientific achievements.

Once a model is constructed and verified, model analysis can be challenging due to the complexity of model design. While fairly simply protocols have been developed for interpretation of alternative modeling methods, such as comparative statics and dynamics for analytical equilibrium models, and interpretation of estimated coefficients for regression models, standardized methods for understanding relationships between the exogenous and endogenous elements of ABMs are not yet developed. Given the purposeful pursuit of nonlinearity and complexity in ABM, it can be especially challenging to map, interpret, and understand model behavior over the entire potential parameter space of the model. From a scientific perspective, this leads some to question the validity of the simulation modeling effort, as the global behavior of models is not well understood. A corollary to this challenge is that it can be challenging to communicate complex model mechanisms and interpret outcomes for nonacademic audiences, such as local stakeholders, planners, politicians, and policymakers. Participatory or companion modeling methods, in which models are co-designed and analyzed, can potentially address this challenge.

SEE ALSO: Agents, ontology-based decision-support systems; Cellular automata; Cognition and spatial behavior; Exploratory spatial data analysis; Geocomputation; Geography and the study of human–environment relations; Microsimulation; Representation: dynamic complex systems; Spatial analysis; Spatial interaction; Time geography and space–time prism; Tobler's first law of geography

Reference

Heppenstall, A., A. Crooks, L. See, and M. Batty. 2012. *Agent-Based Models of Geographical Systems*. Dordrecht: Springer.

Further reading

Alessa, L.N., M. Laituri, and M. Barton. 2006. "An 'All Hands' Call to the Social Science Community: Establishing a Community Framework for Complexity Modeling Using Agent Based Models and Cyberinfrastructure." *Journal of Artificial Societies and Social Simulation*, 9(4): 14.

Bell, A.R., D.T. Robinson, A. Malik, and S. Dewal. 2015. "Modular ABM Development for Improved Dissemination and Training." *Environmental Modelling and Software*, 73: 189–200.

Bousquet, F., and C. Le Page. 2004. "Multi-Agent Simulations and Ecosystem Management: A Review." *Ecological Modelling*, 176(3–4): 313–332.

Filatova, T., P. Verburg, D.C. Parker, and C.A. Stannard. 2013. "Spatial Agent-Based Models for Socio-Ecological Systems: Challenges and Prospects." *Environmental Modelling and Software*, 45: 1–7.

Gilbert, N., and K.G. Troitzsch. 2005. *Simulation for the Social Scientist*. London: Open University Press.

Gimblett, H.R., ed. 2001. *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Understanding Social and Ecological Processes*. Oxford: Oxford University Press.

O'Sullivan, D., and G.L. Perry. 2013. *Spatial Simulation: Exploring Pattern and Process*. Chichester: John Wiley & Sons.

Parker, D.C., S.M. Manson, M.A. Janssen, *et al.* 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): 314–337.

Railsback, Steven F., and Volker Grimm. 2012. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton: Princeton University Press.

Robinson, Derek T., Daniel G. Brown, Dawn C. Parker, *et al.* 2007. "Comparison of Empirical Methods for Building Agent-Based Models in Land-Use Science." *Journal of Land-Use Science*, 2(1): 31–55.

Rounsevell, M., D.T. Robinson, and D. Murray-Rust. 2012. "From Actors to Agents in Socio-Ecological Systems Models." *Philosophical Transactions of the Royal Society B*, 367(1586): 259–269.

Torrens, P.M. 2010. "Agent-Based Models and the Spatial Sciences." *Geography Compass*, 4(5): 428–448.