Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

GUIDELINES FOR DESIGN AND ANALYSIS IN AGENT-BASED SIMULATION STUDIES

Parastu Kasaie

Department of Epidemiology
Bloomberg School of Public Health
The Johns Hopkins University
Baltimore, MD 21205, USA

W. David Kelton

Department of Operations, Business Analytics,
and Information Systems
University of Cincinnati
Cincinnati, OH 45221-0130, USA

ABSTRACT

Agent-based simulation (ABS) continues to grow in popularity and in its fast-expanding application in various fields. Despite the increased interest, however, a common protocol or standard curriculum for development and analysis of ABS models hardly exists. As originally discrete-event simulation (DES) modelers, self-taught and still new to the world of ABS modeling, we have occasionally observed a gap between traditional simulation theory and current practices of ABS in the literature. This points to great unevenness among existing ABS applications in terms of concepts and design, quantitative and computational techniques used in analysis of models, as well as domain-specific issues in different fields. In this paper, we review a number of important topics and issues in the design and analysis of ABS models that deserve attention. Our discussion is supported by some illustrative examples from ABS models of disease epidemics, but it’s applicable to a fairly general class of ABS models.

1 INTRODUCTION

Many different models carry the moniker *agent-based simulation* (ABS) in different disciplines like artificial intelligence, complexity science, game theory, etc. However, there is no universally accepted definition for properties of an object to “deserve” the name of an “agent” (Borshchev and Filippov 2004), and to distinguish the differences among so called “micro-,” “individual-based,” or “agent-based” simulation models. In general, the distinguishing characteristic of ABS is based on decentralization of a system into its constituent components, and representing it through a collection of agents and their environment. The behaviors are programmed at the individual level, and system properties emerge from constituent agent interactions with each other and the environment, also referred to as a *bottom-up* modeling approach (Bonabeau 2002).

A recent sampling of agent-based modeling (ABM) found applications published in many disciplines ranging from archaeology, biology and ecology, supply chains, consumer market analysis, military planning, and economics (Heath, Hill, and Ciarallo 2009). Several indicators of the growing interest in ABM include the number of conferences/workshops on the topic, the growing number of peer-reviewed publications in discipline-specific academic journals across a wide range of application areas and in modeling and simulation journals, as well as the growing number of simulation-software products accommodating ABS modeling (Macal and North 2010).

Despite the substantial and growing interest in this type of modeling and the growing demand for systematic instruction on how to develop and apply ABM techniques, a standard curriculum for teaching ABS has not been established. While there are many simulation courses and training programs focused on more traditional (historical) simulation paradigms such as DES or system dynamics (SD), only a few full-length courses exist on ABS. Given the overlap between ABS and these simulation techniques, there is

978-1-4673-9743-8/15/$31.00 ©2015 IEEE

considerable interest in incorporating ABS into existing simulation courses. The full extent of these relationships, however, is not well defined or fully understood, and many aspects of ABS differ from DES and SD (Macal and North 2010). These include a wide array of application domains, the disciplines and diverse backgrounds of students working with ABS models, as well as the additional and inherent complexity of ABS models that merits more attention in addressing issues for design and analysis of models.

ABS applications range across very diverse fields, and modelers often come to ABS by way of self-study or attendance at tutorials or short courses, and with a wide range of quantitative, programing, and modeling skills (Macal and North 2010; Macal and North 2013). In many applications, the process of modeling and design is an ad hoc activity, driven by the modeler's creativity and natural attitudes toward modeling. Moreover, there is a large discrepancy in the range of quantitative and computational techniques used in analysis of ABS models. Given the short history of ABS applications in many fields, there is also a lack of reporting standards in the literature, especially for description of methods particular to these models (Grimm et al. 2006). For example, a systematic review of individual-based HIV transmission models suggests a substantial discrepancy among studies in terms of providing justification for application of ABS to address questions of interest (provided in 56% of papers), justification of modeling assumptions for choice of the simulation time step (19% of papers) or methods of data conversion (9.4% of papers), carrying out stochastic sensitivity analysis (46% of papers), discussion of applied methods for model validation and implementation (3.1% of papers), etc. (Abuelezam, Rough, and Seage 2013). Such discrepancy further highlights the need for discussion of standardized protocols for development and application ABS models in various fields.

As originally DES-modelers, self-taught and still new to the world of ABS modeling, we have occasionally observed a gap between traditional simulation theory (mostly developed for DES) and current practice of ABS models in the literature. These include the lack of justification for choosing ABS models over simpler modeling paradigms; a general tendency for development of "realistic," heavily detailed ABS models despite the tradeoff between complexity of analysis and transparency of findings; modelers' personal internal assumptions and mechanisms for agents' behavior and interactions without explicit validation and consideration of alternatives; instances of excessively parameter-rich ABS models requiring specialized calibration techniques (not generally discussed for DES models); unclear definition of errors in ABS applications as relevant to different sources of modeling mistakes vs. natural stochastic uncertainties; specific challenges faced in validation of ABS models due to their multi-level structure; as well as further special issues in experimentation and sensitivity analysis of these models.

In this paper, we review a number of important topics and issues in design and analysis of ABS models that deserve attention. While this is not intended to be a survey of all the available literature on ABS modeling or existing methods, our goal is to offer some general guidelines to modelers new to the ABS field (similar to us), and provide reminders of some important notions that merit more attention. In the following sections, we review and discuss several topics and issues for design, parameterization, verification, validation, and experimentation concerning ABS models. Our discussion is supported by a few illustrative examples from ABS models of disease epidemics as relevant to our area of expertise. While this is a small window into the wide literature of ABS applications, our discussion of modeling strategies and analysis techniques is applicable to a fairly general class of ABS models, and we encourage readers to consider this discussion in their specialized fields of interest.

2 DESIGN AND DEVELOPMENT OF THE MODEL

2.1 Choosing the Right Modeling Paradigm and Abstraction Level

The first task in any modeling study is determining the modeling paradigm best suited for addressing the research question. The discussion on the applicability of ABS vs. other modeling techniques, therefore, should focus on the nature of the underlying research questions and not vice versa as "a hammer looking for a nail." DES and SD are two widely-used modeling tools that are alternatives to ABS (Borshchev and Filippov 2004).

SD models a system as a series of stocks and flows in which the state changes are continuous and the model is essentially deterministic. DES, on the other hand, models systems such as networks of queues and activities, where state changes in the system occur at discrete points in time and may affect the system state. Despite major differences in terms of structure and design, both SD and DES follow a top-down modeling approach in representing a system through a global influence diagram or system flowchart, and offer very low flexibility for incorporating individual levels of behavior and micro-dynamics.

In the context of infectious-disease epidemic modeling, for example, such micro-dynamics can relate to human behavior as relevant to transmission of the disease (e.g., role of sexual contact networks in transmission of HIV (Latora et al. 2006)), population mobility patterns relevant to spatial distribution of the disease (e.g., role of the air-transportation network in diffusion of pandemic flu (Epstein et al. 2007)), population heterogeneity relevant to progression and control of the disease (e.g., role of the patient's adherence to medical treatment outcomes (Dimatteo et al. 2002)), etc. In such contexts, ABS offers the most natural way for describing the underlying social/epidemiological system, and provides a flexible and powerful platform for modeling various healthcare interventions and addressing a wide array of policymaking questions. The decentralized bottom-up approach enables the modeler to describe a system from the perspective of its constituent units (agents such as (a) people, (b) pathogens like bacteria, or (c) carriers like dengue mosquitoes), and simulate the behavior of the system through their interactions with each other (e.g., contacts and transmission) as well as their environment (e.g., seasonal migration patterns), capturing emergent disease-diffusion patterns through time and space as the simulation runs. Moreover, the multi-level nature of such models enables explicit definition of various interventions at the individual level (e.g., TB (tuberculosis) contact-tracing programs) and at the population level (e.g., mass vaccination programs), and provides a powerful experimental platform to study the system's behavior and predict future trends. On the other hand, while ABS models provide more realistic representation of the systems and population under study, the realism is not an inherent virtue in and of itself, and it comes at the price of additional complexity in development and analysis of such models. This complexity is beneficial only if it leads to an improvement in the accuracy and validity of the model's predictions as relevant to the goals of the study. Therefore, it's important always to begin with simpler modeling alternatives, and move to more complicated methods only as warranted by the problem. For example, the analyst might start with a simple spreadsheet model and extend it to include additional features of the system until the model appears to be inadequate, and then change to another approach (Siebers et al. 2010).

Once the simulation paradigm is chosen, the next step, as in any modeling practice, is designing the underlying structure of the model, usually referred to as a *conceptual model*. This provides the blueprint of the model's main components and boundaries. Despite the simple definition, the decision on scope and boundary of a model, and specifically what level of detail to include or leave behind, can prove difficult in reality. Especially in the case of ABS, the bottom-up modeling approach provides high flexibility for modeling various aspects of the system and its components at several levels to provide "realistic" representations of reality. In choosing the right level of abstraction, the modeler should note that the main purpose of modeling is not to provide the most elaborate or most detailed (or even elegant) representation of a system, but rather to serve as a tool for addressing specific questions of interest. Therefore, we should start with the simplest form and include those and only those details that are essential to answering the question of interest, or in other words "*make things as simple as possible but no simpler*" (attributed to Albert Einstein).

An example of such issues in the context of epidemic control can be found by comparing models to study TB contact tracing, ranging from simplified global mathematical models, to detailed simulation models of transmission (Begun et al. 2013). An example is an age-structured, socio-demographic, individual-based model capturing patterns of TB transmission in households, schools and workplaces (Guzzetta et al. 2011). Such models offer a realistic view of TB transmission in real social settings, and also offer the advantages for explicit modeling of contact tracing at an individual level, as opposed to a single global parameter change in aggregate mathematical models. In the absence of specific data on household or workplace transmission rates, however, this extra detail can add uncertainty to the results and in turn require further assumptions for calibration of the model. Complex models that are excessively detailed can make unreasonable demands on specifying all the realistic input parameters, probability distributions, and stochastic input processes (e.g., NSPPs). Moreover, they can also be a barrier to translating findings into practice, since they are hard to explain and have computational demands that can limit analysis. Therefore, at present it is not clear that low-level, excessively detailed representations are always beneficial to the assessment of contact tracing, and superior to simpler models (Begun et al. 2013).

2.2 Choosing the Right Mechanisms for Agent Interactions

A common problem in modeling complex systems involving human behavior is identifying, and then conceptualizing, the underlying mechanisms for the human agents' behaviors and interactions. While one aims to develop a representative model of reality, believability is not the main concern. This allows modelers to reduce system complexity by making simplifying assumptions at the individual (and system) level to represent the local dynamics and avoid unnecessary complications. Given the (often) incomplete knowledge of systems, evaluating the impact of such simplifying assumption on final results is important.

For example, Rahmandad and Sterman (2008) study the role of network structure and population heterogeneity on predictions from an individual-based model of disease transmission. They design a controlled set of experiments using models with various configurations of the network (e.g., random, small-world, scale-free) and population heterogeneity (e.g., degree of attendance). Their results suggest that, while all models are calibrated to the same aggregate measures (in terms of the epidemic basic reproduction number R_0), the underlying structure for social norms can have a significant impact on the model's predictions. Now if the goal of the study is to map the short-term patterns of transmission, such impact may be of essential influence. However, in the context of healthcare policymaking, as considered by these authors, subtle changes may not have a significant effect on the final decisions, and simple models with coarser-grained representations may in fact lead to equivalently proper results.

In dealing with uncertain mechanisms, therefore, the modeler should use caution in making simplifying assumptions regarding the mechanics of the system, and consider various alternative approaches for modeling local dynamics. Good practice is to check the sensitivity of the final results to alternative approaches.

3 PARAMETERIZATION AND CALIBRATION

Parameterization is an important issue in the development of any simulation model, and especially ABS models due to their multi-level structure and often parameter-rich nature. The appropriate numbers of parameters and variables depend on the goal of the model, and the degree of realism and accuracy needed. This includes both the categories of so-called *fixed* (given or exogenous) parameters that can be estimated from available data or gleaned from the literature (as deterministic values, stochastic distributions, or dynamic functions), as well as *free* (experimental or variable) parameters that are unknown due to lack of relevant data for fitting or their model-specific definitions (that don't directly correspond to data). In general, including meaningful parameters that can be directly interpreted and understood is favored over "fitted" parameters (Helbing and Balietti 2013). This improves the readability of the code and aids in the calibration procedure.

The simulation literature has long discussed techniques for calibration of fixed parameters in DES models (Law and Kelton 2000; Henderson and Nelson 2006), and covers various topics on choosing the appropriate form, estimating, and assessing the goodness of fit, which are all applicable as well to ABS models. Other sources discuss various methods for calibration of free parameters in simulation models via aggregate model outcomes (examples include several heuristic algorithms (Fabretti 2012; Read et al. 2013), and Bayesian techniques (Kennedy and O'Hagan 2001)). In general, models with better predictive power (evaluated using separate data not used in the calibration) are favored over those with better descriptive power. Finally, given comparable predictive power, the model of simpler form with more analytical tractability is favored (the principle of Occam's razor).

In the case of ABS models, calibration may prove more challenging (Kim, Kim, and Rilett 2005). ABS models usually involve multi-scale parameters defined at a local level but entailing emergent global impact on system behavior (e.g., as seen in multi-agent simulation models of society where an agent's individual behaviors have further impact on the group's interactions and population outcomes (Fehler, Klügl, and Puppe 2005)). Such models would therefore require a multi-level approach to calibration of their parameters. Existing approaches are divided into two categories. First are so-called *blackbox* approaches, obtaining and using an approximate relationship between input and output variables of the simulation for determining the "optimal" input setting. Examples include gradient-based search methods, stochastic-approximation methods, sample-path optimization, response-surface optimization, and heuristic-search methods. A main drawback from this approach for ABS models is that no knowledge about the internal structure and the parameter dependencies of the simulation model is used. The second category includes so-called *whitebox* calibration methods that explicitly use model knowledge to enhance the calibration process. In this approach, the structural properties of the simulation model and knowledge about dependencies between the parameters that are to be calibrated is exploited to reduce the configuration search space, the complexity of parameter dependencies, and consequently decrease the computational cost of parameter configurations and evaluation (Fehler, Klügl, and Puppe 2005). These methods entail several rounds of model decomposition into a hierarchical structure, thus reducing the parameter space for each sub-model and shorter simulation runs due to reduced internal simulation times.

An example of the hierarchical calibration method can be found in the work of Guzzetta et al. (2011) in developing an age-structured, socio-demographic, individual-based model (IBM) with a realistic, time-evolving structure of preferential contacts in a population. After estimating the fixed model parameters from the literature, the IBM model is still left with several variables where no information is available from the literature to determine their values, and the parameters are left free to vary over a range during the calibration procedure. The detailed IBM model is computationally intensive and intrinsically stochastic. Multiple independent realizations (i.e., replications) of the model with the same parameter set are required to obtain stable and precise results in the presence of this random variability, and an exhaustive search of the parameter space with regard to values of all unknown parameters (seven parameters) is infeasible. Instead, the authors use a hierarchical calibration procedure via another sub-model as a special case of the full IBM model with a simpler form (no social structure) to reduce the parameter space. Starting with a global latin hypercube sampling (LHS) of all parameters in the simplified sub-model, they narrow the search to the vicinity of the best-fitted parameter set, and center the local search for the remaining IBM parameters around this point. The ranges of free parameters in the local search are reduced, based on indications from best-fitting simulations in the global search (Guzzetta et al. 2011).

4 MANAGING THE ERRORS: VERIFICATION AND VALIDATION

No measurement of a physical quantity can be entirely accurate. In experimental measurements, the term "error" usually connotes the differences of the measured value from the true value, and "uncertainty" can be defined as an estimate of the error (Coleman and Steele 1989). In the context of simulation modeling and ABS, the term "error" has wide implication for addressing the differences between simulation outcomes and the real phenomena of interest. Such differences, however, entail a wide array of "uncertainties" and "modeling mistakes" that should be assessed separately.

Uncertainty involved in mathematical/computational models and experimental measurements can occur in different contexts (Kennedy and O'Hagan. 2001). This includes *parameter uncertainty* that is implied by inclusion of free parameters with values that are unknown to the modeler, *parametric variability* that is caused by fixed parameters of a stochastic nature, *algorithm uncertainty* associated with numerical errors and numerical approximations in the implementation of the computer code (e.g., numerical integration and infinite-sum truncation that are necessary approximations in numerical implementations), as well as *structural uncertainty* (or model inadequacy) associated with lack of knowledge of the true underlying physics of a system (e.g., agents' definitions and behavioral rules governing their interactions). In contrast, the term "error" is interchangeably used to refer to a separate class of actual modeling mistakes associated with simulation applications. Such mistakes are defined as recognizable deficiencies in any phase of modeling and simulation that are not due to lack of knowledge, and can be recognized upon examination of the code. In general, these entail *compilation errors* that prevent the program from running, e.g., misspelling commands or syntactical errors (e.g., "if" without "end if"), *run-time errors* that occur while the model is running due to the occurrence of some impossible operations (e.g., an unseen division by 0), as well as *logical errors* that prevent the program from doing what was intend by the modeler.

Existing methods for dealing with modeling mistakes and uncertainties are known as *verification and validation* (V&V) methods. Model verification aims to ensure that a computer model/code is carrying out operations in the way that the conceptual model intends. There are several techniques for debugging, testing, and verifying computer programs including, but not limited to, using interactive debuggers to trace compilation errors, walk-through and logic-flow diagrams, face validity to control and reduce compilation error, extreme conditions for run-time errors, and aligning with known scenarios for logical errors (Balci 1995). Simulation validation, on the other hand, aims to ensure that the conceptual model is consistent with the real phenomena of interest. The term validation is used in association with empirical validation or statistical validation concerning the consistency of simulation output with real-world data, conceptual-model validation or operational validation concerning the consistency of concepts and simulated logic with system theories, structural validation or process validation concerning the consistency of model assumptions with how the system operates and is physically arranged, etc., each including several methods for addressing various sources of uncertainty involved in simulation models as mentioned above (Balci 1994; Tesfatsion 2015; Kleijnen 1995).

Despite the availability of several methods, validation (ensuring that the conceptual model accurately represents the reality being modeled) can prove to be a difficult task for ABS models (Klugl 2008; Bharathy and Silverman 2010; Yilmaz 2006). Empirical or statistical validation is possible only if output measures can be found that are able to describe the system appropriately, and only if such measures are available not only from the simulation but also from the real-world system. Such descriptors can be often easily found at aggregate levels (e.g., the number of people developing TB infection over time), but they're not trivial at the agent level, as individual behavior characteristics are hard to capture in a reasonable way (e.g., describing individual contact patterns with other community members throughout a day). Moreover, ABS models are especially apt for studying transient dynamics and answering questions concerning the dynamics and interactions that may or may not lead to an equilibrium or steady state (e.g., models of an Ebola outbreak). Furthermore, while most validation techniques make an implicit assumption that the steady state of the system is what's of interest, and focus on mean behaviors, the model dynamics resulting in those behaviors must be validated as well. In such a case, using procedures for time-series validation may be useful; however, accessing useful real-world time-series data can prove to be difficult for most applications of ABS models. In addition, further complications can arise due to non-linear effects of parameter changes in ABS models (as in many other modeling paradigms), resulting in chaotic effects on simulation outputs. Such chaotic behavior is hard to validate since minimal imprecision in the initial conditions may lead to completely different outputs.

In addition, the multi-level property of ABS models requires a multi-level V&V approach, in which not only input-output relations have to be compared for the overall system, but V&V procedures have to be performed as well for additional sub-ensembles of agents or partial models, down to single agents.

Potential drawbacks from this approach include the large time/effort requirement for performing the task, and most importantly, unavailability of necessary data for empirical validation of model behavior on all levels.

In the context of disease simulation, an example can be found in the development and calibration of the EpiSims model, a stochastic ABS model of airborne infections (such as pandemic influenza or smallpox) in a massive synthetic population of the U.S. (Del Valle et al. 2006). In the absence of explicit contact-network data for such a population, the modelers use a complex transportation network to characterize agents' movements and determine their locations at each point in time (using U.S. Department of Transportation information), and generate synthetic contact networks from the interactions of the agents and their locations. The resulting dynamic bipartite graph is consequently used to simulate the disease spread. The fidelity of such a model therefore relies in part on the credibility of the underlying artificial society and synthetic social networks. Due to the inherent complexity, all model components are separately developed and initially validated with regard to their logic (e.g., calibrating the rules governing an individual's activity schedule against available household transpiration surveys), and parameterization (e.g., calibrating daily work-schedule parameters using available census data on average worker density in the workplace). The validated units are consequently unified, and the model is calibrated against global measures of the epidemic such as temporal disease incidence patterns (Stroud et al. 2007).

Finally, a potential domain of error in assessing the validity of ABS models is due to "over-parameterization" of these models causing an "impossible falsification" effect (Klugl 2008). If the model contains too many degrees of freedom, an automatic optimizing calibration procedure will always be able to fit the model to the data — thus empirical validation is not sufficient. So calibration shares some common aspects with V&V in terms of operational procedures in accomplishing both. However, whereas calibration pertains primarily to the model *inputs*, V&V is rather concerned with the model *outputs*, in particular how closely these outputs align with corresponding metrics from extant real-world systems.

5 EXPERIMENTATION

Once the model is verified and validated, it can be used as a tool for addressing a wide variety of questions relevant to project goals. In the context of epidemic ABS models, this can include questions regarding future predictions of epidemic growth under various scenarios, studying the impact of specific healthcare interventions on future trends, etc. The models can be coupled with optimum-seeking strategies to address more advanced questions such as epidemic resource allocation (e.g., best allocation of a limited budget among alternative healthcare interventions to control the epidemic (Kasaie and Kelton 2013b; Kasaie and Kelton 2013a). Moreover, as with any modeling practice, a typical goal is to study the sensitivity of results with regard to changes in input parameters and assumptions implemented in the model.

In practice, most simulation analysts often use a "hit-or-miss" method for experimenting with simulation models in which they simply try a number of system configurations unsystematically and observe the resulting behaviors. In such a case, a thorough deduction is possible only through a complete exhaustion of all possibilities, which can be time-consuming or actually impossible in any practical sense in many settings. An alternative is application of statistical experimental designs. Statistical design of (simulation) experiments (DOE) is a systematic approach for designing a statistically valid experimental strategy to obtain the required information (for answering the question of interest) with the least amount of simulating (minimal effort). DOE has a well-developed acceptance and track record of success for DES models (Kleijnen 2008; Sanchez and Wan 2009) and some applications of ABS models in the military (Lucas et al. 2003) but could be used to great advantage in other application areas of ABS, notably epidemiology (Wu et al. 2013; Blower and Dowlatabadi 1994).

DOE is a highly useful technique in the context of ABS models that can be used to uncover details about model behavior, assess the relative importance of inputs, provide a common basis for discussing simulation results, and identify problems in the program logic (Kleijnen et al. 2003). Due to some fundamental differences between assumptions made conventionally in DOE and ABS models, however, a straightforward application of DOE to ABS models may not always be appropriate (Sanchez and Lucas 2002). For example, traditional DOE assumptions involve only one response variable, whereas agent-based models typically include many performance measures of interest. Nevertheless, application of DOE can still provide useful information about model behavior that would not be known without a DOE (e.g., with respect to the identification of factor interaction effects).

Sanchez and Lucas (2002) discuss various approaches for more efficient application of DOE to ABS models, including use of more effective measures for data collection using various statistical designs to sample the parameter space efficiently. A typical issue in analysis of the effect of input parameters in ABS models is caused by the large number of free parameters, so that the search spaces that have to be searched by the calibration algorithms is often so big that they cannot be searched exhaustively in any reasonable amount of time (Fehler, Klügl, and Puppe 2005). LHS is a statistical method for generating a sample of plausible collections of parameter value from a multi-dimensional distribution, and provides an efficient approach for sampling large parameter spaces. Another important issue in application of DOE to ABS is regarding the choice of factor levels: for example, if a narrow range of change is imposed on an important factor, but a wide range on an unimportant factor, then the latter could appear be more important than the former (Happe 2005). Moreover, there is an emphasis on application of visual guides for gaining insight into the one-way relationships as well as internal interactions of various factors, such as regression trees, 3-dimensional surface plots, neural networks, etc.

There are further important considerations for treating stochastic simulation models including, but not limited to, choice of the simulation time scale for terminating vs. steady-state goals, followed by specification of transient/warmup/run-length period (in the case of a steady-state goal), choice of the number of simulation replications driven by precision requirements for outputs, use of multiple ranking-and-selection statistical tools for selection of best options from a finite set of alternatives, application of variance-reduction techniques for improving precision, optimum-seeking via metaheuristic approaches, etc. Such DES-mature ideas address fundamental aspects of experimentation with any stochastic simulation model, and therefore are necessary, but not necessarily sufficient, to applications of ABS models.

Finally, one should note the application of various heuristic techniques (e.g., genetic algorithms, evolutionary algorithms, simulated annealing) as efficient alternatives to classical methods for calibration or optimization of simulation models. In such applications, there is a tradeoff between the optimality, completeness, and accuracy of results vs. execution time, which can still favor application of these heuristic methods.

6 CONCLUSIONS

ABS is a relatively novel simulation technique that is growing in popularity and number of applications across diverse fields. Following a bottom-up modeling approach, ABS provides a flexible and powerful tool for modeling complex systems composed of many interacting components, such as those involving human behavior. Due to its comparatively short history and lack of a standardized educational paradigm, there is a wide discrepancy in design and especially experimentation and analysis of ABS models. Focusing on applications to infectious-disease epidemic modeling, the literature contains several instances of ABS exercises lacking the basic practices familiar to a DES-mature audience.

In this paper, we provided a short discussion of some important notions concerning design and analysis in ABS studies that may serve as a guideline for simulation modelers new to ABS. We discussed issues in the design and development of ABS models for appropriate choice of modeling paradigm, abstraction levels, and internal mechanisms used to describe agents' behavior and interactions. We highlighted the important issue of parameterization for detailed ABS models, often requiring specialized calibration methods for reducing the dimension or size of the parameter space. With regard to various sources of error, we discussed methods of verification for addressing modeling mistakes, and provided a short discussion of hierarchical validation techniques applicable to ABS models. Finally, we surveyed a few important notions for experimentation with ABS models using DOE and LHS for efficient analysis of parameter-rich models, as well as further important considerations for the choice of the simulation time scale, number of replications, variance reduction, etc. We provided examples of ABS models of infectious disease epidemics, as relevant to our area of expertise, but the discussion is relevant to the general class of ABS models and applications in other fields.

We don't pretend to have provided here anything like a comprehensive literature survey or a complete prescription for best practices in ABS projects. But our hope is that ABS modelers, both new and seasoned alike, might take away from this paper some considerations and suggestions that could improve their modeling design and parameterization, V&V, as well as design and analysis of the experiments to be carried out using their models as testbeds, to the end of making better decisions.

REFERENCES

Abuelezam, N.N., K. Rough, and G.R. Seage. 2013. "Individual-Based Simulation Models of HIV Transmission: Reporting Quality and Recommendations." *PLoS ONE* 8:e75624.

Balci, O. 1994. "Validation, Verification, and Testing Techniques throughout the Life Cycle of a Simulation Study." *Annals of Operations Research* 53:121–73.

Balci, O. 1995. "Principles and Techniques of Simulation Validation, Verification, and Testing." In *Proceedings of the 1995 Winter Simulation Conference*, edited by D. Goldsman, C. Alexopoulos, K. Kang, and W.R. Lilegdon, 147–54. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Begun, M., A.T. Newall, G.B. Marks, and J.G. Wood. 2013. "Contact Tracing of Tuberculosis: A Systematic Review of Transmission Modelling Studies." *PloS ONE* 8:e72470.

Bharathy, G.K., and B. Silverman. 2010. "Validating Agent Based Social System Models." In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, 441–53. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Blower, S.M., and H. Dowlatabadi. 1994. "Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example." *International Statistical Review/Revue Internationale de Statistique*, 229–43.

Bonabeau, E.. 2002. "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems." In *Proceedings of the National Academy of Sciences of the United States of America*, 7280–87. Washington, DC: National Academy of Sciences.

Borshchev, A., and A. Filippov. 2004. "From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools." In *The 22nd International Conference of the System Dynamics Society*. Oxford, England: Systems Dynamics Society.

Coleman, H.W., and W.G. Steele. 1989. *Experimentation and Uncertainty Analysis for Engineers*. New York: John Wiley & Sons.

Del Valle, S.Y., P.D. Stroud, J.P. Smith, S.M. Mniszewski, J.M. Riese, S.J. Sydoriak, and D.A. Kubicek. 2006. "EpiSimS: Epidemic Simulation System." Los Alamos Unlimited Release (LAUR). Los Alamos, New Mexico: Los Alamos National Laboratory.

Dimatteo, R.M., P.J. Giordani, H.S. Lepper, and T.W. Croghan. 2002. "Patient Adherence and Medical Treatment Outcomes: A Meta-Analysis." *Medical Care* 40: 794–811.

Epstein, J.M., D.M. Goedecke, F. Yu, R.J. Morris, D.K. Wagener, and G.V. Bobashev. 2007. "Controlling Pandemic Flu: The Value of International Air Travel Restrictions." *PLoS ONE* 2 (5).

Fabretti, A. 2012. "On the Problem of Calibrating an Agent Based Model for Financial Markets." *Journal of Economic Interaction and Coordination* 8: 277–93.

Fehler, M., F. Klügl, and F. Puppe. 2005. "Techniques for Analysis and Calibration of Multi-Agent Simulations." In *Engineering Societies in the Agents World V*, 305–21. Berlin-Heidelberg: Springer.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, et al. 2006. "A Standard Protocol for Describing Individual-Based and Agent-Based Models." *Ecological Modelling* 198: 115–26.

Guzzetta, G., M. Ajelli, Z. Yang, S. Merler, C. Furlanello, and D. Kirschner. 2011. "Modeling Socio-Demography to Capture Tuberculosis Transmission Dynamics in a Low Burden Setting." *Journal of Theoretical Biology* 289: 197–205.

Happe, K. 2005. "Agent-Based Modelling and Sensitivity Analysis by Experimental Design and Metamodelling: An Application to Modelling Regional Structural Change." In *XIth International Congress of the European Association of Agricultural Economists*. The Hague: European Association of Agricultrual Economists.

Heath, B., R.R. Hill, and F. Ciarallo. 2009. "A Survey of Agent-Based Modeling Practices (January 1998 to July 2008)." *Journal of Artificial Societies and Social Simulation* 12: 9.

Helbing, D., and S. Balietti. 2013. "How to Do Agent-Based Simulations in the Future: From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design." In *Social Self-Organization*, edited by Dirk Helbing, 25–70. Berlin: Springer.

Henderson, S., and B.L. Nelson. 2006. *Handbooks in Operations Research and Management Science: Simulation*. Amsterdam: Elsevier.

Kasaie, P., and W.D. Kelton. 2013a. "Simulation Optimization for Allocation of Epidemic-Control Resources." *IIE Transactions on Healthcare Systems Engineering* 3: 78–93.

Kasaie, P., and W.D. Kelton. 2013b. "Resource Allocation for Controlling Epidemics: Calibrating, Analyzing, and Optimizing an Agent-Based Simulation." *IIE Transactions on Healthcare Systems Engineering* 3: 94–109.

Kennedy, M.C., and A. O'Hagan. 2001. "Bayesian Calibration of Computer Models." *Journal of the Royal Statistical Society, Series B* 63: 425–64.

Kim, S.J, W. Kim, and L.R. Rilett. 2005. "Calibration of Microsimulation Models Using Nonparametric Statistical Techniques." *Journal of the Transportation Research Board 1935, no. 1*: 111–119.

Kleijnen, J.P.C. 2008. *Design and Analysis of Simulation Experiments*. Berlin-Heidelberg: Springer Science.

Kleijnen, J.P.C., S.M. Sanchez, T.W. Lucas, and T.M. Cioppa. 2003. *A User's Guide to the Brave New World of Designing Simulation Experiments*. Tilburg, The Netherlands: Tilburg University.

Kleijnen, J.P.C. 1995. "Verification and Validation of Simulation Models." *European Journal of Operational Research* 82: 145–62.

Klugl, F.. 2008. "A Validation Methodology for Agent-Based Simulations." *SAC 08 (March)*: 16–20.

Latora, V., A. Nyamba, J. Simpore, B. Sylvette, S. Diane, B. Sylvere, and S. Musumeci. 2006. "Network of Sexual Contacts and Sexually Transmitted HIV Infection in Burkina Faso." *Journal of Medical Virology* 78: 724–29.

Law, A.M., and W.D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd ed. New York: McGraw-Hill.

Lucas, T.W., S.M. Sanchez, T.M. Cioppa, and A.I. Ipekci. 2003. "Generating Hypotheses on Fighting the Global War on Terrorism." In *Maneuver Warfare Science*, edited by G. Horne and S. Johnson, 107–24. Quantico, Virginia: USMC Project Albert.

Macal, C.M, and M.J. North. 2010. "Tutorial on Agent-Based Modeling and Simulation." *Journal of Simulation* 4: 151–62.

Macal, C.M, and M.J. North. 2013. "Successful Approaches for Teaching Agent-Based Simulation." *Journal of Simulation* 7: 1–11.

Rahmandad, H., and J. Sterman. 2008. “Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models.” *Management Science* 54: 998–1014.

Read, M., M. Tripp, H. Leonova, L. Rose, and J. Timmis. 2013. “Automated Calibration of Agent-Based Immunological Simulations.” *Advances in Artificial Life, ECAL* 12: 874–75.

Sanchez, S. M., and T.W. Lucas. 2002. “Exploring the World of Agent-Based Simulations: Simple Models, Complex Analysis.” In *Proceedings of the 2002 Winter Simulation Conference*, edited by E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, 116–26. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Sanchez, S.M., and H. Wan. 2009. “Better than a Petaflop: The Power of Efficient Experimental Design.” In *Proceedings of the 2009 Winter Simulation Conference*, edited by M.D. Rossetti, R.R. Hill, B. Johansson, A. Dunkin, and R.G. Ingalls, 60–74. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Siebers, P.O., C.M. Macal, J. Garnett, D. Buxton, and M, Pidd. 2010. “Discrete-Event Simulation Is Dead, Long Live Agent-Based Simulation!” *Journal of Simulation* 4: 204–10.

Stroud, P., S. Del Valle, S. Sydoriak, J. Riese, and S. Mniszewski. 2007. “Spatial Dynamics of Pandemic Influenza in a Massive Artificial Society.” *Journal of Artificial Societies and Social Simulation* 10: 9.

Tesfatsion, L., 2015. “Verification and Empirical Validation of Agent-Based Models.” http://www2.econ.iastate.edu/tesfatsi/empvalid.htm.

Wu, J., R. Dhingra, M. Gambhir, and J.V. Remais. 2013. “Sensitivity Analysis of Infectious Disease Models: Methods, Advances and Their Application.” *Journal of The Royal Society Interface* 10: 20121018.

Yilmaz, L. 2006. “Validation and Verification of Social Processes within Agent-Based Computational Organization Models.” *Computational & Mathematical Organization Theory* 12: 283-312.

AUTHOR BIOGRAPHIES

PARASTU KASAIE is a Postdoctoral Fellow in the Department of Epidemiology at The Johns Hopkins University Bloomberg School of Public Health. Her research interests are in simulation modeling and statistical analysis of HIV and TB dynamics, with implications for policymaking. Her email address is pkasaie@jhu.edu.

W. DAVID KELTON is a Professor in the Department of Operations, Business Analytics, and Information Systems at the University of Cincinnati, and Visiting Professor of Operations Research at the Naval Postgraduate School. His research interests are in the probabilistic and statistical aspects of simulation, and applications of simulation. His email address is david.kelton@uc.edu.