Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds

MODELING AND SIMULATION GRAND CHALLENGES: AN OR/MS PERSPECTIVE

Simon J. E. Taylor

ICT Innovation Group
Department of IS and Computing
Brunel University, Uxbridge, UB8 3PH
UNITED KINGDOM

Sally Brailsford

Southampton Management School
University of Southampton
University Road, Southampton SO17 1BJ
UNITED KINGDOM

Stephen E. Chick

Technology & Operations Management Area
INSEAD, Boulevard de Constance
77300 Fontainebleau, FRANCE

Pierre L'Ecuyer

DIRO, Université de Montreal
C.P. 6128, Succ. Centre-Ville
Montréal (Québec), H3C 3J7, CANADA

Charles M. Macal

Decision & Information Sciences Division
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439 USA

Barry L. Nelson

Department of Industrial Engineering & Management Sciences
Northwestern University
Evanston, IL 60208 USA

ABSTRACT

Grand challenges are significant themes that can bring together researchers to bring significant change to a field. In 2012 a new initiative to restart the debate on major grand challenges for modeling and simulation (M&S) began. Leading researchers have presented M&S Grand Challenges in areas such as ubiquitous simulation, high performance computing, spatial simulation, big simulation, human behaviour, multi-domain design, systems engineering, cyber systems, network simulation and education. To contribute further to this initiative, this paper presents M&S Grand Challenges from an Operational Research/Management Science (OR/MS) perspective and discusses themes including simulation in healthcare, value of information, data modeling, stochastic modeling and optimization, agent-based simulation and simulation analytics.

1 INTRODUCTION

Grand Challenges can bring together researchers to revolutionize a field. One can turn to astronomy and physics to see how Grand Challenges are being met as scientists collaborate on huge problems that are enabling us to better understand the fabric of our Universe from the large to the small. Modeling & Simulation (M&S) plays a key role in these endeavors as it is in other domains. However, M&S has its own Grand Challenges that can bring great benefit to researchers and practitioners alike. The first major event to reflect on M&S Grand Challenges was the Workshop on Grand Challenge for Modeling & Simulation (M&S), held at Dagstuhl in Germany around a decade ago (Schloss Dagstuhl 2002). To refresh the debate on M&S Grand Challenges, a new initiative began in 2012. Three panels were held. These were at the 2012 Winter Simulation Conference in Berlin, Germany (Taylor et al. 2012), the Symposium on

978-1-4799-2076-1/13/$31.00 ©2013 IEEE

Theory of Modeling and Simulation (TMS'13) during SpringSim 2013 in San Diego, USA (Taylor, et al. 2013a), and at the first SIGSIM-PADS conference in Montreal, Canada (Taylor et al. 2013b).

Grand Challenge themes that are emerging are diverse: overall M&S methodology, interaction of models from different paradigms, coordinated modeling, multi-domain design, parallel and distributed simulation, agent-based M&S, ubiquitous computing, supercomputing, grid computing, cloud computing, cloud-based M&S, big data and complex adaptive systems, "big" simulation applications (data, models, systems), human behavior, model abstraction, replicability, embedded simulation for real-time decision support, systems engineering, simulation on-demand, simulation-based acquisition, simulation interoperability, simulation composability, high speed optimization, web simulation science, spatial simulation, ubiquitous simulation, cyber systems, network simulation, democratization and education.

To continue the Grand Challenge debate, this paper presents views on M&S Grand Challenges from an Operational Research/Management Science (OR/MS) perspective and discusses simulation in healthcare, value of information, data modeling, stochastic modeling and optimization, agent-based simulation and simulation analytics.

2 SALLY BRAILSFORD: ARE GENERIC HOSPITAL SIMULATIONS POSSIBLE?

People who know me well will know that one of my "hobby-horse" topics is the lack of widespread implementation of simulation modeling in healthcare systems (Brailsford, 2005). Nobody would argue with the statement that simulation has been widely applied in healthcare over the past fifty years. Eight years ago, a literature search on Web of Knowledge (WoK) found 1,008 papers describing simulation models of Emergency Departments (Brailsford, 2005). Brailsford et al. (2009) report a wider literature search carried out in 2007 on JSTOR, SCOPUS and WoK, on the search string "*(health-care OR health care) AND (modelling OR modeling OR simulat*OR (system AND dynamic*) OR markov**)", appearing in the title, abstract or keywords. This search resulted in 176,320 hits, and when repeated over successive days the body of literature was found to be expanding at the rate of 30 papers per day. A recent systematic review by Hulshof et al. (2012) contains 658 references. Yet despite this massive literature, reviews of simulation in health dating back to the 1980's (Wilson, 1981; Jun et al., 1999; Fone et al., 2003; Hulshof et al., 2012) all report that despite a plethora of one-off applications in the academic literature, very few papers report the practical outcomes of implementation or sustained adoption of these models. Indeed, Fone et al. (2003) comment that "*… we were unable to reach any conclusions on the value of modelling in health care because the evidence of implementation was so scant*." (p333).

Of course, it is not true to say that simulation modeling is not being used in healthcare systems at all: for example, much of the work in the UK (and, I suspect, the US and other countries as well) is undertaken by business consultancies and therefore does not get reported in the academic literature. However, it is undeniably true that simulation and modeling methods are not routinely embedded in the healthcare management toolbox in the same way that they are in the military or in manufacturing industry (Naseer et al., 2008). This appears to be a global problem, regardless of the way that healthcare is financed, organized and delivered in any particular country.

There are many theories as to why this should be the case. Is healthcare special in some way? Of course, when we are modeling a hospital we are not modeling inanimate widgets on a production line but human beings who are often in extremely stressful situations. The culture in healthcare organizations is definitely different, although surely not so vastly different from the culture in other service organizations where safety is critical, such as the airline industry. Human behavior cannot be neglected, even in manufacturing models.

It is perfectly understandable that academic research should be published before the work has had time to achieve its full benefit. Academic career incentives (publish or perish) are such that in order to gain promotion, researchers have to publish quickly and then move on to the next project. There have been very few incentives to follow up with research users on implementation, although in the UK the 2014 Research Excellence Framework (the process by which the Government assesses the quality of each university's research and allocates future funding) will, for the first time, include "impact" as one of the assessment criteria. UK universities are now frantically chasing their non-academic collaborators for evidence that their research has been used in practice. Moreover, in order to publish in top-quality peer-reviewed journals, academics have to produce models which push the boundaries of research knowledge. These can take years to develop and are often highly complex. Hospital managers however require models (by next week!) which do not require specialist software, can be easily understood and explained to colleagues, and can be easily modified and updated without the need to call the modeler back to write more code. Business consultants of course have other objectives: it is not in their commercial interest to make their models transparent, easy to use or let alone to teach hospitals how to modify them.

Another theory is the "not invented here syndrome". While in some domains people are keen to adopt models and methods which have proved successful elsewhere, there seems to be a barrier in healthcare to doing this. Healthcare modelers often hear the argument that a model which worked for Hospital X would not work for Hospital Y, because "our patients are different, our staff are different, our wards are different, our systems are different…." One wonders how true this really is. Do we really need 1,008 different models of Emergency Departments? Surely all EDs have a great deal in common and have many broadly similar processes. However, in many healthcare modeling interventions, it is the process of stakeholder engagement, and not the end product, that is most effective (Brailsford et al. 2007). In other words, the model itself is often almost just a vehicle for getting the right people together round a table and talking through the problem. So, and this is my "grand challenge", is it necessary to go through this lengthy and painful process every time, or is it possible to develop a generic model, or a suite of generic building blocks, which can be used in plug-and-play mode to speed up the process? Gunal and Pidd (2007) present an example of such a suite of interconnecting models. However, I am not aware that this has been widely implemented since then.

Bowers et al. (2012) discuss this issue and compare three different approaches to modeling. Firstly, a bespoke simulation model developed for an ED in Fife, Scotland; secondly, the re-use of this model (with minor adaptation) for an orthopedic outpatient clinic in the same hospital; and thirdly, a "generic" ED model, developed jointly by Lancaster University and the UK Department of Health (Fletcher and Worthington, 2009) and intended to be freely available for use by any National Health Service hospital. Bowers et al. comment that "*Where resources allow, the bespoke route can be attractive*." (p1465), suggesting that in an ideal world, this would still be the method of choice. The generic model had all sorts of issues with acceptance, and Bowers et al. state that "*There is a danger that even if it is technically valid, staff may be reluctant to accept any intervention on the basis of a simulation that is perceived to represent external practice.*" (p1465). The re-use model appears to be a "*reasonable compromise*" (p1465), as it still allowed an element of stakeholder engagement, although there were still some issues with acceptance, and of course it was only re-used within the same hospital….

My challenge therefore is, is it possible to build a generic model, or a suite of generic building blocks, which are software-platform independent, and easily understood by clinical and managerial stakeholders in healthcare systems, and which would enable hospital simulations to be built rapidly and re-used, and would lead to wide uptake and acceptance of simulation models? Could simulation models ever be as pervasive as spreadsheets in healthcare organizations?

3 STEPHEN CHICK: VALUE OF INFORMATION AND DATA MODELING

In what is described here, I focus on some end-user benefits and possibilities enabled by trends in data and networking. Certainly, however, I won't claim to be an expert in them all. Grand challenges by definition require a contribution from a broad variety of experts. I have several somewhat disjointed observations to make about the role of simulation in decision making. The observations are influenced by my path: which a couple decades ago included work in materials handling engineering in the automotive sector doing discrete event simulation, to a transition to the faculty in industrial engineering with work on manufacturing and on epidemic transmission systems, and to a decade ago to a business school.

Value of information. If simulation is to be used in a decision making process, some specific benefit must be derived. Decisions that are made on the basis of any decision support tool, including simulation, are often made in order to improve some metric. If we define benefits as an increase in such a metric, then the development and use of a decision support tool can be considered to be an option to gain information that reduces uncertainty about what decision is best. Quite a bit of work has focused on this from statistical and from stochastic optimization perspectives, with various goals in mind. Bayesian, decision-theoretic work takes the perspective that simulation sampling gives the option to learn about alternatives, that sampling therefore reduces the risk of an incorrect selection (in expectation), although it may come at a cost. There are opportunities in this line of work in several directions. One opportunity is that formal mechanisms to link more narrowly defined operational objectives to broader economic objectives (with potentially risks that are not commonly modeled in simulations today). For example, how does average throughput rate influence the economic impact of a decision taken with simulation? This means stepping back from operational details of a system being modeled and making the link to macro implications of decisions (does the increase in production rate actually improve the firm's 5-year discounted NPV?).

Another direction is to extend current theory on expected value of information approaches, which has seen 'good' results in a number of contexts, to very large scale problems (combinatorial number of alternatives) where correlation structures in beliefs of the performance of alternatives and in random variate usage may be at play. This value of information thus raises a number of other interesting theoretical issues, involving sequential sampling and sequential decision making in environments with many decision variables and complex correlation structures. The dynamic programming approach (and related approximations) to sequential learning in simulation contexts has shown some interesting progress, but more remains to be done. This may prove useful not only in improving the speed at which complex simulated systems can be optimized, but might also prove useful in terms of spill-over effects to sequential clinical trial or health technology assessment applications. Another area worthy of exploration is a greater emphasis of the role of risk of decisions (from stochastic risks associated with random outcomes even when parameters are known, from epistemic risks associated with uncertainty about what the parameters of a system are, and third from environmental risks associated with the fact that entire systems are changing around us). I have spoken about stochastic and epistemic risks in past WSCs. This third risk, environmental risk, I use to refer to scenario planning exercises that firms are increasingly using to characterize the many shifts, often unpredictable, in environments. This third risk seems to offer opportunities to link simulation in the discrete sense to simulation in the business sense of developing monitoring and action plans associated with industrial, economic and other structural uncertainties. That is, can simulation be used to improve 'optimal' system control in a non-stationary sense: the optimal control being a policy that adapts control to events that arise.

A very different benefit of the simulation modeling process is its use in getting people to share a common vision to a set of activities (which will hopefully be organized into a process). The process of articulating assumptions in a group to develop a common vision and goal structure should not be underestimated when improving the functioning of effective processes. This so-called soft operations research has incredible practical value, and is by no means easy. In addition to research within the field of simulation and operational research, scholars in the field of strategy (strategy execution) and organizational behavior (team alignment, etc) have also studied this problem. One 'challenge' is to extend the reach of current theory across these domains to provide greater theory (falsifiable hypotheses) as to when simulation initiatives (of the group decision making process or organizational alignment sort) do or do not create value.

Data modeling. Rare is the day when samples are truly independently sampled. While a joint independence assumption for stochastic outcomes may be a useful approximation to reality, there is often correlation across activities in service or manufacturing systems: a patient with a complicated operation is likely to have complicated billing activities: service times in the operating room are likely correlated with service times in administrative processes. Moreover, there may be correlations in time (a tired worker might have several service times that are slow in a row; a server with a long queue of customers might work faster), and across people. An example of autocorrelation across people is that infectious disease outcomes are correlated across individuals due to the nonlinear dynamics of disease transmission. These correlations can be shown to lead to very different conclusions for some important decision problems (such as some in infectious disease control) as compared to a model with independence assumptions. Modeling of correlation and of 'real' distributions is therefore important. Social network theory is increasingly being applied to indicate that behavior may be correlated for other phenomenon besides disease transmission: such theory attempts to explain why outcomes also appear correlated in some non-communicable disease categories (e.g., peer pressure in social networks about smoking; or behavior for activity in groups that may reduce risks of obesity and related comorbidities). Developing effective data models of human behavior in our increasingly networked world seems to be a worthwhile goal. This could involve theoretical statistical models, estimation models, and efficient variate generation tools. The agent based world has made some neat progress here, and there seems to be more that can be done. This is particularly true with the open data movement and other massive data sets that can describe actual human behavior.

In summary, I think there are a number of interesting opportunities for simulation modelers in the areas of use of massive amounts of data, of development of theories with falsifiable hypothesis of human interactions (both in the decision making process and in the human interaction processes in simulated human systems), and in the area of linking value creation and risk management goals of the decisions being made with the use of simulation as a decision support tool.

4 PIERRE L'ECUYER: STOCHASTIC MODELING AND OPTIMIZATION

Available computing power to perform simulations increases relentlessly and fast, mostly due to the possibilities for parallel and distributed simulation, via cloud computing, multicore processors, and general purpose graphical processing units (GPGPUs), for example. Larger and more complicated models can be simulated, or more simulation runs can be performed to obtain tighter confidence intervals on certain unknown quantities. But these results can be meaningless and misleading if the model is not sufficiently realistic and accurate. Simulation can be useful only to the extent that we can trust the model (Law 2007). Building trustable stochastic models of large complicated systems is one of the biggest challenges we face in modeling and simulation research. A key reason for this is that there are many sources of uncertainty, which are generally not independent, and the dependence is often very hard to model. In my experience, lousy and unconvincing stochastic modeling, often with lack of appropriate data given as an excuse, is (unfortunately) a major weakness in many application-oriented articles submitted to simulation journals.

In parallel with the increasing compute power, huge amounts of data are currently becoming available at a rate never seen before and that increases exponentially. Exploiting this massive data to build credible and valid stochastic models of complex systems leads to statistical modeling and estimation problems related to current research in Bayesian statistics, data mining, and machine learning, for example. Bayesian statistics are involved when we need to effectively update the models as new data comes in. This is typically needed when simulation is used periodically, in a dynamic fashion, to update decisions. In fact, Monte Carlo simulation methods are often used in these areas for parameter estimation and for "learning" the models based on available data (Robert and Casella 2004). It is important to emphasize that realistic modeling is generally much more complicated than selecting univariate distributions and estimating their parameters. Model inputs are often multivariate distributions and stochastic processes, with hard-to-model (but important) dependence between them.

Let us give some examples of large stochastic systems that are difficult to model. As a first example, consider modeling and simulating the climate change on our planet, to study the potential regional impacts. For this we need (among other things) three-dimensional dynamic models for concentrations of greenhouse gases such CO2, CH4, N2O, freons, water vapor, for the air and water temperature, water currents, surface ice and snow, the clouds, etc. This involves uncertainty which is difficult (although essential) to model.

Such difficulties also occur for example if we want to simulate a whole human body or one of its parts, or the spread of infectious diseases by virus or bacteria in populations of humans, animals, and plants, with resistance of bacteria to antibiotics, or social networks and crowd behavior, where stochastic behaviour of interacting humans is involved, or communication and transportation networks, which involve various devices and stochastic demands and travel times, or stochastic models in finance, where external knowledge might be better exploited and where dependence between asset prices in a portfolio (beyond correlations) can have a large impact on risk, for example, or health-care facilities (clinics, hospitals, ambulances, patient evolution), where stochastic aspects are ubiquitous.

To examine a few concrete illustrations more closely, consider a telephone call center, where calls arrive one by one according to some stochastic process. Such centers currently employ around 3% of the workforce in North-America, so they are important economically (Aksin, Armony, and Mehrotra 2007, Koole 2013). They are used by small and large organizations for marketing and sales, customer service, billing and recovery, public services, emergency calls, taxis, fast food ordering, etc.

Important stochastic aspects to model in those centers include for example the call arrival processes, the service time distributions, and the abandonments. Simplistic modeling often assumes that the arrivals are from a stationary Poisson process, that the service times are i.i.d. exponential random variables, and that patience times (the time the customer is ready to wait for an agent before hanging up and abandoning) are also i.i.d. exponentials. But these assumptions are highly unrealistic.

Arrival rates in call centers are highly non-stationary; they depend on the time of the day, day of the week, type of day (holiday, special day, etc.), and other seasonal effects. The arrival rates are also highly stochastic. With deterministic time-varying rates, the number of arriving calls within any given time period(the arrival count) would have a Poisson distribution, with variance equal to the mean, but the observed variance of the counts over periods of a few hours is usually much larger than the mean, sometimes by factors of over 50, and these counts are also dependent across periods (Avramidis, Deslauriers, and L'Ecuyer 2004; Channouf and L'Ecuyer 2012; Ibrahim, L'Ecuyer, Régnard, and Shen 2012; Oreshkin, L'Ecuyer, and Régnard 2013; Steckley, Henderson, and Mehrotra 2005).

To account for this high dispersion, one may assume that the arrival rate is random, for example constant in each 30-minute time period, and with some dependence structure across time periods that could be modeled via a copula. One difficulty in estimating such a model is that we do not observe these rates, but only the arrival counts. This generally makes maximum likelihood parameter estimation much more difficult. In fact, the available information is often only the arrivals count in each time period. Incoming calls are also partitioned into different types, often several dozen types or more, where each type can only be handled by a subset of the agents that have the corresponding skills, and the arrival processes are dependent across call types (Jaoua, L'Ecuyer, and Delorme 2013). This brings a further level of difficulty. Modeling this dependence across periods and across call types simultaneously would give rise, in general, to highly-dimensional copulas that are very challenging to estimate.

Service time distributions (call durations) depend of course on the call type but also on the individual agents, with means that can sometimes vary by up to a factor of 2 or 3 between two agents for the same call type (Gans, Liu, Mandelbaum, Shen, and Ye 2010; Ibrahim and L'Ecuyer 2013). It may also depend on the number of call types that the agent is handling, and may change with time (due to learning effects, motivation and mood of the agents, etc.). These can be important factors to model and to consider when making work schedules for agents, but modeling them properly is not easy.

Similar arrival-process and service time modeling problems occur in other settings that involve humans, such as customer arrivals at stores, incoming demands for a product, arrivals at hospital emergency, etc. The uncertainty in those settings tends to be more important and more complicated than in industrial automated systems that involve only machines. Modeling it in a trustable way is still a significant challenge (and is likely to remain so for many years ahead) in several areas of application.

5 CHARLES MACAL: AGENTS – THE FUTURE OF SIMULATION?

5.1 Perspective

To frame our discussion, I'd like to ask: *What is simulation?* Because people often think of different things when discussing simulation, Pritsker (1979) was one of the first to present a compilation of definitions of simulation, and Oren (2011) ultimately came up with about 400 definitions. Based on the commonalities of these definitions, one could conclude that a *simulation is a model of a dynamic process*, i.e., a process that is "animated" over time, and often, more recently, over space (Law 2007). This notion of simulation is broad enough to encompass traditional discrete event simulation (DES), system dynamics (SD), and agent-based simulation (ABS). If uncertainty is included in the simulation, Monte Carlo (MC) simulation plays an essential role (Rubinstein and Kroese, 2008). I also believe that *simulation is what "simulationists" do -- or will be doing in the future*. By simulationists I mean simulation scientists, engineers, or just regular people that develop or use computer simulation models. One way to look at a grand challenge is to enumerate the things related to simulation that people are already spending, or intending to spend, a lot of time and effort on, perhaps pointing out a few bottlenecks to progress and problems that might require focused efforts of the simulation community in the future.

5.2 Grand Challenges

With these arguments in place, we can ask equivalently, *What is the future of simulation?*, here focusing on the OR/MS perspective, and *What will simulationists be doing in the future?* There is much discussion in the simulation community about all possible pair-wise combinations of the relationship among DES, SD, ABS, and MC simulation. Each approach carries its own community, including specialists, accepted practices, software, and common understandings of what constitutes good models and good applications. I propose the first grand challenge is for the simulation and modeling community to come to a common understanding of the various ways of how these simulation approaches fit together in applications to solve important problems of national and global significance.

But to my main point, I believe the future of simulation is agent-based simulation. In the future, simulationists will be building, using, and reporting results from agent-based simulations to decision and policy makers; simulationists will team with scientists across many disciplines to build agent-based simulations and use them as electronic laboratories to advance scientific knowledge. In the future, all simulations will be agent-based simulations, or agent-based simulations will be components of larger simulation systems. In the course of developing many agent-based simulation applications, I have observed that it is not unusual in the current environment to be approached by potential collaborators, customers, and clients who are interested in developing agent-based simulation. A common refrain is: *"We do not know what agent-based simulation is, … but we know we need it!"* (As to why people think they need agent-based simulation, and whether it is appropriate or not, is beyond the scope of the present discussion.)

We cover details of agent-based simulation elsewhere (Macal and North, 2010). In a nutshell, agents have behaviors and states upon which their behaviors are based. An agent's behavior, if truly dynamic and giving an agent the capability to sense the environment, deliberate, and respond to interactions with other agents, is based on the state of the agent at a particular time. An agent-based simulation consists of a set of agents, a set of agent interactions, and the mechanisms that update the agent states as a result of the agents' behaviors and interactions. It is worth noting that agent-based simulation overflows the boundaries of traditional simulation, and in my opinion ABS is a more general modeling technology. Agent interaction algorithms are the basis for decentralized system optimization algorithms, such as ant colony optimization and particle swarm optimization (Barbati, Bruno and Genovese, 2011).

I will enumerate four areas I think are important for the future of agent-based simulation: behavior (including connections to Big Data and Data Analytics), emergence, design, and the micro-to-macro connection.

5.2.1 Behavior

One of the reasons that people come to agent-based modeling is because they would like to include truer representations of the *behaviors* of entities, or agents, into their models (e.g., Ferguson 2007). An agent-based model can be used to investigate whether modeling agent behavior matters at all, what micro-level agent behaviors would improve macro-level system performance, and when collective behaviors may emerge among interacting agents, among other interesting research questions. Here, there is a connection between ABS and Big Data Analytics. For example, Kosinaki et al. (2013) demonstrate how behavioral attributes (potentially akin to market segmentation and an antecedent to developing behavioral models) can be identified from digital records. The inclusion of real-time data for updating agent states and blending that information with projections or forecasts of agent activities is another promising area (Bengtsson et al. 2011). Ultimately, there appears to be a natural motivation to model and simulate all of society, which would have a variety of beneficial applications and uses (Epstein and Axtell, 1996). A fanciful vision of possibilities is provided by Asimov, when Hari Seldon invents the field of "psychohistory" that enables him to predict precisely the aggregate behaviors of large populations of people (Asimov 1988). Also interesting is the media's reaction to mining social network data (The Economist 2013). Individual-based agent models approaching a global scale have already begun to be developed (Parker and Epstein 2011).

5.2.2 Emergence

I see very little recognition of, or appreciation for, emergence that occurs in ABS by the M&S community. Emergence refers to the generation of order in the form of patterns and structures that self-organize endogenously within a model. ABS originated in the fields of study called *complex adaptive systems* and *artificial life* (Macal 2009). In these fields, emergence and adaptation of natural systems are central concepts. People often come to ABS to study emergence by modeling "systems from the ground up," meaning that the agents and their interactions produce structures and patterns that they have not intentionally programmed into their models. For example, in the famous "Boids" model (Reynolds 1987), mobile agents interact with their neighbors through simple behavioral interaction rules that influence each other's speed and direction. Most people find it surprising that for certain parameter settings related to the strength of agent interactions, such agents form coherent swarms, reminiscent of fish schooling or birds flocking, that had not been explicitly programmed in to the model. Most agent-based models exhibit some form of emergence whether it is recognized or not by the modelers. A grand challenge is to develop methods that recognize endogenous emergent effects in models, create higher-level structures of the kind that would emerge in real-world processes, and have those structure endogenously interact with agents.

5.2.3 Design

Agent-based models often just happen – they are not the result of a deliberate design process -- but they should be and will be in the future. Agent-based modeling should become more and more of an exercise in assembling pre-designed (and potentially validated) modules into a coherent model to solve a problem. Enormous progress has been made in the computer science community, for example, in object-oriented design, where the objective of reusability has resulted in the efficient development of programs consisting of millions of lines of highly reliable computer code. In the agent-based realm, specifications such as the ODD (Overview, Design concepts, and Details) protocol provide a framework for communicating model purpose and design (Grimm et al. 2006). The key principle of design is to separate model specification from model implementation. In other words, the model design is a standalone entity, and implementation can proceed in any computer language according to the unambiguous design specification. Conceptually, every decision made in the design process can be encapsulated and preserved for future use by others (North and Macal 2013).

5.2.4 Micro to Macro

Another reason that people come to agent-based simulation is the challenge of connecting the micro (agent behaviors) to the macro (population and system-wide behaviors). Agent-based modeling brings with it explicit models for agent behaviors, which may not be present in other types of simulations. Analytic treatments of the micro-to-macro connections have so far been unavailable, but there may be room for progress, either analytically or computationally. Related to this situation, is the need for breakthrough developments for the proper treatment of uncertainty in simulation models, specifically methods that related model inputs, data, and parameters to a scientifically valid characterization of model outputs. Characterizing uncertainty of simulation model outputs that correctly characterizes the uncertainties surrounding simulation input data as well as structural uncertainties in the model, such as those stemming from the uncertainties around behavior representations, are grand challenges. One approach to this may be to develop equivalent classes of more aggregate and less computationally demanding meta-models that mimic large-scale stochastic models within estimated uncertainty bounds.

5.3 CONCLUSION

The grand challenges above are among many challenges for agent-based simulation in the future. Other grand challenges such as agent-based simulation validation, with the need to validate embedded behavioral models as well as emergent processes, will ensure there is much exciting work to do in simulation for many decades to come.

6 BARRY L. NELSON: SIMULATION ANALYTICS

As a simulation research and practice community we need to rethink the reporting of simulation results, or perhaps the tools we give users to create their own reports. At first blush this may seem like less than a "grand challenge," but I am firmly convinced of two things: (1) Many, if not most, simulation users misinterpret or misunderstand the results they are getting now; and (2) this misinterpretation or misunderstanding is not harmless. The most straightforward evidence for (1) is that if users really understood what they were getting then they would almost certainly be asking for more, and as far as I can tell they are not. I will say more about (2) later.

If I am correct then this is discouraging. Simulation is increasingly the *only* method capable of analyzing, designing, evaluating or controlling the large-scale, complex, uncertain systems in which we are interested. The available software is up to the *model-building* task, and the use of animation can inspire confidence that, say, an integer nonlinear programming formulation never will. However, the success of data analytics in providing fine-grained, effective and profitable prescriptions based on transactional data is going to put increasing pressure on *all* mathematical models, including stochastic simulation, to do the same. It is not a stretch to describe what we do as "data analytics on conceptual systems," that is, systems for which no actual data yet exists. Why should simulation be held to a lesser standard? We have probably been able to slide by mostly reporting means and confidence intervals because, until now, no one expected the real-world implementation to behave exactly as the simulation predicts, and even badly interpreted simulation results nudged users toward better choices. That may no longer cut it.

I want to be clear that I am not beating up on the software vendors. The research community has not made a strong enough case for why this matters, nor have we provided very many good solutions. In this section I will identify a few of the challenges we should address.

Here is the briefest possible summary of inferential statistics, and probably what we should assume as the starting point for many simulation users: You have a data set $y_1, y_2, \ldots, y_n$ on some variable quantity Y and your goal is to estimate one or more properties of Y. In the simulation context Y is a system performance output and $y_1, y_2, \ldots, y_n$ are simulation-generated observations of it. The long-run average and percentage chance something happens are familiar properties. To provide a concrete context I will refer to this real-world example: *Prior to the U.S. elections of 2012 there was concern about whether the design and staffing of polling places would allow everyone to exercise their right to vote with tolerable waiting. Historical voting patterns including numbers of voters by precinct, actual voter turnout, and the percentage of the vote cast by time of day were available, as well as ballot-dependent models of the time to cast a vote. Simulations were constructed to assess poll performance for any given staffing level, equipment capacity or registration policies.*

Here are some reasons why reporting the results of stochastic simulation experiments is not easy:

Simulations generate a lot of data: Even this simple example generates delay data for each of hundreds to thousands of voters; congestion data for the polling place, the registration desk and the voting devices; and utilization data for poll workers and voting devices. A complex manufacturing or supply chain simulation could easily increase the amount of data by one to two orders of magnitude. The immense volume of data is one reason that simulation software invariably provides summary reports, typically averaging everything in sight; but averaging introduces its own problems.

Averaging *through time* masks time-dependent effects: A critical feature of voting is time-varying load. In some precincts voting primarily takes place before or after working hours; but a precinct serving a retirement community will experience a very different distribution of load. The delay experienced by voters, averaged throughout the voting day, will hide the long delays experienced by some voters during peak load periods by averaging them with brief delays for voters off-peak. *While averaging across replications is always statistically helpful, averaging within replications may be bad.*

The performance that interests us is often related to *time*: Recall our generic output data $y_1, y_2, \ldots, y_n$. These might be voter waiting times during the simulated day. To be very specific, in a typical simulation these could be the waiting times of n voters *in the order that they completed waiting*. Notice first that the index is *not* the order of arrival to the polling place, nor the order of departure from it. But even if the output data were sorted in one of these two ways, it is unlikely that we are interested in, say y_{32}, the waiting time of the 32nd voter of the day to arrive. Instead, the anticipated waiting time as a function of the *time* of arrival is what is relevant. For instance, what delay might a voter expect who arrives at noon as opposed to 5 PM? *I am willing to argue that for nearly all systems with time-varying load, measures of delay should be indexed by time.* But this a more subtle measure to define and estimate.

A related problem is reporting results for continuous-time outputs, like queue length $Q(t)$. These outputs are naturally indexed by time, so it is easy to generate a within-replication time plot of $Q(t)$ vs. t as nearly all simulation languages do. Averaging such data across replications to get a better estimate of mean queue length as a function of time is certainly meaningful, but how do you do it? The most straightforward way requires saving lots of data from each replication (the values of $Q(t_i)$ and t_i at each point in time t_i when it changed) and doing a substantial amount of post-run computation. If a measure of variability around this mean queue length is reported, as it should be, then there is even more work. Yet this is exactly the sort of information that would be meaningful in the voting model.

Simulations are conducted in a sequential, exploratory manner: In classical design of experiments, important factors are identified, levels of the factors are chosen, and all of these predefined experiments are executed and the results analyzed. However, since simulations are not physical experiments, data ("runs") are often cheap, and results can be summarized and displayed immediately. Thus, it is natural to conduct some simulation studies in a more exploratory fashion. In the voting simulation we might pick a particular layout and staff, make a few runs, look for problems, and adjust the staffing and capacity where we see congestion or idleness. How do we summarize a user-guided sequence of experiments on different scenarios so as to insure that they did not miss something or get mislead?

We focus on reporting error instead of risk: We have a historical focus on measures of error while decisions should usually be based on measures of risk; even worse, users often confuse error with risk. A confidence interval (CI) is a measure of error; measures of error answer the question "have we done enough simulation to estimate whatever?" The answer should be related to the specific context, like "I am comfortable being off by ±$1000, but not ±$50,000." The width of a CI says nothing about the natural system variability; that is, it says nothing about risk, although many people think that it does. A prediction interval (PI) is a measure of risk; it answers the question "what performance will we most likely see when we implement?" Measures of risk should be standard, but they are not. And when reporting either type of measure we should avoid statistically meaningless precision. See Nelson (2008), Song and Schmeiser (2009) and Wieland and Nelson (2009) for some ideas about how to do this.

Simulations involve multiple objectives and competing scenarios: It is a rare simulation that produces only a single output measure, and for each output there may be more than one property we care about. Trade-offs between costs, quality of service, physical space, etc. are common. For some performance measures, the mean (long-run average) and a measure of stability (standard deviation, tail percentile, PI) are both relevant. And all of these occur across multiple scenarios or system designs. This is an area in which our reporting is particularly weak. Even the formal statistical comparisons we provide typically focus on a single measure, and nearly always compare means rather than measures of risk.

This list of analytics challenges is not exhaustive, but it is a good start. We are currently giving simulation users amazing tools for generating data, but not nearly as much help in getting the most out of these data.

7 CONCLUSIONS

This paper has continued the discussion of Grand Challenges in M&S by discussing these from an Operational Research/Management Science (OR/MS) perspective. It has discussed themes including simulation in healthcare, value of information, data modeling, stochastic modeling and optimization, agent-based simulation and simulation analytics.

ACKNOWLEDGMENT

The work described in C. Macal's section work is supported by the U.S. Department of Energy under contract number DE-AC02-06CH11357. P. L'Ecuyer's work is supported by grants from NSERC-Canada and a Canada Research Chair.

REFERENCES

Aksin, O. Z., M. Armony, and V. Mehrotra. 2007. "The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research". Production and Operations Management 16 (6): 665–688.

Asimov, I. 1988. *Prelude to Foundation*. Doubleday.

Avramidis, A. N., A. Deslauriers, and P. L'Ecuyer. 2004. "Modeling Daily Arrivals to a Telephone Call Center". Management Science 50 (7): 896–908.

Barbati, M., Bruno, G., Genovese, A. 2011. "Applications of agent-based models for optimization problems: a literature review," *Expert Systems with Applications*, doi: 10.1016/j.eswa. 2011.12.015

Bengtsson L., X. Lu, A. Thorson, R. Garfield, and J. von Schreeb. 2011. "Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti." *PLoS Med* 8(8): e1001083.

Bowers J., Ghattas M. and Mould G. 2012. Exploring alternative routes to realising the benefits of simulation in healthcare, *Journal of the Operational Research Society*, 63:1457-1466.

Brailsford S.C. 2005. Overcoming barriers to the implementation of OR simulation models in healthcare. *Journal of Clinical Investigative Medicine*, 28:312-315.

Brailsford S.C., Bolt T., Connell C., Klein J. and Patel B. 2009. Stakeholder engagement in health care simulation. In: Rossetti MD, Hill RR, Johansson B, Dunkin A and Ingalls RG (eds). Proceedings of the 2009 Winter Simulation Conference. Austin, TX. IEEE: Piscataway, NY, pp 1840–1849.

Brailsford S.C., Bolt T.B., Bucci G., Chaussalet, T.M., Connell N.A., Harper P.R., Klein J.H., Pitt M. and Taylor M. 2013. Overcoming the barriers: a qualitative study of simulation adoption in the NHS. *Journal of the Operational Research Society* 64: 157-168.

Brailsford S.C., Harper P.R., Patel B and Pitt M. 2009. An Analysis of the Academic Literature on Simulation and Modelling in Healthcare. *Journal of Simulation* 3:130-140.

Channouf, N., and P. L'Ecuyer. 2012. "A Normal Copula Model for the Arrival Process in a Call Center". International Transactions in Operational Research 19:771–787.

Epstein, J. M., and R. Axtell. 1996. *Growing Artificial Societies: Social Science From The Bottom Up*. Cambridge, MA: MIT Press.

Ferguson, N. 2007. "Capturing Human Behavior." *Nature* 446:733, April.

Fletcher A. and Worthington D. 2009. What is a 'generic' hospital model?—a comparison of 'generic' and 'specific' hospital models of emergency patient flows. *Health Care Management Science* 12: 374–391.

Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G and Palmer S, 2003. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health Medicine* 25: 325–335.

Gans, N., N. Liu, A. Mandelbaum, H. Shen, and H. Ye. 2010. "Service Times in Call Centers: Agent Heterogeneity and Learning with some Operational Consequences". Manuscript.

Grimm, V., et al. 2006. "A standard protocol for describing individual-based and agent-based models." *Ecological Modelling* 198 (1-2), 115-126.

Gunal M.M. and Pidd M. 2007. Interconnected DES models of emergency, outpatient, and inpatient departments of a hospital. In: Henderson SG, Biller B, Hsieh M-H, Shortle J, Tew JD, and Barton RR (eds). Proceedings of the 2007 Winter Simulation Conference. Washington, DC. IEEE: Piscataway, NY, pp 1461–1466.

Hulshof P.J.H, Kortbeek N., Boucherie R.J., Hans E.W. and Bakker P.J.M., 2012. Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. *Health Systems* 1: 129-175

Ibrahim, R., and P. L'Ecuyer. 2013. "Forecasting Call Center Arrivals: Fixed-Effects, Mixed-Effects, and Bivariate Models". Manufacturing and Services Operations Management 15 (1): 72–85.

Ibrahim, R., P. L'Ecuyer, N. Régnard, and H. Shen. 2012. "On the Modeling and Forecasting of Call Center Arrivals". In Proceedings of the 2012 Winter Simulation Conference: IEEE Press. paper inv243.

Jaoua, A., P. L'Ecuyer, and L. Delorme. 2013. "Call type dependence in multiskill call centers". Simulation. see http://sim.sagepub.com/content/early/2013/04/01/0037549713479405.

Jun J.B., Jacobson S.H. and Swisher J.R. 1999. Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society* 50: 109–123.

Koole, G. 2013. Call Center Optimization. MG books, Amsterdam.

Kosinaki, M., D. Stillwell, and T. Graepel. 2013. "Private records and attributes are predictable from digital records of human behavior," *Proc. National Academy of Sciences*, 110(15): 5802-5805 (April).

Law, A. M. 2007. Simulation Modeling and Analysis. Fourth ed. New York: McGraw-Hill.

Macal, C. M., 2009. "Agent-based modeling and artificial life," in *Encyclopedia of Complexity and System Science*, Robert A Meyers (ed.), pp. 119-131, Springer.

Macal, C. M., 2010. "To Agent-based Simulation from System Dynamics." in *Proceedings of the 2010 Winter Simulation Conference*. B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, eds. pp. 371-382. Wiley-IEEE Press.

Macal, C. M., and M. J. North. 2010. "Tutorial on agent-based modelling and simulation," *Journal of Simulation*, Special Issue on Agent-Based Modelling 4(3): 151–162, September.

Naseer A., Eldabi T. and Jahangrian M. 2008. Cross-sector analysis of simulation methods: A survey of defense and healthcare. *Transforming Government* 3:181–189.

Nelson, B. L. 2008. "The MORE Plot: Displaying Measures of Risk & Error from Simulation Output. *Proceedings of the 2008 Winter Simulation Conference*, Edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 413–416. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

North, M. J., and C. M. Macal. 2013. "Product and Process Design Patterns For Agent-Based Modeling," *Journal of Simulation*, April.

Ören, T. 2011. "A Critical Review of Definitions and About 400 Types of Modeling and Simulation." *Modeling & Simulation Magazine.* Society for Computer Simulation. No. 3 (July).

Oreshkin, B., P. L'Ecuyer, and N. Régnard. 2013. "Rate Based Arrival Process Models for Modeling and Simulation of Call Centers". manuscript.

Parker, J., and J. M. Epstein. 2011. "A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission" *ACM Trans. Model. Comput. Simul.* 22(1): 1-25.

Pritsker, A. A. B. 1979. "Compilation of Definitions of Simulation." *Simulation*. 33: 61-63 (August). doi:10.1177/003754977903300205.

Robert, C. P., and G. Casella. 2004. Monte Carlo Statistical Methods. second ed. New York, NY: Springer-Verlag.

Reynolds, C. 1987. "Flocks, herds and schools: A distributed behavioral model." ACM SIGGRAPH '87: Proc. 14th Annual Conference on Computer Graphics and Interactive Techniques, pp. 25–34, doi:10.1145/37401.37406.

Rubinstein, R. Y., and D. P. Kroese. 2008. *Simulation and the Monte Carlo Method*. 2nd ed. Wiley.

Schloss Dagstuhl. 2002. www.dagstuhl.de/02351.

Song, W. T., and B. W. Schmeiser. 2009. Omitting meaningless digits in point estimates: The probability guarantee of leading-digit rules. *Operations Research* 57:109–117.

Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2005. "Performance measures for service systems with a random arrival rate". In Proceedings of the 2005 Winter Simulation Conference, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 566–575: IEEE Press.

Taylor, S.J.E., Fishwick, P.A., Fujimoto, R., Page, E.H., Uhrmacher, A.M., Wainer, G. 2012. Panel on Grand Challenges for Modeling and Simulation. In *Proceedings of the Winter Simulation Conference 2012*. ACM Press, NY.

Taylor, S.J.E., Khan, A., Morse, K.L., Tolk, A., Yilmaz, L, Zander, J. 2013a. Grand Challenges on the Theory of Modeling and Simulation. In *Proceedings of the 2013 Symposium on the Theory of Modeling and Simulation*. SCS, Vista, CA. *To appear.*

Taylor, S.J.E., O. Balci, W. Cai, M. Loper, D. Nicol and G. Riley. 2013b. Grand Challenges in Modeling and Simulation: Expanding Our Horizons. In *Proceedings of the 2013 SIGSIM-PADS Conference, to appear.*

The Economist. 2013. "*Dr Seldon, I presume*." Feb. 23.

Wieland, J. R., and B. L. Nelson. 2009. "How Simulation Languages Should Report Results: A Modest Proposal." In *Proceedings of the 2009 Winter Simulation Conference*, Edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 709–715. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wilson J.C.T. 1981. Implementation of computer simulation projects in health care. *Journal of the Operational Research Society* 32: 825–832.

AUTHOR BIOGRAPHIES

SALLY C. BRAILSFORD is Professor of Management Science at the University of Southampton, UK. She received a BSc in Mathematics from the University of London, and MSc and PhD in Operational Research from the University of Southampton. Her research interests include simulation modeling methodologies, system dynamics, health service research and disease modeling, and the modeling of human behavior in healthcare systems. She is chair of the European Working Group on OR Applied to Health

Services (ORAHS) and is an Editor-in-Chief of Health Systems. She is on the editorial boards of Health Care Management Science, the Journal of Simulation and Operations Research for Health Care. Her email address is s.c.brailsford@soton.ac.uk.

STEPHEN E. CHICK is a professor of Technology and Operations Management at INSEAD, and the Novartis Chair of Healthcare Management. He has worked in the automotive and software sectors prior to joining academia, and now teaches operations with applications in manufacturing and services, particularly the health care sector. He enjoys Bayesian statistics, stochastic models, simulation, and the music of the Grateful Dead. His email address is stephen.chick@insead.edu.

PIERRE L'ECUYER is Professor in the DIRO, at the Université de Montréal, Canada. He holds the Canada Research Chair in Stochastic Simulation and Optimization. He is a member of the CIRRELT and GERAD research centers. His main research interests are random number generation, quasi-Monte Carlo methods, efficiency improvement via variance reduction, sensitivity analysis and optimization of discrete event stochastic systems, and discrete-event simulation in general. He is currently Editor-in-Chief for ACM Transactions on Modeling and Computer Simulation, and Associate/Area Editor for ACM Transactions on Mathematical Software, Statistics and Computing, International Transactions in Operational Research, and Cryptography and Communications.

CHARLES M. MACAL, PhD, PE, is a senior system engineer at Argonne National Laboratory and Senior Fellow of the Computation Institute, University of Chicago. He has developed many applications of agent-based simulation, discrete event simulation, system dynamics, and optimization models of economic equilibrium systems, and is a principal investigator for the Repast Simphony agent-based modeling toolkit (macal@anl.gov).

BARRY L. NELSON is the Walter P. Murphy Professor and Chair of the Department of Industrial Engineering and Management Sciences at Northwestern University. He is a Fellow of INFORMS and IIE. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems. His e-mail and web addresses are nelsonb@northwestern.edu and www.iems.northwestern.edu/˜nelsonb.

SIMON J E TAYLOR is the Founder and Chair of the COTS Simulation Package Interoperability Standards Group (CSPI PDG) under SISO. He is the Editor-in-Chief of the Journal of Simulation. He was Chair of ACM SIGSIM from 2005 to 2008. He is a Reader in the School of Information Systems, Computing and Mathematics at Brunel and leads the ICT Innovation Group. He has published over 150 articles in modeling & simulation and distributed computing. His recent work has focused on leading the development of Grand Challenges in Modeling & Simulation and the development of standards for distributed simulation and cloud computing in industry. His email address is Simon.Taylor@brunel.ac.uk.