Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

ANALYSIS METHODOLOGY: ARE WE DONE?

Sigrún Andradóttir
David Goldsman

School of Industrial & Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

Bruce W. Schmeiser

School of Industrial Engineering
Purdue University
W. Lafayette, IN 47907, U.S.A.

Lee W. Schruben

Dept. of Industrial Engineering & Operations Research
University of California, Berkeley
Berkeley, CA 94720, U.S.A.

Enver Yücesan (Moderator)

INSEAD
Boulevard de Constance
77305 Fontainebleau Cedex, FRANCE

ABSTRACT

Since 1967, the Winter Simulation Conference has been a forum for the introduction of innovative approaches to effectively analyze discrete-event simulation experiments. The goal of this panel is to bring together key contributors to analysis methodology research in order to clarify areas that they think are essentially complete, and identify areas that need more work. In doing so, we hope to help provide direction to younger researchers looking for the "right" problems to work on.

1 SIGRÚN ANDRADÓTTIR

This write-up provides brief descriptions of three broad problems within the analysis methodology area that are highly deserving of additional research. These three problems are all receiving a considerable amount of attention by the research community at the current time, and I believe this will and should continue in the next several years. The problems fall within the input/output analysis, variance reduction, and optimization sub-fields of simulation, respectively.

The first problem involves the development of enhanced methods for ensuring that the results obtained from a simulation experiment are accurate in that they in fact satisfy the properties they are stated to have. For example, does the confidence interval for the value of a performance measure of interest in fact cover the unknown value of that quantity with the stated probability? This can be difficult to ensure due to various sources of errors and approximations arising in the modeling, implementation, and analysis processes. Consequently, there is great need for enhanced techniques that assess the actual coverage of confidence intervals and produce confidence intervals that in fact do have the stated coverage. Also of importance are techniques that help a practitioner determine when it is appropriate to terminate a simulation experiment and expect the results produced to have the desired properties. Finally, techniques that reduce the errors and approximations made in the modeling, implementation, and analysis processes, such as improved methods for modeling and generating dependent random variables and stochastic processes, would be valuable contributions.

Another broad area of interest involves the development of new and improved techniques for increasing the efficiency of simulation experiments. Although computers are constantly becoming more powerful, it is well known that the precision of simulation estimates only increases slowly with the computational effort provided. Consequently, it is important for our community to continue developing smart techniques that yield more precise estimators without increasing the associated computational budget (e.g., through a better allocation of sampling effort, or through smoothing the performance measure over the sample space, and consequently reducing its variance). This is particularly important in the context of rare event simulation.

The final topic I wish to comment on in this write-up is the area of simulation optimization. This research area has received a considerable amount of attention in recent years, but given the current state of the art and the fact that the ultimate goal of many (if not most) simulation studies involves some amount of system design or optimization, it is crucial for our community to continue working towards the development of improved techniques for optimizing systems through the use of simulation. This includes both enhanced techniques for solving optimization problems that possess very little structure (beyond a certain amount of smoothness in the performance measures of interest over the parameter space) that are consequently suitable for inclusion in general purpose simulation languages, and also techniques that exploit special structure present in the problem at hand to complete the optimization process more efficiently (in fact, the identification of classes of simulation optimization problems that are of sufficient importance to warrant special attention and at the same time possess structure that can be exploited by an optimization algorithm is by itself a worthy research contribution).

2 DAVID GOLDSMAN

The last thirty years have brought forth significant advances in the ways that we conduct formal mathematical analyses of simulation processes. In this position statement, I will discuss some of the traditional analysis methodology research areas that we have "put to bed," along with some others that have emerged or have become invigorated. I will break things down into discussions involving (i) the statistical aspects of modeling and generation of simulation processes, and then (ii) the analysis and optimization of simulation processes.

2.1 Generating Randomness

A simulation must be driven by random variates that adequately mimic the real-world (or virtual) system under consideration. In particular, these random variates ought to be generated quickly and must follow the probability distributions that they are supposed to represent. Traditionally, we have subdivided this category of research into two pieces: generation of uniform(0,1) random numbers, and generation of random variates from all other distributions. This makes sense, since the uniforms are the building blocks that are used to generate everything else. However, I will also add a third, emerging subdivision – random process generation.

2.1.1 Uniform Generation

In past years, one could easily be satisfied with a simple, reliable uniform generator such as $X_{i+1} = 16807\, X_i \bmod(2^{31}-1)$, which is fast, has reasonable independence and uniform properties, and has a period of about two billion numbers before cycling occurs. But times have changed, and we often need many, many more random numbers for our simulation models. The good news is that there are now a number of quick, statistically nice, extremely long-cycling random number generators on the block, often with periods on the order of 2^{100}; see, e.g., L'Ecuyer (1999). So, practically speaking, the problem of providing an adequate supply of uniforms is solved. Perhaps in the spirit of calculating decimals of π, it will be interesting to see how much farther researchers can go in this area.

2.1.2 Random Variate Generation

Simple generation methods for the standard distributions have been available for a long time. Any good simulation textbook such as Law and Kelton (2000) (and even many statistics texts) will provide a list of easy-to-use algorithms, ranging from simple inverse-transform-based techniques to more-sophisticated combinations of techniques. There is still probably some call for the continued development of even-more-efficient random variate generation methods for standard univariate distributions, but, for the most part, it seems that much of the important work has been done.

So what can one do if a non-standard distribution is encountered? The easiest option is to punt and work with the empirical distribution. On the other hand, Wagner and Wilson (1996) present an easy-to-use graphical method, based on rigorous theory, that allows users to interactively model "complicated" distributions. This type of methodology reflects an encouraging trend.

The problem of generating univariate random variates serves as a stepping stone for some timely extensions to the multivariate / time series case.

2.1.3 Random Process Generation

Nowadays, it is more important than ever to be able to generate random vectors and/or time series quickly and efficiently. This type of work is finding its way into the textbooks, but perhaps a starting point is Schmeiser and Song (1989), which details a number of clever techniques for producing various common stochastic processes commonly encountered, for example, the M/M/1 queue waiting time process.

It is apparent that applications areas are driving the types of stochastic processes under study. Melamed (1991) proposed the TES methodology for the modeling of telecommunications traffic, which is often ``bursty'' in nature. Along the way, he found that he could generate stochastic processes having fairly general (but approximate) correlation structure with given marginals. Another seminal reference along the same lines is Cario and Nelson (1997), who propose the NORTA (NORmal To Anything) technique for generating stochastic processes with desired correlation and marginal properties. With the rise of quantitative analysis on Wall Street, an important application area lies in financial analysis. Here, we might be interested in generating geometric Brownian motion or its extensions for use in certain options pricing models. See Glasserman (2004) for a compendium of relevant techniques.

The point is that this is an ongoing active research area, which will undoubtedly be fruitful for years to come.

2.2 Analysis of Randomness

Assuming that we are satisfied by the quality of the simulation model and of the random variate generation techniques driving the model, the task now boils down to analyzing subsequent simulation output, and then, possibly, attempting to optimize the model. Here, I will break the discussion into two parts: analysis of univariate output and optimization methods.

2.2.1 Output Analysis Techniques

Suppose for simplicity that we concentrate on the problem of estimating the variance of the sample mean – which leads to statements about the precision of the sample mean as well as confidence intervals for the mean. Tremendous progress continues to be made in this mature area – even with respect to old tried-and-true methods such as batch means.

Over the last few years, researchers have looked at an amalgamation of methods involving batch means, overlapping batch means, standardized time series, spectral, and regenerative estimators (among others) for variance of the sample mean. A number of general trends seem to be emerging, all of which are made possible by the computer's ability to do more work in less time:

- Overlapping (in fact, resampling) estimators are good. Beginning with the classic paper Meketon and Schmeiser (1984), researchers have found that the variance of non-overlapping batched variance estimators is almost always reduced by overlapping, while bias properties are preserved. Problems with computation inefficiencies are being removed, so this opens the door for the development of new low-variance estimators (Damerdji, Henderson, and Glynn 1997).
- Dynamic procedures. Fishman and Yarberry (1997) and Steiger et al. (2005) propose procedures that dynamically call for more and more data until particular stopping criteria are met – for example, has a certain confidence interval achieved the desired precision? All indications show that procedures such as these perform well in practice, which bodes well for the adoption of such techniques in commercial software packages.
- Ties with ranking-and-selection and optimization methods, to be discussed next.

2.2.2 Ranking-and-Selection and Optimization

Assuming that we are capable of conducting a reasonable analysis involving a univariate measure such as a steady-state mean, the next task might be to find the simulation alternative that possesses the ``best'' such measure. This question can sometimes be addressed via a ranking-and-selection formulation, or in some other optimization framework.

Ranking and selection was first formulated as a methodology in the early 1950's. It seeks to answer the question of which of a (limited) number of options is the best? Unfortunately, along the way the field fell into disfavor among classical statisticians, and is only now enjoying a resurgence, thanks in part to some of the work done in the simulation field. In the context of simulation, we are interested in addressing problems involving correlated data – an application which was ignored by the statistics community – and this have paved the way for the development of innovative procedures. One that immediately comes to mind is Kim and Nelson (2001), which gives a fully sequential procedure for "indifference-zone" selection in steady-state simulation systems. This procedure is much more efficient than the 1970's-style procedures heretofore adopted for use in simulation.

Another area for advance concerns simulation optimization, surveys of which are widely available. Instead of selecting among a few options, simulation optimization typically searches a large parameter space. There is a large body of literature on the topic, but a particularly interesting paper that merges ranking-and-selection techniques with optimization is that of Boesel, Nelson, and Kim (2003).

The bottom line is that this entire line of research appears to be wide open for future development.

3 BRUCE SCHMEISER AND LEE SCHRUBEN

3.1 The Past

To answer whether we are done, we first must consider what we have been doing. Discrete-event simulation analysis methodology inherited much of its context from queueing theory. In establishing academic credibility, simulation researchers needed to demonstrate that they had something to offer in addressing the "interesting" questions posed by researchers in queueing theory, such as estimating mean values and confidence intervals for steady-state performance measures, determining the optimal numbers and types of resources in stationary queues, and estimating the steady-state probabilities of delays or rare events (such as buffer overflows). Simulation analysis methodology had to demonstrate that, at least from a practical point of view, simulation models, when subjected to rigorous analysis, can produce system-performance estimators that are as good as those obtained from queueing theory.

Reviewing the challenges in Conway's (1963) classic paper, we see that many of the issues in analyzing steady-state queues have been resolved more-or-less satisfactorily. We have workable methods for initializing and terminating runs, methods for constructing reasonably accurate confidence intervals for steady-state performance measures, and methods for optimizing stationary simulation responses.

3.2 The Present

For practical purposes, simulation experiments have replaced queueing analysis in the market place of solving real-world problems. Although most methods fail in extreme situations, by and large there are well-established simulation methods for addressing most classical queueing problems. Simulation analysis techniques have found their way into the hands of many (but not enough) simulation professionals, textbooks, and commercial software packages.

Most major university programs have one or more faculty members who participate in simulation-analysis research. Most credible university degree programs have one or more courses in simulation that go beyond model building to teach the effective design and analysis of simulation experiments. Most simulation textbooks (as distinct from software manuals) have reasonable coverage of analysis issues and methodologies. Simulation-analysis researchers have established intellectual credibility with their academic colleagues in applied probability and statistics.

In most business and engineering disciplines, ad-hoc simulation modeling and heuristic analysis approaches have had considerable impact. Simulation is also a primary tool in national and industrial research laboratories. In a report from the Lawrence Berkeley Labs, simulation was given a status "on a par with theory and physical experimentation" (Yarris 1996). Examining the programs for the national conferences in Biological, Mechanical, Electrical, Civil, Industrial, and Chemical Engineering, one sees a clear dominance of simulation as the fundamental methodology for problem-specific research and practice.

Although simulation methodology is an established academic research area, there is still work to do. Despite simulation-research successes, at the INFORMS National Meeting, there are still more papers on queueing than on simulation. But if the goal was to replace queueing theory as the primary tool for mean-performance analysis of steady-state service and production systems, then yes, we are done.

3.3 The Future

The academic foothold for simulation has been established, but we are yet to reach our full potential. The questions addressed in simulation analysis methodology research can be extended beyond those formulated by queueing theorists to some really hard questions. For example: Given a complex simulated system, determine whether it is stable.

Perhaps a question that is more appropriate than: "are we done?" is "have we started?" Now that simulation methodology research is a viable path to university tenure, we would like to see academia embrace more ad-hoc simulation methodologies, temporarily foregoing mathematical rigor (when it has to be established) at the expense of efficient engineering solutions to pressing and important practical problems. This does not mean that the research questions asked and problems posed should not be well-defined with explicit specification of measures of what constitutes answers and solutions. It is embarrassingly hard to find explicit statements of fundamental simulation research questions: "the initialization bias problem", "the run duration problem", or "the input modeling problem". With well-defined, meaningful, problem statements in hand, we would then like to see more integration of simulation modeling and analysis methodology research with the decision, social, and behavioral sciences.

There is now very little research on appropriate analysis methodology for many (if not most) modern systems-analysis problems. The literature has little to offer concerning systems that are never in steady state. Not only are the underlying driving processes for these systems changing in time, the rules and laws governing system behavior may also be changing. In the parlance of simulation, the systems we should be studying have dependent input processes, are open with state-dependent feed forward and feedback, and have non-stationary performance measures that are often, at least in the long-run, unstable. None of the commercial discrete-event simulation modeling packages and few of the output-analysis methodologies address these issues. Indeed, many problems that researchers need to address have not yet been satisfactorily defined. In posing new research questions, interactions between simulation-analysis methodology researchers and simulation professionals will become even more important than in the past.

In practice, simulations are used in *forecasting*: trying to asses what might happen if certain decisions are made. Confidence intervals for steady-state performance measures are of little practical value. Indeed, simulation output analysis within the dynamic decision context of most real-world systems, with changing and unknown risks and rewards, is almost non-existent. There is much that simulation has to offer for which reliable analysis methodologies are yet to be developed. Here are four.

1. Input modeling. Analysis methodologies need to be developed to better study the environments in which a future system might operate. The external driving processes for simulation models are typically modeled as independent identically (IID) distributed random variables. There are several easy-to-use and readily available distribution-fitting software packages on the market or included with commercial simulation languages. These packages take what is presumed to be an IID sampled data set and select, from a set of common scalar probability distributions, those that give the better values for several goodness-of-

fit statistics (with little consideration of the questionable realm of hypothesis testing for which these test statistics were developed). Unless used carefully, these packages can remove most of the useful information in the data. Real data typically have cycles and trends that might otherwise be exploited in forecasting if they were not filtered by the distribution-fitting software. We need to develop more-robust methodology for assessing how input stochastic processes may be inter and serially dependent among themselves, on time, and on the system's state and its environment, as well as how these processes might effectively be modeled.

We have inherited the homogeneous Poisson process as the workhorse for input modeling from queueing theory. We need to expand our repertory. The kind of input modeling tool we would like to see included in commercial simulation software is something along the line of multivariate, dynamic Bézier models, like those developed for scalar variates by Wagner and Wilson (1996). These curves should be dynamic in that they can change over time and/or with changes in the system state. Furthermore, to be useful they should be structured to easily facilitate sensitivity analysis.

2. Sensitivity Analysis: We would like to see it established that sensitivity analysis is always done before data collection. We are of the opinion that many, if not most, data sets are dated, distorted, dependent, damaged, deleted, or, yes, deceptive (Barton et al. 2002). Doing sensitivity analysis before collecting data tells us what might be important to know about a system's future environment. Most of the real-world data that can be collected is on processes that either do not influence decisions very much, or are so critical as to be the subject of other projects that will change them by the time the recommendations of the simulation study are implemented. Any data collected should be used to forecast what might happen in the future, not to fit probability models to the past. We don't believe explicit questions for sensitivity analysis have yet been formulated from either a research or practical point of view. For example: asking what might happen if "demand doubles" could mean that customers would arrive twice as fast, or that customers with the same arrival pattern would order twice as much. System performance is different depending on what the question means.
3. Output analysis: We also would like to see output-analysis research similarly develop beyond the study of stationary scalar processes ("easy" problems for which we have pretty good methodologies) to modeling non-stationary multivariate response surfaces over finite horizons, perhaps using innovative 3-D computer graphics. The aim might be to forecast the potential impacts of different decisions under different scenarios. Rather than merely estimating the future average performance of a system given a particular design and control strategy, however, why not the inverse problems? That is, why not ask for the likelihood of a scenario where the environment changes so that a particular decision that is good under current conditions becomes a disaster. The "robustness" research by Kleijnen and his colleagues (for example, Gaury and Kleijnen 1998) and the recent Penn State PhD thesis by Govind (2004) are along these lines.
4. Modeling and Analysis. Integrating experimentation and output analysis with modeling has important advantages. For example: time-dilation experimental techniques are easily implemented in event-scheduling models (Hyden and Schruben, 2000). Another example: the conclusion that the indirect estimation of Q from L via Little's law is a good idea could be wrong when you account for the fact that Q can be directly estimated orders of magnitude more efficiently than L using an event scheduling model than with a classical job-driven process-interaction model.

Are we done? Yes, we are finished with what simulation-analysis methodology researchers have primarily been trying to do for the past few decades: impress our academic colleagues in statistics, probability, and queueing theory. Have we started? No, we have not yet formulated explicit research questions and engineered practical integrated analysis and modeling methodologies that attack important problems in a real-world social and economic context.

REFERENCES

Barton, R. R., R. C. H. Cheng, S. E. Chick, S. G. Henderson, A. M. Law, L. M. Leemis, B. W. Schmeiser, L. W. Schruben, and J. R. Wilson. 2002. Panel on current issues in simulation input modeling. *Proceedings of the 2002 Winter Simulation Conference* (eds. E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes), 353–369. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Boesel, J., B. L. Nelson, and S.-H. Kim. 2003. Using ranking and selection to `clean up' after simulation optimization. *Operations Research*, 51:814–825.

Cario, M. C. and B. L. Nelson. 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

Conway, R. 1963. Some tactical problems in digital simulation. *Management Science*, 10: 47–61.

Damerdji, H., S. G. Henderson, and P. W. Glynn. 1997. Computational efficiency evaluation in output analysis, *Proceedings of the 1997 Winter Simulation Conference* (eds. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson), 208–215. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Fishman, G. S. and L. S. Yarberry. 1997. An implementation of the batch means method. *INFORMS Journal on Computing*, 9(3): 296-–310.

Gaury, E. G. A. and J. P. C. Kleijnen. 1998. Risk analysis of robust system design, *Proceedings of the 1998 Winter Simulation Conference* (ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan), 1533–1540. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*, Springer. New York.

Govind, N. 2004. Robust parameter design with imperfect experimental control of noise. Unpublished PhD Dissertation, The Pennsylvania State University.

Hyden, P. and L. W. Schruben. 2000. Improved decision processes through simultaneous simulation and time dilation, *Proceedings of the 2000 Winter Simulation Conference* (eds. J. A. Joines, R. R. Barton, K. Kang, and P.A. Fishwick), 743–748. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11:251–273.

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

L'Ecuyer, P. 1999. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research*, 47:159–164.

Meketon, M. S. and B. W. Schmeiser. 1984. Overlapping batch means: Something for nothing? *Proceedings of the 1984 Winter Simulation Conference* (eds. S. Sheppard, U. W. Pooch, and C. D. Pegden), 227–230, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Melamed, B. 1991. TES: a class of methods for generating autocorrelated uniform variates. *ORSA Journal on Computing*, 3:317-329.

Schmeiser, B. W. and W.-M. T. Song. 1989. Inverse-transformation algorithms for some common stochastic processes, *Proceedings of the 1989 Winter Simulation Conference* (eds. E. A. MacNair, K. J. Musselman, and P. Heidelberger), 490–496, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Steiger, N. M., E. K. Lada, J. R. Wilson, J. A. Joines, C. Alexopoulos, and D. Goldsman. 2005. ASAP3: A batch means procedure for steady-state simulation analysis. *ACM Transactions on Modeling and Computer Simulation*, 15:39-73.

Wagner, M. A. F. and J. R. Wilson. 1996. Using univariate Bézier distributions to model simulation input processes. *IIE Transactions,* 28 (9): 699–712.

Yarris, L., Computing resources available to Lab researchers, *Berkeley Lab Currents,* Lawrence Berkeley Labs, July 12, 1996.

AUTHOR BIOGRAPHIES

SIGRÚN ANDRADÓTTIR is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. She received a B.S. in Mathematics from the University of Iceland in 1986, an M.S. in Statistics from Stanford University in 1989, and a Ph.D. in Operations Research from Stanford University in 1990. Her research interests include stochastic optimization, simulation, and stochastic processes. Her e-mail address is sa@isye.gatech.edu.

DAVID GOLDSMAN is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. in Operations Research and Industrial Engineering from Cornell University. His research interests include simulation output analysis and ranking and selection. He is an active participant in the Winter Simulation Conference, having been Program Chair in 1995, and having served on the WSC Board of Directors since 2002. His e-mail address is <sman@isye.gatech.edu>, and his web page is www.isye.gatech.edu/~sman/.

BRUCE SCHMEISER is Professor of Industrial Engineering at Purdue University. His research interests center on developing methods for better simulation experiments. He is a Fellow of INFORMS, is a Fellow of IIE, and has been active within the Winter Simulation Conference for many years, including being the 1983 Program Chair and chairing the Board of Directors from 1988–1990. During the Fall of 2003, he was on sabbatical to the University of California, Berkeley.

LEE SCHRUBEN is Chancellor's Professor in the Department of Industrial Engineering and Operations Research at the University of California, Berkeley where he served as Department Chairman from 2001 to 2005. He is a Fellow of INFORMS and senior member of IIE.

ENVER YÜCESAN is Professor of Operations Management at INSEAD, Fontainebleau, France. He received his Ph.D. in Operations Research and Industrial Engineering from Cornell University. His research interests include simulation output analysis and ranking and selection. He is an active participant in the Winter Simulation Conference, having served as the Proceedings Co-Editor in 2002. His e-mail address is enver.yucesan@insead.edu.