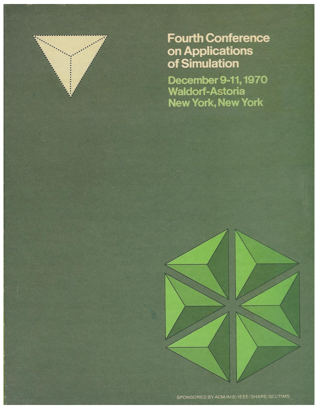
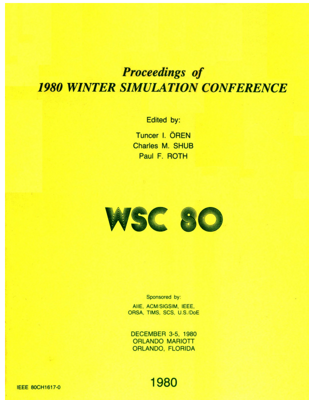
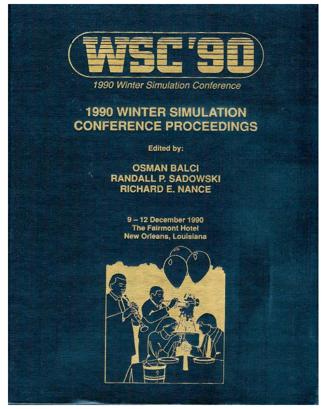
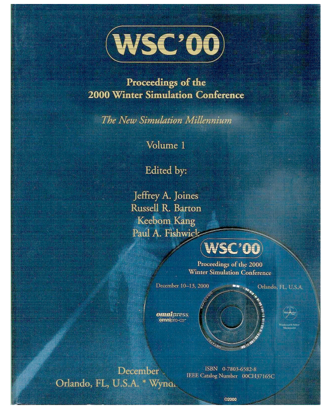
Proceedings of the 2007 Winter Simulation Conference
S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds.









FORTIETH ANNIVERSARY SPECIAL PANEL: LANDMARK PAPERS

David Goldsman

H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

James O. Henriksen

Wolverine Software Corporation
3131 Mount Vernon Avenue
Alexandria, VA 22305-2640, U.S.A.

Pierre L'Ecuyer

Département d'Informatique et de Recherche Opérationnelle
Université de Montréal, C.P. 6128, Succ. Centre-Ville
Montréal (Québec), H3C 3J7, CANADA

Barry L. Nelson

Department of Industrial Engineering and
Management Sciences
Northwestern University
Evanston, IL 60208-3119, U.S.A.

David H. Withers

8815 SE 132nd Place
Summerfield, FL 34491, U.S.A.

Nilay Tanık Argon
(Moderator)

Department of Statistics and Operations Research
University of North Carolina
Chapel Hill, NC 27599-3180, U.S.A.

ABSTRACT

To celebrate the fortieth anniversary of the Winter Simulation Conference, we have selected ten landmark papers from the four decades of the conference. In this article, we review these landmark papers and discuss their impact on the theory, education, and practice of simulation.

INTRODUCTION

This year we are celebrating the fortieth anniversary of the Winter Simulation Conference (WSC). Since its early years, WSC has maintained its reputation for being the premier international conference for disseminating advances in the field of simulation, especially discrete-event simulation. Every year, WSC gathers hundreds of simulation practitioners, researchers, and vendors from various disciplines and from the industrial, governmental, and academic sectors. Over the forty years, the number of papers presented at the conference grew considerably reaching approximately 320 papers in 2006.

To celebrate these forty productive years of WSC, we organized a special panel to discuss some of the *landmark* papers published in the 1968–2006 proceedings. After months of discussions, we selected ten papers among thousands. The selection process was difficult. We solicited nominations from the simulation community, gathered information from our colleagues, and also reviewed the WSC proceedings and citation indices. Based on this, we selected the following list of papers (in alphabetical order):

Boyle, P., M. Broadie, and P. Glasserman. 1995. Recent advances in simulation for security pricing. In *Proceedings of the 1995 Winter Simulation Conference*, C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, eds., 212–219

Glynn, P. W. 1986. Stochastic approximation for Monte Carlo optimization. In *Proceedings of the 1986 Winter Simulation Conference*, J. Wilson, J. Henriksen, and S. Roberts, eds., 356–365.

Goldsman, D., B. L. Nelson, and B. Schmeiser. 1991. Methods for selecting the best system. In *Proceedings of the 1991 Winter Simulation Conference*, B. L. Nelson, W. D. Kelton, and G. M. Clark, eds., 177-186.

L'Ecuyer, P. 1986. Efficient and portable 32-Bit random variate generators. In *Proceedings of the 1986 Winter Simulation Conference*, J. Wilson, J. Henriksen, and S. Roberts, eds., 275–277.

Meketon, M. S., and B. W. Schmeiser. 1984. Overlapping batch means: Something for nothing? In *Proceedings of the 1984 Winter Simulation Conference*, S. Sheppard, U. W. Pooch, and C. D. Pegden, eds., 227–230.

Reitman, J., D. Ingerman, J. Katzke, J. Shapiro, K. Simon, and B. Smith. 1970. A complete interactive simulation environment: GPSS/360-Norden. In *Proceedings of the Fourth Annual Conference on Applications of Simulation*, P. J. Kiviat, M. Araten, eds., 260–270.

Sargent, R. G. 1984. A tutorial on verification and validation of simulation models. In *Proceedings of the 1984 Winter Simulation Conference*, S. Sheppard, U. W. Pooch, and C. D. Pegden, eds., 114–121.

Schriber, T. J., and D. T. Brunner. 1994. Inside simulation software: how it works and why it matters. In *Proceedings of the 1994 Winter Simulation Conference*, J. D. Tew, S. Manivannan, A. A. Sadowski, and A. F. Seila, eds., 45–54

Shahabuddin, P., V. F. Nicola, P. Heidelberger, A. Goyal, and P. W. Glynn. 1988. Variance reduction in mean time to failure simulations. In *Proceedings of the 1988 Winter Simulation Conference*, M. Abrams, P. Haigh, and J. Comfort, eds., 491–499.

Pritsker, A. A. B., D. L. Martin, J. S. Reust, M. A. Wagner, O. P. Daily, A. M. Harper, E. B. Edwards, L. E. Bennett, J. R. Wilson, M. E. Kuhl, J. P. Roberts, M. D. Allen, and J. F. Burdick. 1995. Organ transplantation policy evaluation. In *Proceedings of the 1995 Winter Simulation Conference*, C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, eds., 1314-1323.

We believe that there are many papers that could be selected as landmark papers and our choices are somewhat arbitrary from this large pool. During our selection process, we paid particular attention to choosing a diverse set of papers. WSC is a unique conference for bringing practitioners, educators, researchers, and vendors together for mutual benefit. It is not primarily a conference for scholarly publication although that is certainly an important part of the program. WSC is also not primarily a trade show, but the vendors certainly play a very important role in the success of the conference. Similarly, the reports of successful practice serve not only as a base for lessons-learned, but also as inspiration and motivation for the field in general. Based on this, we selected landmark papers so that there is at least one paper in each category: (i) vendor contributions, (ii) successful practice, (iii) educational contributions, and (iv) scholarly research contributions.

In the following, we review the selected papers and discuss their importance. We start by reviewing a landmark paper with vendor contributions, which also happens to be the earliest paper that we selected (Reitman et al. 1970). We then review two famous tutorials that appeared multiple times in the WSC proceedings and continued to draw a large audience at the conference (Sargent 1984, and Schriber and Brunner 1994). Following these tutorials, we discuss Pritsker et al. (1995), which is an excellent example of a WSC paper on practice of simulation with a high impact. We continue our review with another landmark paper that focuses on an important application area of simulation — financial engineering (Boyle et al. 1995). In the second half of our paper, we review five landmark papers with scholarly research contributions at least one of which has also substantial impacts on the education of simulation. The first paper that we review in this category is Meketon and Schmeiser (1984), which is a seminal work on simulation output analysis. In the following section, we review a paper on combined congruential generators that are widely used for random number generation in simulation packages (L'Ecuyer, 1986). The next paper that we recognize is Shahabuddin et al. (1988), which is one of the earlier papers among a stream of papers on rare-event simulation in reliability estimation and state-dependent importance sampling. Finally, we review two papers that are recognized with their impact on using simulation as an effective tool for decision making and optimization (Goldsman et al. 1991 and Glynn 1986).

REITMAN ET AL. (1970) REVIEW BY J. O. HENRIKSEN

This landmark paper provides an excellent snapshot of the state of discrete event simulation in 1970. It describes a very early simulation environment that included animation, visual data input and editing, and a degree of interactive control of simulation execution. While describing problems confronted by early users of simulation, it also anticipates developments that would occur later in the 1970's.

The authors explain at some length "two opposing fronts" that were prominent in 1970. On the one hand, a great deal of effort had to be put into development of simulation technology, and on the other hand, effort had to be put into promoting the application of simulation and gaining acceptance by operations research professionals, engineers, and management. One can easily appreciate that the former dominated the latter. If technology does not exist, one must first develop it before it can be applied. The authors lament that "... the process has been too slow, and has left too many discouragements in its path." The authors advocate a very proactive approach to widening the acceptance of simulation:

"Simulation has a reputation for being one of the most expensive Operations Research techniques. Even though the cost reductions and increases in efficiency possible through simulation studies can be great, they are often indirectly achieved and not easily verifiable. Thus, confidence in the technique *follows* [emphasis added] completion of a successful project. However, approval for initial funding that some amount of confidence must exist *prior to* [emphasis added] the project. It is therefore necessary to show the potential user a proposal which [sic] he can understand and perceive its practicality."

Although we have made great progress over the last 35 years, these are still word to live by.

Let us pause for a moment to compare the tools that were available in 1970 with those of the present. Present day CPUs are faster by a factor of roughly 3,000. Improvements in software technology, e.g., compiled code vs. interpretive implementation, add another factor of 5, yielding total speed improvements of a factor of 15,000. In 1970, the vast majority of computers did not have virtual memory, and real memory capacities of more than several megabytes were considered large. Graphical displays attached to mainframes were very expensive. For example, the IBM 2250 display unit rented for around $5,000 *per month*. No wonder simulation was expensive. Today's one second run on a $1,000 laptop would have been a big deal in 1970.

In this environment, the authors pioneered the use of animation, and they developed visually-based utilities for managing simulation data. The latter capabilities would have been completely subsumed by modern spreadsheet software, but of course, that did not exist in 1970. When technology does not exist, someone has to develop it. The use of animation with simulation underwent a meteoric rise with the advent of the personal computer. Indeed, many would assume that animation began with the PC; however, the authors made productive use of animation well before the advent of the PC.

This paper provides present day simulationists with a wonderful glimpse into the past. While the authors were necessarily forced to wring maximal performance out of incredibly modest (by present standards) resources, they followed a game plan that addressed the big picture.

SARGENT (1984) REVIEW BY D. H. WITHERS

We believe that landmark papers should appeal to a wide audience and/or provide very significant contributions to a particular area of simulation and modeling. The tutorial series on verification and validation by Bob Sargent does both. Every modeler can benefit from his guidance and this series of papers constitutes a good working reference for the subject. This paper is the first of a tutorial series by Sargent and later Balci that summarize and codify the important role of verification and validation of simulation models. Sargent put rigor to what had previously been an ad hoc, informal set of rules on a little-known subject. Verification and validation had been the least understood and most frequently overlooked phase of the modeling process. The introduction to the paper positions these activities and motivates the reader to pursue to the topic:

"Simulation models are often used to aid in decision making and problem-solving. The users of these models are rightly concerned with whether the models and information derived from them can be used with confidence. Model developers address this concern through model validation and verification. Model validation is usually defined to mean 'substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model' [Schlesinger, et al. (1979)] and is the definition used here. Model verification is frequently defined as ensuring that the computer program of the computerized model (i.e., the simulator) and its implementation is correct and will be the definition used here."

Each of the papers in this series includes a practical reference to validation techniques and a guide to verification procedures. Readers (or attendees at a conference presentation) can take the information contained in one of these tutorials and proceed with a high level of competency to investigate and certify the applicability and utility of their model.

The positive impact of these papers is demonstrated by the number of citations in applications publications, see, e.g., Nayani (1998). Generally, application presentations have referred to the most recent version of the tutorial series. This seems to show continuing interest in these landmark papers. For many simulation practitioners, this tutorial series has been their introduction and their only reference to verification and validation. The presentation sessions at the conferences were always well attended and generated much positive feedback from the attendees.

We are pleased to recognize these tutorials as landmarks in the history of the Winter Simulation Conference.

SCHRIBER AND BRUNNER (1994) REVIEW BY J. O. HENRIKSEN

This sequence of papers (1994–2006) has become a fixture of the Winter Simulation Conference, and rightfully so. It has to be one of the most aptly titled works of all time.

Simulation practitioners must master a wide variety of skills to perform their work. Even slight lapses of attention to detail in any of a multiplicity of areas can lead to inaccuracies or outright errors. One of the examples I'm fond of using when teaching simulation arises in the granddaddy of all simulation models, the one-line, single-server queueing system. If we use exponentially-distributed service times, what is the maximum service time that we'll see? The algorithm used in my company's software yields a maximum value of roughly 23 times the mean. If we're running Tom Schriber's barbershop model, where the mean time for a haircut is 15 minutes, this means that a single haircut could take nearly six hours. Of course, the maximal time occurs with probability 2^{31}, so this value is "extremely unlikely." How do we know if such a value occurs? If we don't look for it, we may never know.

The authors provide standard nomenclature for describing how discrete-event simulation software works. They do so in product-neutral fashion, allowing them to describe a multiplicity of simulation software tools in unbiased fashion. Each year, they select several software tools, and they compare basic operations and describe carefully selected interesting sample problems. For example, assume that a consumer of services in a model relinquishes a server and immediately attempts to reacquire the same server. If there are other would-be consumers in a queue for the server, which consumer acquires the server? Is it the consumer who just relinquished the server? If you need to have one kind of behavior, and your simulation software by default provides another, you'll have to work around the problem to achieve the desired results.

The answers to such questions vary to a surprising extent across the software tools described. This sequence of papers has contributed to the Winter Simulation Conference in three ways:

- For the beginning practitioner, it provides an overview of how things work.
- For a practitioner who is familiar with one or two simulation tools, it provides the invaluable insight that there may be other ways of doing what they take for granted.
- For everyone, it raises a giant red flag. Sometimes things may not work the way we expect them to. Details of how software architecture and implementation can profoundly affect the correctness of a simulation. We may have a perfectly valid model, but if our implementation of the model fails to properly execute the rules embodied in the model, incorrect results will be obtained.

Exhaustively cataloguing the differences among simulation software tools, even if confined only to the most popular tools, is an impossible task. However, the authors have done the next best thing by heightening the awareness of the simulation community that differences do exist, and no matter which tool is used, there's no substitute for knowing what you're doing.

PRITSKER ET AL. (1995) REVIEW BY D. H. WITHERS

This paper is recognized as one of the landmark applications in the history of the Winter Simulation Conference for its very significant value (lives are saved) and the publicity the work received in the popular press and within U.S. government policy-making and legislative organizations. From a press release in Richmond, Virginia by UNOS (United Network for Organ Sharing) 2/7/1997:

"A unique computer simulation has been used for the first time to help establish national health care policy for organ allocation. This policy development is considered a breakthrough in addressing a major public health crisis in this country: the critical shortage of transplantable organs. The demand for organs continues to grow faster than the number of organ donors available. As a result, some people will die waiting because an organ was not available. The new, computer-based policy will increase access to transplantation and save lives ..."

The first result of the model described in this paper was a significant contribution to a modification of the national liver allocation policy. The model predicted that 100 fewer deaths annually will occur among patients who are transplanted. Previously, policy-making had relied largely on intuition, experience and judgment of those involved. Computer modeling offered a more scientific approach to policy-making, providing additional advantages of precision, quantification and probabilistic reasoning. The policy resulting from this analysis was debated by members of the transplantation community; primarily in UNOS committees.

In 1997, the UNOS Board of Directors adopted a new liver policy that rejected the sickest-patient-first approach while creating new status definitions for patients.

The regulation was published in the Federal Register as OPTN (Organ Procurement and Transplantation Network): Final Rule [1]. Congressional hearings were held to discuss and analyze the Final Rule in June 1998 before a Joint Hearing of the House Committee and Senate Labor and Human Resources Committee (Pritsker 1998).

The study reported in this paper was an excellent example of the application of best practices for simulation and up-to-date modeling technologies. The animation showing livers flying across the country is hard to forget!

The liver transplant modeling was followed by similar work on cadaveric kidneys in 2000 and Johns Hopkins University researchers have continued to use simulation to influence transplant policies for kidneys with reports as recent as 2006 (Science Daily 2006; original reference by Montgomery et al. 2006).

Based upon the value of the model in saving lives and the positive public relations resulting from this work, we are pleased to include this paper in the list of landmarks for the Winter Simulation Conference.

BOYLE, BROADIE, AND GLASSERMAN (1995) REVIEW BY D. GOLDSMAN

Financial engineering problems have come to the fore in recent years, garnering a great deal of attention in the operations research and management science literature in general, as well as the simulation literature in particular. One of the basic goals of the area is that of calculating fair values for various financial instruments such as options and other derivatives. For certain simple cases, one can come up with exact values, but this ability to calculate exact prices breaks down very quickly if we consider anything but the easiest cases. Even a straightforward instrument such as an American call option presents a problem — it is not possible to obtain closed-form value expressions for an instrument that allows the holder multiple opportunities to exercise. Luckily, such problems are amenable to analysis via clever simulation methods.

Among the contributions of the outstanding landmark paper of Boyle, Broadie, and Glasserman (1995, BB&G from here on) were that (1) it provided a terrific, clear introduction of security pricing via simulation methods to many researchers in the simulation and general operations research communities; (2) it presented a state-of-the-art tutorial for applicable methods to attack pricing problems; and (3) it set the stage for a great deal of future research in the area — in fact, there has been a boom of outstanding work in the field over the last decade, thanks in good part to this paper.

BB&G begin with an introduction to arbitrage-free markets, citing the famous Black-Scholes model as a typical model for pricing securities,

$$dS_t \;=\; \mu S_t\,dt + \sigma S_t\,d\mathscr{W}_t,$$

where S_t is the underlying stock (or other asset) price at time t, σ is the volatility, and $\mathscr{W}$ is a standard Brownian motion process (see Black and Scholes 1973 and Black 1976). Under the arbitrage-free assumption (essentially a fair and efficient market assumption, where one cannot make money "for free"), one takes $\mu = r$, the risk-free interest rate currently available. Suppose we own a European style option to buy one share of the underlying stock at time T at price K; then of course the option will pay us $(S_T - K)^+$ at time T, where $(x)^+ \equiv \max(0,x)$. It can easily be shown that the arbitrage-free expected present value of our option is

$$C \;\equiv\; \mathsf{E}[e^{-rT}(S_T - K)^+],$$

which can be evaluated in closed-form via the Black-Scholes equation (also see Hull 2006).

BB&G then go on to show how one would compute the expectation using simulation, providing wonderful motivation for attacking more-difficult problems. For example, what happens if the volatility is unknown? What if we have multiple option exercise opportunities, as given by American style options?

Since computationally intensive simulation becomes one of the weapons of choice for evaluating these more-complicated scenarios, BB&G recognize the need for efficient implementation of any simulations; and so they present a very nice tutorial on variance reduction techniques that are appropriate for use with such financial analysis problems.

Although the majority of the discussion in the paper concerns the use of Monte Carlo simulation for pricing securities, BB&G are also interested in the evaluation of price sensitivities — which almost always cannot be given by exact methods and therefore need to be computed. Such "derivatives of a derivative security's price" with respect to various model parameters are usually referred to as "Greeks" since their monikers are given by Greek letters, the most important of these being delta, the derivative of the price of a contingent claim with respect to the current price of an underlying asset. For instance, the delta of a stock option is the option price's derivative with respect to the current stock price. BB&G give an insightful discussion on the evaluation of such price sensitivity matters, after which they conclude the paper with some research topics of future interest. Indeed, these topics have yielded fruitful research in the ensuing years (see Glasserman 2004).

What makes this landmark paper especially important is that it introduced an entirely new field to many simulation researchers, while generously offering interesting problems to study — thus jump-starting an exciting and rich body of work.

MEKETON AND SCHMEISER (1984) REVIEW BY D. GOLDSMAN

Meketon and Schmeiser (1984, M&S from here on) was one of the breakthrough papers in the area of simulation output analysis, and to this day continues to garner citations in articles from archival journals.

One of the most challenging (and ongoing) problems in our field is that of analyzing the output process from a complicated stochastic simulation. The problem is that typical simulation output — such as a series of consecutive queue waiting times — never conforms to the standard assumptions of independent and identically distributed (i.i.d.) normal random variables. In fact, processes such as waiting time output are more likely to be positively autocorrelated, skewed to the right, and nonstationary. Even in the ideal case of steady-state output analysis, the task is difficult enough to render classical statistical analysis useless.

The most common goal in simulation output analysis is probably to provide information about the unknown mean μ of the underlying steady-state output process $\{Y_i : i = 1,2,\ldots,n\}$. The sample mean $\bar{Y}_n \equiv \sum_{i=1}^{n} Y_i/n$ is the obvious point estimator for μ, but we would be well-advised to provide a measure of the sample mean's precision — as a precursor to, for instance, confidence intervals for μ.

Over the years, researchers have come up with a number of techniques to analyze the output from steady-state discrete-event simulations. For example, the following well-known methods of steady-state simulation output analysis are discussed in popular textbooks such as Bratley, Fox, and Schrage (1987) and Law (2007): nonoverlapping batch means (NBM); overlapping batch means (OBM); spectral analysis; regenerative analysis; autoregressive modeling; and standardized time series (STS). These techniques can all be adapted to estimate the sampling error of the sample mean from a steady-state simulation, or almost equivalently, the so-called *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} n\mathsf{Var}(\bar{Y}_n)$; and then one can form confidence intervals for the unknown mean μ.

A common strategy used by some of the above methodologies, e.g., NBM, OBM, and STS, requires *batching* the observations. The concept of batching is simple—instead of considering the entire simulation-generated time series $\{Y_i\}$ all at once, we break up this data set into smaller batches or subseries that are composed of consecutive observations (where the batches may be disjoint or overlapping, depending on the analysis method); and then we perform the relevant analysis on each batch separately. For instance, in the NBM method we split the observations into adjacent disjoint (nonoverlapping) batches; then we assume that the resulting sample (batch) means computed from each batch are approximately i.i.d. normal; and finally we apply "standard" or "naive" variance estimation techniques to the batch means.

In the OBM method introduced in M&S's remarkable 1984 paper, we form overlapping batches, with the full realization that the associated overlapping batch means are *not* independent (though they are identically distributed and asymptotically normal):

$$\begin{array}{ll}
\text{OBM } 1: & \bar{Y}^O_{1,m} \equiv (Y_1 + Y_2 + \cdots + Y_m)/m \\
\text{OBM } 2: & \bar{Y}^O_{2,m} \equiv (Y_2 + Y_3 + \cdots + Y_{m+1})/m \\
\vdots & \\
\text{OBM } n-m+1: & \bar{Y}^O_{n-m+1,m} \equiv (Y_{n-m+1} + \cdots + Y_n)/m,
\end{array}$$

where m is the batch size. Then we compute the OBM estimator of σ^2 based on the sample variance of the overlapping batch means,

$$\mathscr{V}(b,m) \;\equiv\; c(b,m) \sum_{i=1}^{n-m+1} (\bar{Y}^O_{i,m} - \bar{Y}_n)^2,$$

where $b \equiv n/m$ and $c(b,m)$ is an appropriate scaling factor designed to make the estimator $\mathscr{V}(b,m)$ asymptotically unbiased for σ^2.

M&S recognized that this seemingly problematic technique exploits key results from the theory of spectral analysis to yield an estimator of the variance parameter σ^2 that is provably superior to the NBM variance estimator, at least asymptotically for many types of serially correlated simulation output processes.

What do we mean by a superior estimator of the variance parameter σ^2? We most often care about the bias and variance of an estimator for σ^2, as well as the resulting mean squared error (MSE), that is, the estimator's variance plus the square of the estimator's bias. Batching typically increases bias but decreases variance so that its net effect on MSE requires careful analysis. What is nice about the OBM estimator is that asymptotically it has the same bias as, but smaller variance than, the NBM estimator when the sample size $n \to \infty$. Thus, OBM gives better performance than NBM based on the same simulation-generated time series $\{Y_i : i = 1,2,\ldots,n\}$ of length n, provided that n is sufficiently large.

Besides introducing the OBM methodology in this seminal article, M&S also provided major contributions to the literature by

- Motivating how the estimator works;
- Showing that the estimator is asymptotically unbiased for σ^2 as b and m become large;

- Arguing that the estimator has an asymptotic variance that is only two-thirds that of NBM (hence, the title "Something for Nothing");
- Noting and addressing computational efficiency issues (one of the first such insights in the literature);
- Noting and addressing storage space issues;
- Establishing a link between OBM and spectral theory-based estimators, in particular, Bartlett's estimator (see also Welch 1987);
- Setting the stage for a great deal of future research in the area.

With the last point in mind, it turns out that M&S spawned a great deal of work in the area. Welch (1987) relates OBM to certain spectral estimators and looks into the effects of partial overlapping. Goldsman and Meketon (1986), Song (1988), and Song and Schmeiser (1993) derive bias and variance properties of OBM estimators, among others. Song and Schmeiser (1993) also give additional insight by plotting the coefficients of the estimators' quadratic-form representations; included in the presentation are the OBM estimators as well as overlapped versions of certain STS estimators. Further early work on the subject is undertaken by Pedrosa and Schmeiser (1993, 1994), who establish covariance properties between OBM estimators and subsequently propose a batch-size determination algorithm. In a terrific series of papers, Damerdji (1991, 1994, 1995) establishes consistency results (both in the strong and mean-square senses) for a variety of variance estimators, including OBM and an overlapping version of a certain STS estimator. In the spirit of Welch (1987), Damerdji also establishes a formal linkage between the spectral method and simulation analysis methods based on overlapping batches. And even to this day, direct descendants of M&S continue to appear in the literature, e.g., Alexopoulos et al. (2006, 2007).

L'ECUYER (1986)
REVIEW BY B. L. NELSON

Random numbers play a central role in the validity of stochastic computer simulation experiments because they provide the underlying source of randomness that insures that realizations (outputs) of the simulation are consistent with the stochastic process that the modeler intended to construct. This permits simulations to be appropriately analyzed as statistical experiments over a well-defined space of possible outcomes. Neither expert modeling of the physics of the real system nor statistically sound experiment design and analysis can compensate for defective random numbers.

Simulation experiments run in modern simulation environments invariably employ deterministic pseudorandom-number generators rather than obtaining numbers that are in any sense truly random, and this facilitates code debugging, model portability and sharper comparison of system designs via simulation. Pseudorandom-number generators produce sequences of values in the interval $[0,1]$ that, for a good generator, appear to be independent samples from the uniform distribution on that interval. They are then transformed into the service times, product demands, machine failures, etc. specified in the model.

The critical importance of pseudorandom-number generation is why L'Ecuyer (1986, L86 from here on) is a landmark WSC paper. L86 was the first published paper on pseudorandom-number generation by the prolific researcher Pierre L'Ecuyer, and it provided the foundation for a family of pseudorandom-number generators and a deep theory of such generation that followed. The impact of this work is hard to avoid: The pseudorandom-number generators in the simulation packages Arena®, Automod™, Simul8®, Witness® and probably others; in the statistical analysis package SAS®; in numerous video/gambling/lottery games; and advocated by numerous textbooks (including such staples as Law 2007, and Banks et al. 2005) trace their origins to this landmark paper.

Prior to L86, the most popular pseudorandom-number generation scheme was the multiplicative linear congruential generator (MLCG), which takes the form

$$\begin{aligned} s_i &= a s_{i-1} \mod m \qquad (1) \\ U_i &= s_i/m \end{aligned}$$

where the modulus m, the multiplier $a < m$ and the initial seed $0 < s_0 < m$ are nonnegative integers, while U_i is the ith pseudorandom number returned by the generator. Essential properties for any pseudorandom number generator are long period and apparent uniformity and independence of the generated values. Because the maximal period of such a generator (number of distinct values generated before the sequence repeats) is $m-1$, long period necessarily means large modulus m. Unfortunately, this leads to serious implementation problems on a digital computer if m is too large. Other desirable properties include portability and the facility to easily jump ahead in the generated sequence so as to imitate the existence of independent streams of random numbers. MLCGs are ideal for jump ahead, but large m is a barrier to portability.

In L86, L'Ecuyer observed that it is possible to obtain long periods—much longer than could practically be obtained via increasing the modulus m in (1)—by *combining* good MLCGs with much smaller modulus. To facilitate this he undertook a search for pairs (a,m) yielding MLCGs that exhibit good performance as measured by the spectral test (a test for uniformity) when m is prime, less than $2^{31}-1$, and $a < \sqrt{m}$, features that facilitate stable implementation.

L86 provided a general recipe for combining generators, and also a specific combination of two generators:

$$\begin{aligned} s_{1,i} &= 40692 s_{1,i-1} \mod 2147483399 \\ s_{2,i} &= 40014 s_{2,i-1} \mod 2147483563 \\ s_i &= (s_{1,i} + s_{2,i} - 2) \mod 2147483562 \\ U_i &= s_i / 2147483563. \end{aligned}$$

Combinations of this sort retain the facility to jump quickly ahead in the sequence, yet the period of this combined generator is approximately 2.3×10^{18}.

A sequence of papers, refined generators and theory for construction and testing of generators followed L86, including the popular implementation in L'Ecuyer et al. (2002). This generator, which combines three MLCGs, has period 2^{191}, which is about 3×10^{57}. To gain some understanding of just how large this period is, notice that if one could generate 2 billion pseudorandom numbers per second—the entire period of a typical MLCG—it would take approximately 4.6×10^{40} *years* to exhaust the period of this generator, a time longer than the age of the universe which is a mere 2×10^{10} years. An overview of the follow-on work and references can be found in L'Ecuyer (2006).

SHAHABUDDIN ET AL. (1988) REVIEW BY B. L. NELSON

Shahabuddin et al. (1988, SNHGG88 from here on) addresses statistical efficiency in estimating the mean time to failure (MTTF) of a highly-reliable system. The description of the problem context and approach that follows are based on Nelson (2004).

Consider using simulation to estimate the probability that a very unlikely event occurs, for instance having probability on the order of 10^{-9} of occurring. An excessive number of simulated trials would be required to observe even a small number of these events, and conducting so many trials is impossible if each trial requires even a moderate amount of simulation effort. But if the simulation model could be changed so that the rare event occurs much more often, say exactly 1,000,000 times more often, then the event would be observed more frequently, giving a much better estimator of the probability that it occurs. Of course, the estimator would be *wrong*, but since we know that it is 1,000,000 times wrong we can correct for the bias. This is essentially the idea behind the variance reduction technique of importance sampling (IS). Unfortunately, if the probability dynamics of the system are changed crudely to increase the frequency of the rare event, then the result can be a significant variance increase.

SNHGG88 considered systems of many components that are subject to random failure, but can also be repaired. When enough components in certain combinations fail, then the entire system fails. Let $\{Y_t, t \geq 0\}$ be a stochastic process with state space $\mathscr{E}$ representing the status of the components at time t, where state 0 indicates any starting state (e.g., all components functional) and there is a subset of states $\mathscr{F} \subset \mathscr{E}$ that corresponds to system failure. SNHGG88 considered systems for which Y_t is a continuous-time Markov chain with generator $\mathbf{Q}$. IS for such problems is conceptually simple: change the failure rates (and perhaps the repair rates) to make a system's failure much more likely (e.g., replace the generator $\mathbf{Q}$ by a different generator $\mathbf{Q}'$ that has larger component failure rates and smaller component repair rates). The idea is very powerful, but the reality is that finding a $\mathbf{Q}'$ that guarantees a variance reduction in a dynamic, stochastic simulation is not easy, and in fact the authors discovered that a proven IS technique for estimating steady-state system availability failed to work for estimating MTTF, motivating this paper.

Let α_B represent the first passage time of the system from state 0 to some subset of states B. Then SNHGG88 observed that

$$\mathrm{E}[\alpha_F] = \frac{\mathrm{E}[\min(\alpha_0, \alpha_F)]}{\Pr\{\alpha_F < \alpha_0\}} \tag{2}$$

where $\mathrm{E}[\alpha_F]$ is the MTTF. They demonstrated empirically that substantial variance reductions (4 orders of magnitude in the confidence interval width relative to straightforward simulation) could be obtained by applying different IS changes to estimate the numerator and denominator of (2) separately: In fact, straightforward simulation could be used to estimate the numerator, while an intuitively appealing IS strategy could be applied to estimate the denominator.

This landmark paper was one of the earliest in a long line of highly influential publications on rare-event simulation by various subsets of the authors. A summary with references can be found in Juneja and Shahabuddin (2006). In fact, a number of WSC papers could have been picked as the "landmark" paper on this topic, including Goyal, Heidelberger and Shahabuddin (1987) from the previous year.

The work in this paper also motivated Shahabuddin (1994), which won the Nicholson Prize from the Institute for Operations Research and the Management Sciences (INFORMS) as the best student paper in 1990, and the INFORMS College on Simulation Publication Award in 1996. Versions of these and later ideas were implemented in IBM's SAVE availability modeling package, providing a substantial practical impact on system design applications.

GOLDSMAN, NELSON, AND SCHMEISER (1991) REVIEW BY P. L'ECUYER

Simulation is often presented as a way to estimate an unknown number (or vector of numbers) expressed as the mathematical expectation of a random variable that we know how to simulate, or the ratio of two such expectations, or something similar. These expectations are estimated by generating several copies of the random variable and taking the average; this is the Monte Carlo method. But the ultimate goal of most simulation projects is decision making. We may want to choose the "best" of a few systems or configurations, or we may want to optimize some discrete or continuous parameters of the system, or even try to optimize some complex decision policies to control the system, where each decision may depend on the entire state of the system. These problems are quite difficult in general, but they are extremely important from the practical viewpoint.

Goldsman, Nelson, and Schmeiser (1991) consider the first type of situation, where one wishes to select the best system among a few alternatives, or at least a system whose performance is close enough to the best, assuming that the performance is defined by a mathematical expectation that can only be estimated by simulation. The easiest approach one could think of in this context is to simulate each alternative a large number of times, large enough to estimate the performance with negligible statistical error, and then select the winner. But this brute force approach is often much too inefficient. Better methods save work by simulating each system just enough to make a good decision with high enough confidence. The actual definitions of *good decision* and *high confidence* may depend on the method and on the user.

In their paper, Goldsman, Nelson, and Schmeiser (1991) discuss three different methods for deciding which system is best. Each of the three authors explains and argues for one of these methods: Schmeiser is a proponent of *interactive analysis* (IA), Goldsman is for *ranking and selection* (RS) procedures, and Nelson defends *multiple comparison* (MC) procedures. The authors have agreed on a small example in which one must choose between four configurations of an airline reservation system. The performance measure (to be maximized) is the expected time to failure (ETTF). Important characteristics of this example are that simulation runs are not cheap (about 30 seconds of CPU time, on the computers used at that time), and that the standard Monte Carlo estimator has substantial relative error and is highly non-normal. The use of common random numbers across systems is also disallowed. As a consequence, one has to spend a significant amount of CPU time to obtain good confidence that the system with the empirical best performance is truly the best, or close enough to the best. Each author applied his favorite method to the selected example, and summarized in the paper a log of all the steps he made for this experiment (including pilot runs, and so on). These details give the reader a clear understanding of how each method works in a concrete situation. This makes the paper a wonderful tutorial on the main issues involved in selecting the best system, and on the comparisons and relationships between the proposed methods.

In the IA method, the analyst first makes some simulation runs to get a rough idea on the means and variances for the different configurations, and on how much time it takes to simulate the system, and then uses his judgment and common sense to determine (heuristically and informally) what to do next. The proponent of IA stays away from a rigorous statistical analysis, arguing that this is not necessary. This appears sensible for the small example considered in the paper, and the method has the advantage of being simple and flexible, but its performance in general may depend too much on the good judgment of the analyst. The other two procedures provide more specific guidance. Also, IA does not provide a well-defined measure of confidence in the retained selection.

In RS procedures, the analyst specifies an *indifference gap* $\delta > 0$ and a probability P^*, and the procedure returns, with probability at least P^* under some appropriate assumptions, a system whose performance measure is no more than δ away from the best. In general, the assumptions involve independence, normality, equality of variances, and the like. Various RS procedures have been designed that operate under different sets of assumptions. They usually have two stages: a set of pilot runs to estimate the required sample size for each system, and the production runs to collect these sample sizes and make the final decision. In practice, the RS procedures tend to be conservative, in the sense that the probability of correct decision is often larger than P^*.

The MC procedures compute simultaneous confidence intervals on the differences between the performance of each system and that of the best minus δ. When none of the interval contains zero except for the current best system, the procedure stops. The version adopted in the paper assumed equal variance, and this was achieved (approximately) via batching of observations.

The paper lists the required assumptions for each method and made insightful comparison between the methods, while giving the advantages and disadvantages of each. The detailed report of the process followed with each method makes the paper a nice tutorial on how to apply these methods in practice. This paper played a key role in teaching the simulation community about the methods for selecting the best system, and their philosophy. It contributed to raising interest in these methods and to motivate further work on this important problem, as can be appreciated, e.g., from Kim and Nelson (2006).

GLYNN (1986) REVIEW BY P. L'ECUYER

Glynn (1986) addresses a different class of optimization problems, where one wishes to optimize a continuous decision parameter θ of a regenerative stochastic process, over a bounded interval, assuming that the performance measure $\alpha(\theta)$ is a steady-state average that can be written as a ratio of two expectations. The derivative $\alpha'(\theta)$ is written as a ratio whose numerator $g(\theta)$ (say) is a function of four expectations, and the optimization is achieved by finding a root of this numerator. Two optimization algorithms are proposed. Both are *stochastic approximation* procedures, moving by small steps toward the optimum: At each step they estimate $g(\theta)$ and make a small move in the opposite direction. They differ essentially by the way they estimate $g(\theta)$. The first procedure uses finite differences; it applies to a rather general class of stochastic processes. The second procedure assumes that the process is an irreducible Markov chain whose behavior depends on θ in a smooth way. It uses a cleverly-designed unbiased estimator based on a likelihood ratio method, and using pairs of independent simulation runs. Almost sure convergence is proved for the two algorithms. The bounded interval is mapped to the entire real line by a nonlinear transformation and the algorithms actually work with the transformed parameter on the real line.

This innovative paper was the first to provide a rigorous mathematical analysis and a convergence proof for this type of optimization algorithm for regenerative systems. Later experiments have shown that the proposed algorithms, as they stand, can be very noisy in practice, due to the nonlinear transformation of the parameter and the high variance of the likelihood ratio gradient estimator, especially when the regenerative cycles are long. But the main contribution of this paper was to motivate and show the way to further research in which a whole bunch of improved algorithms were designed, rigorously analyzed, and compared empirically on various types of examples. Hundreds of research articles and several Ph.D. theses came out of this effort. The follow-up work includes the analysis of optimization algorithms based on infinitesimal perturbation analysis and on finite differences with common random numbers, for example. And the further results include not only convergence proofs, but also convergence rates and central limit theorems for those algorithms, in a variety of contexts.

REFERENCES

Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, and J. R. Wilson. 2007. Efficient computation of overlapping variance estimators for simulation. *INFORMS Journal on Computing*, 19(3), to appear.

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2006. Overlapping variance estimators for simulation. *Operations Research*, to appear. Available online via <ftp.ncsu.edu/pub/eos/pub/jwilson/ovestv72.pdf> [accessed June 17, 2007].

Banks, J., J. S. Carson, B. L. Nelson and D. M. Nicol. 2005. *Discrete-Event System Simulation*, 4th edition. Upper Saddle River, NJ: Pearson Prentice Hall.

Black, F. 1976. The pricing of commodity contracts. *Journal of Financial Economics* 3: 167–179.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–659.

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A Guide to Simulation*, 2nd ed. New York: Springer-Verlag.

Damerdji, H. 1991. Strong consistency and other properties of the spectral variance estimator. *Management Science* 37: 1424–1440.

Damerdji, H. 1994. Strong consistency of the variance estimator in steady-state simulation output analysis. *Mathematics of Operations Research* 19: 494–512.

Damerdji, H. 1995. Mean-square consistency of the variance estimator in steady-state simulation output analysis. *Operations Research*, 43: 282–291.

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. Heidelberg: Springer-Verlag.

Goldsman, D., and M. S. Meketon. 1986. A comparison of several variance estimators. Technical report J-85-12. School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Goyal, A., P. Heidelberger and P. Shahabuddin. 1987. Measure specific dynamic importance sampling for availability simulation. In *Proceedings of the 1987 Winter Simulation Conference*, A. Thesen, H. Grant and W. D. Kelton, eds., 351–357.

Hull, J. 2006. *Options, Futures, and Other Derivatives*, 6th edition, Englewood Cliffs, New Jersey: Prentice-Hall.

Juneja, S. and P. Shahabuddin. 2006. Rare-event simulation techniques: An introduction and recent advances. Chapter 11 in *Handbooks in Operations Research and Management Science*, S. G. Henderson and B. L. Nelson, eds. Amsterdam, The Netherlands: North-Holland, Vol. 13: 291–350.

Kim S.-H. and B. L. Nelson. 2006. Selecting the best system. Chapter 17 in *Handbooks in Operations Research and Management Science*, S. G. Henderson and B. L. Nelson, eds. Amsterdam, The Netherlands: North-Holland, Vol. 13: 501–534.

Law, A. M. 2007. *Simulation Modeling & Analysis*, 4th edition. New York: McGraw-Hill.

L'Ecuyer, P. 2006. Uniform random number generation. Chapter 3 in *Handbooks in Operations Research and Management Science*, S. G. Henderson and B. L. Nelson, eds. Amsterdam, The Netherlands: North-Holland, Vol. 13: 55–81.

L'Ecuyer, P., R. Simard, E. J. Chen, and W. D. Kelton. 2002. An object-oriented random-number package with many long streams and substreams. *Operations Research* 50(6): 1073–1075.

Montgomery, R. A., S. E. Gentry, W. H. Marks, D. S. Warren, J. Hiller, J. Houp, A. A. Zachary, J. K. Melancon, W. R. Maley, H. Rabb, C. Simpkins, and D. L. Segev. 2006. Domino paired kidney donation: a strategy to make best use of live non-directed donation. *Lancet* 368: 419–421.

Nayani, N., and M. Mollaghasemi. 1998. Validation and verification of the simulation model of a photolithography process in semiconductor manufacturing. In *Proceedings of the 1998 Winter Simulation Conference*, D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, eds., 1017–1022.

Nelson, B. L. 2004. Stochastic simulation research in *Management Science*. *Management Science* 50: 855–868.

Pedrosa, A. C., and B. W. Schmeiser. 1993. Asymptotic and finite-sample correlations between OBM estimators. In *Proceedings of the 1993 Winter Simulation Conference*, G. W. Evans, M. Mollaghssemi, E. C. Russell, and W. E. Biles, eds., 481–488.

Pedrosa, A. C., and B. W. Schmeiser. 1994. Estimating the variance of the sample mean: Optimal batch-size estimation and 1-2-1 overlapping batch means. Technical Report SMS94-3. School of Industrial Engineering, Purdue University, West Lafayette, IN.

Pritsker, A. A. B. 1998. Organ transplantation allocation policy analysis. *ORMS Today* 25 (4): 22–28.

Schlesinger, S. et al. 1979. Terminology for model credibility. *Simulation* 32 (3): 103–104.

Science Daily. July 30, 2006. 'Domino' transplant program makes best use of altruistic donated kidneys.

Shahabuddin, P. 1994. Importance sampling for the simulation of highly reliable Markovian systems. *Management Science* 40: 333–352.

Song, W.-M. T. 1988. Estimators of the variance of the sample mean: Quadratic forms, optimal batch sizes, and linear combinations. Ph.D. dissertation. School of Industrial Engineering, Purdue University, West Lafayette, IN.

Song. W.-M. T., and B. W. Schmeiser. 1993. Variance of the sample mean: Properties and graphs of quadratic-form estimators. *Operations Research*, 41: 501–517.

Welch, P. D. 1987. On the relationship between batch means, overlapping batch means and spectral estimation. In *Proceedings of the 1987 Winter Simulation Conference*, A. Thesen, H. Grant, W. D. Kelton, eds., 320–323.

AUTHOR BIOGRAPHIES

DAVID GOLDSMAN is a Professor in the H. Milton Stewart 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@gatech.edu⟩, and his web page is ⟨http://www.isye.gatech.edu/~sman⟩.

JAMES O. HENRIKSEN is the president of Wolverine Software Corporation. He was the chief developer of the first version of GPSS/H, of Proof Animation, and of SLX. He is a frequent contributor to the literature on simulation and has presented many papers at the Winter Simulation Conference. Mr. Henriksen has served as the Business Chair and General Chair of past Winter Simulation Conferences. He has also served on the Board of Directors of the conference as the ACM/SIGSIM representative. He was named a Titan of Simulation at the 2006 Winter Simulation Conference. His e-mail address is ⟨mail@wolverinesoftware.com⟩

PIERRE L'ECUYER is Professor in the Département d'Informatique et de Recherche Opérationnelle, at the Université de Montréal, Canada. He holds the Canada Research Chair in Stochastic Simulation and Optimization. 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 Associate/Area Editor for *ACM Transactions on Modeling and Computer Simulation*, *ACM Transactions on Mathematical Software*, *Statistical Computing*, *International Transactions in Operational Research*, *The Open Applied Mathematics Journal*, and *Cryptography and Communications*. He obtained the *E. W. R. Steacie* fellowship in 1995-97, a *Killam* fellowship in 2001-03, and became an INFORMS Fellow in 2006. His recent research articles are available on-line from his web page: ⟨http://www.iro.umontreal.ca/~lecuyer⟩.

BARRY L. NELSON is the Charles Deering McCormick Professor in the Department of Industrial Engineering & Management Sciences at Northwestern University. His research interests are the design and analysis of stochastic simulation experiments, particularly issues of multivariate input modeling, optimization via simulation and metamodeling. He is currently Editor in Chief of *Naval Research Logistics*, and a member of the Board of Directors of the Winter Simulation Conference, IIE (senior member), INFORMS (Fellow), ACM and ASA.

DAVID H. WITHERS is retired following a technical management career with The U.S. Coast Guard, IBM, LexisNexis, and Dell Computer. He received a B.S. in Engineering from the U.S. Coast Guard Academy, and M.S. degrees in mathematics and computer science from Rensselear Polytechnic Institute. He has also been active in public service as a fire fighter and emergency medical technician. His publications include contributions in the Proceedings of the Winter Simulation Conference, the Journal of Computational Physics, and the IBM Journal of Research and Development. He is a member of INFORMS, and the INFORMS College on Simulation. He was General Chair for the 1997 Winter Simulation Conference. His e-mail address is ⟨davidwithers@msn.com⟩

NİLAY TANIK ARGON is an Assistant Professor in the Department of Statistics and Operations Research at the University of North Carolina at Chapel Hill. She received her Ph.D. degree in Industrial Engineering in 2002 from the Georgia Institute of Technology. Her primary research interests include the modeling and analysis of stochastic systems and statistical simulation output analysis. She is a member of INFORMS. Her e-mail address is ⟨nilay@unc.edu⟩, and her web page is ⟨http://www.unc.edu/~nilay⟩.